Hindawi
Advances in Meteorology
Volume 2020, Article ID 8385416, 17 pages
https://doi.org/10.1155/2020/8385416

Research Article

Monthly Mean Meteorological Temperature Prediction Based on VMD-DSE and Volterra Adaptive Model

Guohui Li, Wanni Chang, and Hong Yang

School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi 710121, China

Correspondence should be addressed to Guohui Li; lghcd@163.com

Received 7 October 2019; Revised 27 December 2019; Accepted 7 January 2020; Published 21 March 2020

Academic Editor: Herminia García Mozo

Climate is a complex and chaotic system, and temperature prediction is a challenging problem. Accurate temperature prediction is also concerned in the fields of energy, environment, industry, and agriculture. In order to improve the accuracy of monthly mean temperature prediction and reduce the calculation scale of hybrid prediction process, a combined prediction model based on variational mode decomposition-differential symbolic entropy (VMD-DSE) and Volterra is proposed. Firstly, the original monthly mean meteorological temperature sequence is decomposed into finite mode components by VMD. The DSE is used to analyze the complexity and reconstruct the sequences. Then, the new sequence is reconstructed in phase space. The delay time and embedding dimension are determined by the mutual information method and G-P method, respectively. On this basis, the Volterra adaptive prediction model is established to modeling and predicting each component. Finally, the final predicted values are obtained by superimposing the predicted results. The monthly mean temperature data of Xianyang and Yan'an are used to verify the prediction performance of the proposed model. The experimental results show that the VMD-DSE-Volterra model shows better performance in the prediction of monthly mean temperature compared with other benchmark models in this paper. In addition, the combined forecasting model proposed in this paper can reduce the modeling time and improve the forecasting accuracy, so it is an effective forecasting model.

1. Introduction

In the past hundred years, human beings have been in the stage of rapid development, but the global mean surface temperature has increased in this stage [1]. Climate change caused by global warming leads to frequent occurrence of extreme weather events, and the impact of meteorological disasters on human development is increasing [2], Extreme weather will lead to different degrees of impact on agriculture and economic development [3]. Temperature is considered to be one of the factors affecting humans, plants, animals, and climate change. Therefore, it is of great significance to predict the temperature accurately [4, 5].

Climate is a complex and chaotic system; Blöschl et al. [6] analyzed and discussed 23 unsolved problems in hydrological research, one of which is how to solve the uncertainty of the prediction model. The modeling method proposed in this paper provides a new way to solve this problem. The change of temperature is closely related to hydrology. It has always been a challenging problem to establish a certainty modeling method for temperature prediction. Accurate temperature prediction is also concerned in the fields of energy, environment, industry, and agriculture. At present, most of the prediction models of temperature time series are based on statistics. It has included the autoregressive model (AR), autoregressive-sliding average model (ARMA), autoregressive integrated moving average (ARIMA), autoregressive fractionally integrated moving average (ARFIMA), and seasonal autoregressive integrated moving average (SARIMA) model. AR, MA, and ARMA models are often used to predict stationary series. ARIMA is an important method to study nonstationary time series with trends. SARIMA is an important method to study nonstationary time series with trends and periodicity. In order to predict the temperature accurately, Papacharalampous et al. [7, 8] have studied the hydrological

time series prediction of temperature, precipitation, water flow, and so on, which is the largest research so far. A large number of algorithms and multiple time series are used to compare the statistical models and the application of machine learning models in hydrological time series prediction. Ma et al. [9] used the ARIMA model to analyze the trend and prediction of local environmental temperature changes in the Three Gorges reservoir area. Murty et al. [10] used SARIMA to predict the monthly mean maximum and minimum temperatures in India. Papacharalampous et al. [11] used the ARFIMA model to apply the automatic univariate time series prediction method to predict the monthly mean temperature and precipitation. The statistical model is easy to use and has a small amount of calculation. It is widely used in the prediction of temperature series. In recent years, with the development of time series prediction research, the artificial intelligence model adopts new research methods and machine learning to solve complex prediction work. Therefore, the artificial intelligence model has been applied in prediction research by more scholars. Some typical artificial intelligence models, such as artificial neural network (ANN), echo state networks (ESNs), and random forest model, have been applied to the prediction of temperature series. Tyralis and Papacharalampous [12] used large-scale temperature series data to evaluate the one-step prediction performance of random forest. Suzulmus [13] proposed to use artificial neural network to predict mean temperature. Although the neural network has good self-learning ability and nonlinearity in prediction, it also has some problems such as easy to fall into local optimum, slow convergence speed, and easy to oscillate, and the number of units in the hidden layer is difficult to determine. Huang et al. [14] proposed a recursive Bayesian state algorithm based on an echo state network model to predict temperature time series. As a new type of recurrent network, the echo state network overcomes the problem that traditional recurrent network is easy to fall into local optimum and its training algorithm is complex, but the random generation reservoir in the echo state network is not related to the specific issues and the parameters are difficult to determine. The above prediction model can effectively predict temperature. But according to research, the temperature series has nonstationary characteristics [15], and the prediction essence of chaotic time series is to reconstruct the phase space of the original dynamic system. In various nonlinear prediction models, the Volterra adaptive prediction model is applied in different fields because of its small training samples and easy programming [16, 17].

Because a single prediction model cannot achieve the best prediction performance in temperature prediction, the hybrid prediction model combined with the signal decomposition method is often used to improve the prediction accuracy. EMD proposed by Huang et al. [18] is a data-driven signal decomposition method, which has been proved to be a good time-frequency analysis tool. Cai et al. [19] proposed EMD combined with SVM to predict the monthly mean temperature in Nanjing. The intrinsic mode function (IMF) obtained by EMD can better reflect the physical characteristics of the system. However, since EMD is sensitive to noise, there is mode mixing and endpoint effects [20, 21]. Dragomiretskiy and Zosso [22] proposed variational mode decomposition (VMD) in 2014. The results show that VMD is widely used in many fields because of its strong decomposition ability, better noise robustness, and fast processing speed in signal decomposition [23, 24]. VMD is used to decompose complex signals into simple components which are easy to process, and then a prediction model is established for each component to improve the prediction accuracy. For example, Ali et al. [25] used VMD to decompose wind speed data into multiple intrinsic narrow band components to facilitate their prediction, and the experimental results verify the effectiveness of the method. Wu and Lin [26] combined VMD with wavelet decomposition to predict air quality index and improve the prediction accuracy of AQI.

A hybrid prediction model combined with mode decomposition can effectively improve prediction accuracy. However, if each mode component is modeled and predicted, the calculation scale and prediction error will increase. Shannon proposed to use entropy [27] to measure the irregularity and complexity of signals. The results show that the prediction performance can be improved by combining the mode components with approximate complexity and building prediction models for each new subsequence. Wei and Wang [28] proposed the mode component merging method based on sample entropy (SE). Although SE has powerful functions, its calculation speed is relatively slow. Cao et al. [29] proposed an IMF component merging method based on permutation entropy (PE) for ultra-short-term wind power prediction. The authors think that the sample entropy will produce the disadvantage of inaccurate estimation in the case of small samples, which is compensated by the PE. Chen et al. [30] proposed a method of mode component merging based on dispersion entropy (DE), but the DE does not take into account the influence of signal fluctuation on the mapping. Yao et al. [31, 32] proposed differential symbolic entropy (DSE). Li et al. [33] proposed that the combination of DSE and mode decomposition can improve the recognition rate of underwater acoustic signal feature extraction. DSE is a new tool to quantify the complexity of nonlinear time series. The local nonlinear dynamic information is extracted from three adjacent elements and the flexibility of nonlinear complexity detection is improved by using adjustable control parameters. Therefore, DSE is not only computationally efficient, but also can measure the complexity of short datasets.

In order to improve the accuracy of monthly mean temperature prediction and reduce the calculation scale, a combined prediction model based on variational mode decomposition-differential symbolic entropy (VMD-DSE) and Volterra is proposed. Firstly, the original temperature time series is decomposed into finite stationary mode components by the VMD method. Secondly, DSE is used to analyze the complexity of subsequences, and the mode components with similar entropy values are merged and superimposed. Then, the time delay and embedding dimension of each new sequence are calculated separately and predicted by the Volterra prediction model. Finally, the

predicted values of each component are superimposed. The experimental results show that the combined prediction model proposed in this paper improves the prediction accuracy compared with other benchmark models and can effectively predict the monthly mean temperature. In addition, the proposed model has a better reference value for hydrological time series prediction.

2. Materials and Methods

2.1. Phase Space Reconstruction of Monthly Mean Temperature Sequence. Chaotic systems extend one-dimensional time series to high-dimensional phase space by reconstructing phase space of monthly mean temperature series. The original dynamic characteristics of the time series are restored to reveal its internal regularity. In the reconstructed phase space, the time series of monthly mean temperature is set as $x_i, i = 1, 2, ...n$, and the appropriate delay time τ and embedding dimension m are selected. The phase space is reconstructed as shown in equation (1). τ is the moving value of time. The selected τ must keep the components of reconstructed phase space independent of each other. If the value of delay time is too large, any two adjacent delay coordinate points of the time series will not be related, which cannot reflect the characteristics of the whole system. If the value of delay time is too small, the data will be redundant. The embedding dimension m determines the dimension of reconstructed phase space. The embedding dimension of data series is different, and the dimension of reconstructed phase space is also different.

$$y_t = \left\{x_t, x_{t+\tau}, x_{t+2\tau}, \ldots, x_{t+(m-1)\tau}\right\}, \quad (1)$$

where $t = 1, 2, \ldots, n - (m-1)\tau$ and y_t is the phase point of m-dimensional phase space.

Packard et al. [34] suggested using delay coordinates to reconstruct phase space. Takens proved the embedding theorem [35], which shows that when the dimension of the embedding space is at least twice that of the attractor, the reconstructed phase space is topologically equivalent to the state space of the original system. Therefore, a suitable embedding dimension is found, which satisfies $m \geq 2\ d + 1$, where m is the dimension of reconstructed phase space and d is the correlation dimension of the original state space. At present, there are two methods to select the time delay: autocorrelation method [36] and mutual information method [37]. The mutual information method corresponding to the first minimum has less mutual information, and the beginning of folding can be clearly distinguished. Therefore, the dynamic characteristics of attractors can be quantitatively and qualitatively analyzed by reconstructing phase space. In this paper, the G-P method [38] is chosen to calculate the embedding dimension. The mutual information method [39] calculates delay time τ. The G-P method has been proved to be a simple and effective method to calculate the embedding dimension. In addition, this method can also be used to judge the chaos of time series. If the correlation dimension reaches saturation value with the increase of embedding dimension, it shows that the time series has chaotic characteristics. The mutual information method is a nonlinear method which is widely used to solve the time delay of phase space reconstruction. The autocorrelation method only describes the linear relationship between the two variables, which is difficult to extend to high-dimensional space, while the mutual information method describes the overall relationship between the two variables.

2.2. Variational Mode Decomposition. VMD is an adaptive signal decomposition method proposed by Dragomiretskiy and Zosso [22] in 2014. The decomposition process is essentially an iterative process of solving a special variational model. The frequency center value and bandwidth of each mode component are solved by the iterative process; the objective of the iteration process is to minimize the sum of estimated bandwidth for each mode.

In order to accurately describe the change of frequency with time, a better adaptive and intuitive instantaneous frequency analysis method is needed. VMD is a typical instantaneous frequency analysis method. The main function of the VMD method is used for stabilizing the signal. The fluctuation or change trend of the original signal under different frequencies is decomposed to produce a series of sequences with different characteristics, and each sequence is called an intrinsic mode function (IMF). The complex nonstationary signals can be decomposed into a series of simple stable signals with different frequencies by VMD. At this time, the prediction modeling of each simple component can effectively reduce the difficulty of prediction and improve the accuracy of prediction.

Firstly, before the variational model of signal decomposition is established, each IMF obtained by VMD is defined as amplitude-modulated-frequency-modulated (AM-FM) signal as shown in the following equation:

$$u_k(t) = A_k(t)\cos\left(\phi_k(t)\right), \quad (2)$$

where $u_k(t)$ is the k-th IMF component, $A_k(t)$ and $\phi_k(t)$ are the instantaneous amplitude and phase of $u_k(t)$, respectively, and the instantaneous frequency $\omega_k(t) = \mathrm{d}\phi_k(t)/\mathrm{d}t$. So, $u_k(t)$ is a harmonic signal with amplitude $A_k(t)$ and frequency $\omega_k(t)$.

On the basis of the above definition, VMD theory assumes that the input signal $x(t)$ is composed of a finite number of IMF with limited bandwidth and different frequencies. Under the constraint that the sum of each IMF component is equal to the input signal $x(t)$, the variational model of signal decomposition is constructed with the goal of minimizing the sum of estimated bandwidth of each IMF. The process of establishing the variational model is as follows [40]:

(1) Hilbert transform is applied to each mode component, and the unilateral spectrum of the mode function is obtained by Hilbert transform and construction of analytic signals.

(2) The analytic signal of each mode component is mixed with the corresponding center frequency, thereby moving the spectrum of each mode component to the baseband.

(3) Estimate the bandwidth of each mode component.

(4) Constrained variational model is constructed by introducing constraints. The concrete structure is as follows:

$$\left\{ \begin{array}{l} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \dfrac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \quad \sum\limits_{k=1}^{K} u_k(t) = f(t) \end{array} \right\}, \tag{3}$$

where $\{u_k\}$, $\{\omega_k\}$ represents the set of subsignals and their center frequencies, K represents the total number of subsignals, and $\delta(t)$ represents the Dirac distribution.

For solving variational problems, the quadratic penalty factor α and Lagrange multiplier $\theta(t)$ are introduced. In order to transform the constrained problem into a nonconstrained variational problem, the extended Lagrangian function $L(\{u_k\}, \{\omega_k\}, \{\lambda\})$ shown in equation (4) is required to be expressed as follows:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \theta) = {} & C \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ & + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 \\ & + \left\langle \theta(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle. \end{aligned} \tag{4}$$

The variational mode decomposition uses the alternate direction method of multipliers (ADMM) to solve equation (4), and u_k^{n+1}, ω_k^{n+1}, and θ^{n+1} are updated alternately, where n denotes the number of iterations. The equations are as follows:

$$\widehat{u}_k^{n+1}(\omega) = \frac{\widehat{f}(\omega) - \sum_{k=1}^{K} \widehat{u}_k(\omega) + (\widehat{\theta}(\omega)/2)}{1 + 2C(\omega - \omega_k)^2}, \tag{5}$$

$$\omega_k^{n+1} = \frac{\int_0^{\infty} \omega \left| \widehat{u}_k(\omega) \right|^2 \mathrm{d}\omega}{\int_0^{\infty} \left| \widehat{u}_k(\omega) \right|^2 \mathrm{d}\omega}, \tag{6}$$

$$\widehat{\theta}^{n+1}(\omega) = \widehat{\theta}^{n+1}(\omega) + \tau \left[\widehat{f}(\omega) - \sum_{k=1}^{K} \widehat{u}_k^{n+1}(\omega) \right]. \tag{7}$$

Given a discriminant accuracy $e > 0$, the convergence condition for stopping iteration is as follows:

$$\sum_{k=1}^{K} \frac{\left\| \widehat{u}_k^{n+1} - \widehat{u}_k^{n} \right\|_2^2}{\left\| \widehat{u}_k^{n} \right\|_2^2} < e. \tag{8}$$

Inverse process of decomposition is that the sequences with approximate entropy are classified into one category, and the new sequence is obtained by adding the subsequence of the same category. The process of decomposition and reconstruction of temperature series is described in Section 4.3.

2.3. Differential Symbolic Entropy. Shannon proposed using entropy to measure the degree of signal confusion and uncertainty [27]. For any random variable x, when the uncertainty of x increases, so does the value of entropy. Taking differences between adjacent elements into account, Yao and Wang [31, 32] proposed DSE as a tool to quantify the complexity of time series. This symbolization attributes its complexity detection to detailed local dynamic information. The calculation process of DSE is encoding the data sequence after symbolization. Then, get the final DSE by calculating the probability of each code. The flowchart is shown in Figure 1. The specific process is as follows:

(1) Process of symbolization: considering time series $X = \{x_1, x_2, \ldots, x_i, \ldots\}$, the differences between current element and its forward and backward ones are expressed as $D_1 = \|x(i) - x(i-1)\|$ and $D_2 = \|x(i+1) - x(i)\|$. The 4-symbolic differential symbolization is obtained by the following equation:

$$S_i(x_i) = \begin{cases} 0\text{: } \mathrm{diff} \geq \alpha \cdot \mathrm{var}, \\ 1\text{: } 0 \leq \mathrm{diff} < \alpha \cdot \mathrm{var}, \\ 2\text{: } -\alpha \cdot \mathrm{var} < \mathrm{diff} < 0, \\ 3\text{: } \mathrm{diff} \geq -\alpha \cdot \mathrm{var}, \end{cases} \tag{9}$$

where $\mathrm{diff} = D_1 - D_2$ and $\mathrm{var} = \sqrt{(D_1^2 + D_2^2)/2}$. Parameter α in could be adjusted from 0.3 to 0.6.

(2) Symbolic sequence coding process: construction of symbolic sequences, or words, is the next step by collecting groups of symbolics together in temporal order. This coding process is to create symbolic templates or words with finite symbolics and has some similarities to embedding theory for phase space construction. Symbolic sequence will be coded into m-bit series $C(i)$ and there are 4^m symbolics in coded series considering the 4-symbolic differential symbolization. The symbolization and coding process of DSE is shown in Figure 2. Firstly, the original sequence is symbolized according to equation (9) to form symbol series $S(i)$. Secondly, the symbol series is encoded. If the symbolic series are "$\alpha\beta\gamma$," respectively, the code series can be expressed as

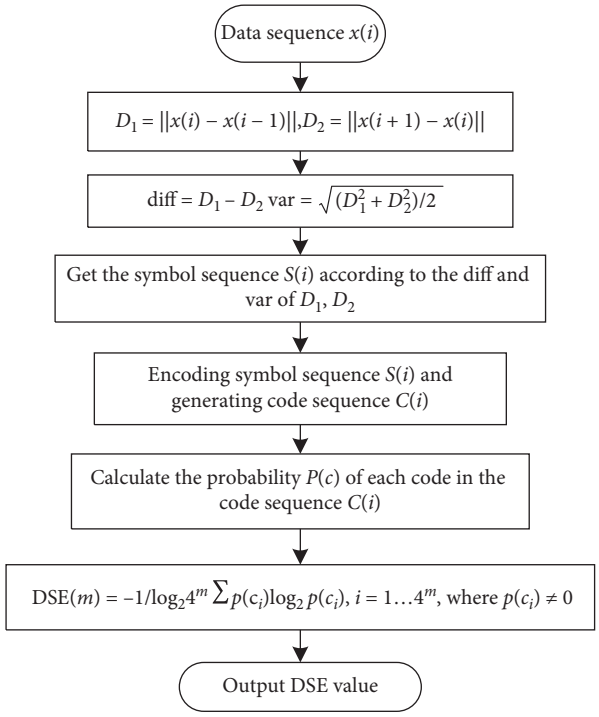

Figure 1: Process of DSE.

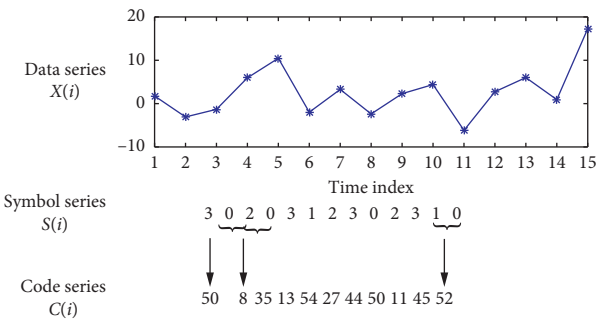

Figure 2: Process of symbolization and coding (data and symbolics in virtual frame will not be symbolized or encoded). In creation of code series, symbolic length m is 3. The first and last elements will not be transformed according to the determination of symbolization, and the last $n-1$ bit symbolics are not encoded for this encoding process.

$$c(i) = \alpha \cdot n^2 + \beta \cdot n + \gamma, \tag{10}$$

where "n" should not be less than the number of symbol types. Take Figure 2 as an example; the number of symbol types is 4, and the first three bit symbol series is 3, 0, and 2, and the corresponding code is 50 according to equation (10). After calculating the code series, the probability of calculating each code symbol is $P(c) = [p(c_1), p(c_2), \ldots, p(c_{4^m})]$.

(3) Calculate the DSE value: DSE is obtained by computing Shannon entropy from the probability distribution for all the words, as shown in equation (11). Its normalized form is $h(m) = H(m)/\log_2 4^m$.

$$\text{DSE}(m) = -\frac{1}{\log_2 4^m} \sum p(c_i)\log_2\ p(c_i), \quad i = 1, \ldots, 4^m, \text{ where } p(c_i) \neq 0. \tag{11}$$

According to Reference [31], the symbolic length m and the adjustment factor α of DSE are set to 3 and 0.6, respectively.

2.4. Volterra Adaptive Prediction Model. The Volterra model can fit any dynamic systems, without necessity to assume the model structure [41]. Suppose that the input of the nonlinear discrete dynamical system is $X(n) = [x(n), x(n-1), \ldots, x(n-N+1)]$ and the output is $\hat{x}(n+1)$. From Takens embedding theorem, it can be seen that the phase space reconstruction of chaotic time series is essentially to reconstruct the dynamics model $F(*)$ of the system through the state of the dynamic system. The equation is established as follows:

$$\hat{x}(n+1) = F(X(n)). \tag{12}$$

Then, the third-order Volterra adaptive series is expanded as follows:

$$\hat{x}(n+1) = h_0 + \sum_{m}^{N_1-1} h_1(m)x(n-m) + \sum_{m_1=0}^{N_2-1}\sum_{m_2=0}^{N_2-1} h_2(m_1, m_2), \quad x(n-m_1)x(n-m_2) + \sum_{m_1=0}^{N_3-1}\sum_{m_2=0}^{N_3-1}\sum_{m_3=0}^{N_3-1} h_3(m_1, m_2, m_3), \quad x(n-m_1)x(n-m_2)x(n-m_3). \tag{13}$$

According to equation (13), h_1, h_2, and h_3 are the Volterra cores corresponding to each order, N_1, N_2, and N_3 are the lengths of filter, and m is the optimal embedding dimension of power system. Equation (12) is extended to linear state form, as shown in the following equation:

$$\hat{x}(n+1) = H^T(n)U(n), \tag{14}$$

where $\hat{x}(n+1)$ is the predicted output value of the signal, $U(n)$ is the effective input signal vector of the filter, $H^T(n)$ is the signal core coefficient vector, and $U(n)$ and $H^T(n)$ are expressed as follows:

$$U_{(n)} = [1, x(n), x(n-1), \ldots, x(n-m+1), x^2(n), x(n)x(n-1), \ldots, x^2(n-m+1)]^T, \tag{15}$$

$$H_{(n)} = [h_0, h_1(0), h_1(1), \ldots, h_1(m-1), h_2(0,0), h_2(0,1), \ldots, h_2(m-1, m-1)]^T. \tag{16}$$

For Volterra adaptive filter, the normalized least mean square algorithm with better adaptive performance is used to

process the filter. The filter vector of the filter coefficient is $H^T(n)$ and the input signal effective vector is $U(n)$. The least mean square algorithm is described as follows:

$$e(n) = \hat{x}(n+1) - H^T(N)U(n), \tag{17}$$

$$U(n+1) = U(n) + \frac{\mu}{H^T(n)H(N)} e(n)H(n), \tag{18}$$

where μ is the convergence step $(0 < \mu < 2)$.

3. The Proposed Method for Monthly Mean Temperature Prediction

The flow chart of the combined prediction model of VMD-DSE-Volterra is shown in Figure 3. The concrete steps are as follows:

Step 1: VMD Process

In order to establish an effective prediction model and improve the prediction accuracy of time series. The time series $x_t\,(t = 1, 2, \ldots, N)$ of monthly mean temperature is decomposed by VMD, and a group of mode components $\mathrm{IMF}_{jt}\,(j = 1, 2, \ldots, n)$ with different frequency and amplitude is obtained, where n represents the number of mode components. Each mode component is added to get the original signal $x_t = \sum_{j=1}^{n} \mathrm{IMF}_{jt}$.

Step 2: DSE Process

DSE is used to calculate the entropy value of IMF1, IMF2... IMFn and analyze the complexity. Then, the IMF with similar entropy value is added to get the new subsequence $\mathrm{DSE-IMF}_{it}\,(i = 1, 2, \ldots, s)$, where s is the number of mode components after merging. The purpose is to solve the problem of overdecomposition and computing burden. The decomposition and recombination process of VMD-DSE is shown in Section 4.3.

Step 3: Volterra Predicted Process

(1) The monthly mean temperature historical data are selected as training data, and the training data are reconstructed in phase space to obtain the predicted input vector.

(2) Initialize the coefficient vector of the Volterra model and train the coefficient vector by the time orthogonal adaptive algorithm until convergence is reached.

(3) The phase space contained at a moment at the point to be measured is input into the trained Volterra filter, and the prediction results are obtained.

Step 4: Hybrid Process

$\hat{\mathrm{DSE-IMF}}_{it}\,(i = 1, 2, \ldots, s)$ is the prediction result of $\mathrm{DSE-IMF}_{it}\,(i = 1, 2, \ldots, s)$, which will be superposed to output the final prediction result $\hat{x}_t = \sum_{i=1}^{s} \hat{\mathrm{DSE-IMF}}_{it}$. $\hat{x}_t$ represents the predicted value of the original data.

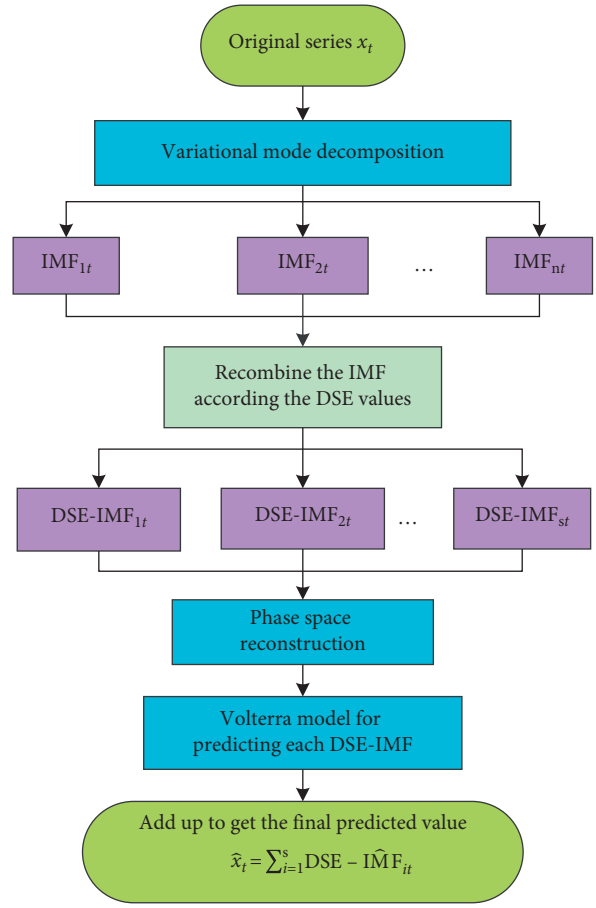

Figure 3: The flowchart of the VMD-DSE-Volterra model.

4. Results and Discussion

4.1. Data Sources and Simulation Software. Shaanxi Province is located in the northwest of China, spanning the north temperate zone and subtropical zone. It has a continental monsoon climate as a whole. Xianyang city is located in the Guanzhong Basin of Shaanxi Province, and the season of high temperature is also the season of rainfall. The annual mean temperature ranges from 9.0 to 13.2°C, the annual minimum temperature is −18.6°C, and the annual maximum temperature is 41.2°C. Yan'an city is located in the southern half of northern Shaanxi Province. Its landform is dominated by Loess Plateau and hills. It belongs to the warm temperate zone with semihumid and drought-prone climate. The annual climate change is subject to monsoon circulation. The annual minimum temperature is −25.4°C, and the annual maximum temperature is 39.7°C.

This paper chooses the monthly mean temperature data of Xianyang city from 1958 to 2017 for 60 years and Yan'an City from 1951 to 2017 for 67 years. The data come from China Meteorological Data Network (http://data.cma.cn/). The time series diagram of the monthly mean temperature in Xianyang and Yan'an is shown in Figure 4. In addition, the software environment of this paper is Matlab 2014b. The analyses and visualizations have been performed in Matlab Programming Language. We have used the following

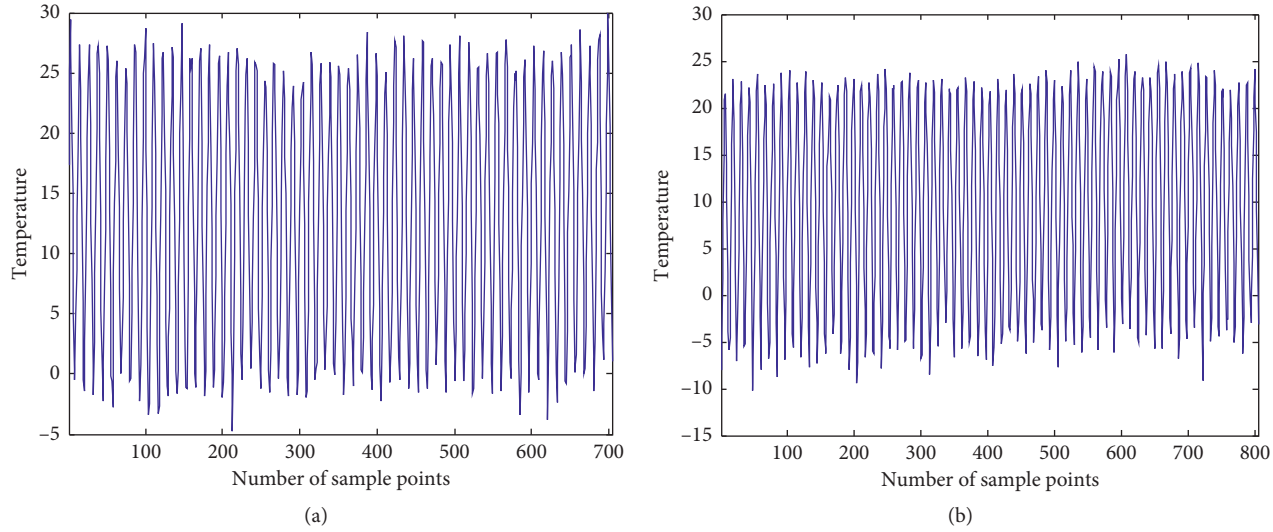

Figure 4: Temperature time series: (a) Xianyang area; (b) Yan'an area.

contributed Matlab toolbox: chaotic time series analysis and prediction Matlab toolbox—trial version 2.9 [42] and signal processing toolbox [43]. In the prediction part of the SARIMA model, SPSS 25 [44] software is used to process the data.

4.2. Phase Space Reconstruction of Monthly Mean Temperature. Taking the temperature series in Xianyang area as an example, this paper selects the mutual information method to calculate the delay time, and the correlation dimension is determined by the G-P method. Figure 5(a) shows the delay time of monthly mean temperature in Xianyang. When the delay time is 2, the mutual information quantity reaches the minimum for the first time. After this value, the trend of mutual information is nearly stable, so the time delay is 2. Figure 5(b) shows the relationship curve between correlation dimension and embedding dimension. The variation curve of correlation dimension determined by the G-P method is shown in Figure 5(c). It can be seen from 5(b) that with the increase of embedding dimension m, the D-M curve is saturated. When m continues to increase, the correlation dimension d converges to about 2.756. According to $m \geq 2\ d + 1$, the embedding dimension m of the monthly mean temperature time series in Xianyang area is determined to be 7. At the same time, the d value of correlation dimension in Figure 5(b) is saturated, which indicates that there is chaos in the temperature series in the Xianyang area.

4.3. VMD-DSE Processing. VMD needs to set the mode number K before decomposition; when the K value is too large, it will cause additional noise or cause mode repetition. When the K value is too small, the mode under-decomposition will be caused. So, the selection of the appropriate mode number K is very important for the decomposition results. Because the central frequency of each mode number is different, the central frequency method is used to determine the mode number K in this paper, and the corresponding mode central frequency of each K value is tested. If the central frequency value approaches the mode value, it will be regarded as overdecomposition [45]. After repeated experiments, the subsequent tend to approximate when $K > 6$, so $K = 6$ is chosen in this paper. Similarly, the temperature series in Yan'an area is decomposed into 7 subsequences by VMD. The decomposition results of VMD and EMD for temperature series in Xianyang and Yan'an are shown in Figure 6.

As shown in Figure 6(a), IMFs obtained from EMD of temperature series in Xianyang area have different characteristics. Because IMF1, IMF2, IMF3, and IMF4 components cannot be effectively separated in the process of decomposition, the frequency mixing phenomenon of mode components will occur. At the same time, the IMFs obtained by EMD decomposition is disturbed by noise, and there is a relatively large swing at both ends, which affects the whole component sequence. So, it will cause large errors and make the decomposition result seriously distorted. Figure 6(b) shows that the monthly mean temperature time series of Xianyang is decomposed into six IMFs by VMD. From the decomposition results, it can be seen that the VMD improves the mode mixing and waveform distortion, and the decomposed IMF is clearer, which is more conducive to improving the prediction accuracy. For temperature series of Yan'an, EMD results are shown in Figure 6(c) and VMD decomposition results are shown in Figure 6(d). Compared with EMD, VMD overcomes the problem of mode mixing and produces better decomposition effect. After VMD decomposition, seven IMFs are obtained in Yan'an area. VMD can obtain finite mode components with approximate complexity by decomposition. If the Volterra model is used to predict each subsequence directly, the computation scale will increase. In order to reduce the computational scale, DSE algorithm is used to analyze the complexity of each mode subsequence. The subsequences of DSE approximation are

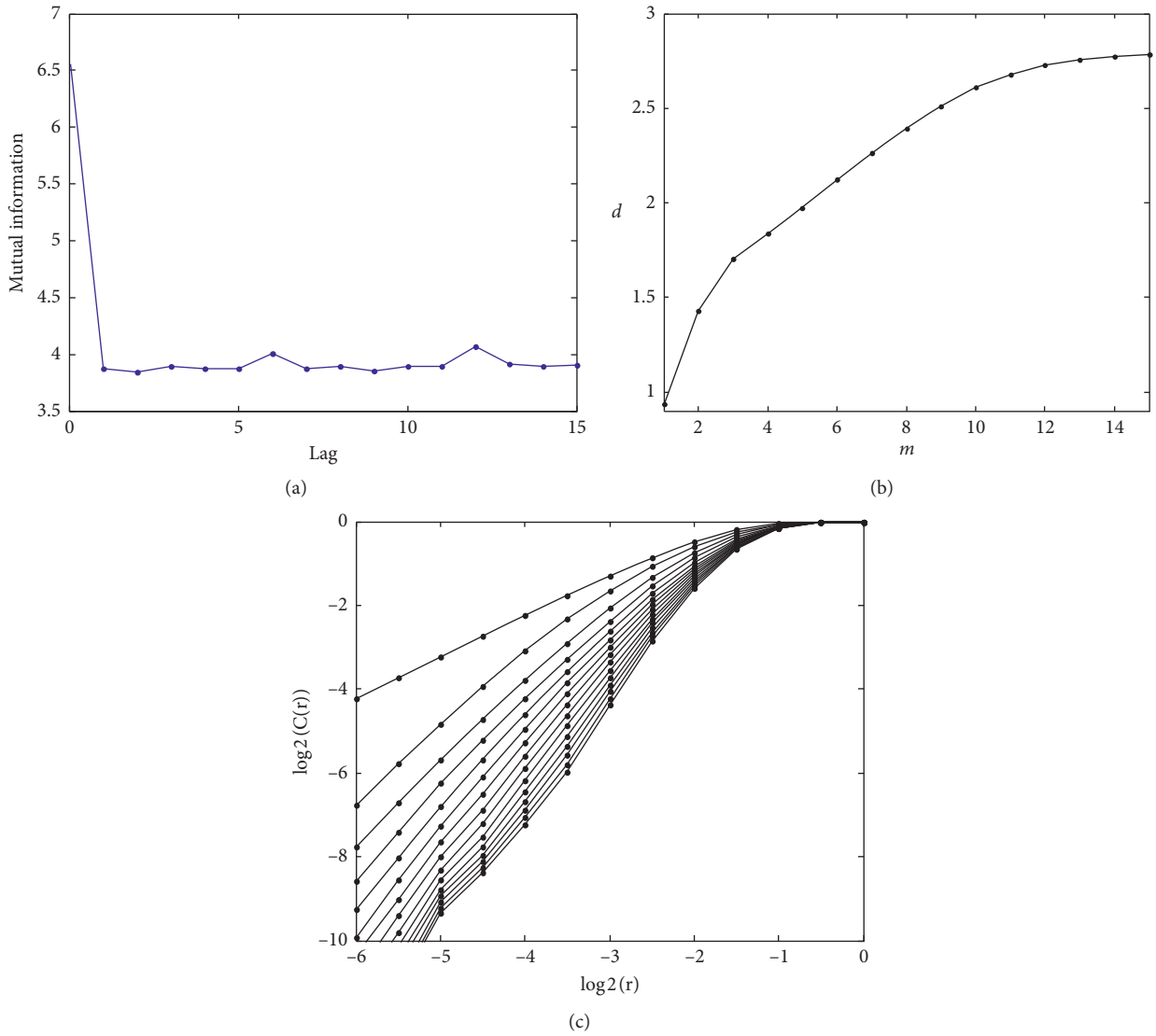

Figure 5: (a) Delay time; (b) D-m curve; (c) $\log_2 c(r) - \log_2 r$ curve.

merged and superposed to reduce mode components. The DSE calculation results of each IMF component in Xianyang area are shown in Table 1.

From Table 1, the monthly mean temperature in the Xianyang area is divided into six mode components, of which the DSE difference between IMF 2 and IMF 4 is 0.0303, and IMFs 2 and 4 are merged into a new subsequence as input to Volterra. For the same reason, the value of DSE between IMF 3 and IMF 6 is 0.0375, so IMF 3 and IMF 6 can be merged into a new mode component. The DSE value of IMF1 and IMF5 is obviously different from other modes, so it can be used as a new subsequence. The temperature series of Xianyang after VMD-DSE decomposition and reconstruction is shown in Figure 7(a).

The DSE values of each component after VMD decomposition of temperature series in Yan'an area are shown in Table 2. It can be seen from Table 2 that the DSE values of IMF2 and IMF3 are close, and the DSE values of IMF5 and IMF6 are close, so we merge the mode components with close complexity. The DSE value of IMF 1, IMF4, and IMF7 is obviously different from other modes, so it can be used as a new subsequence. The temperature series of Yan'an after VMD-DSE decomposition and reconstruction is shown in Figure 7(b).

4.4. Performance Indicators of Prediction Accuracy. In this paper, three error indicators are selected to measure the prediction effect of the proposed prediction model: mean absolute error (MAE), root mean squared error (RMSE), and correlation coefficient (R). MAE and RMSE are used to quantify the error of the predicted value; the smaller the numerical value, the better the prediction accuracy. R represents the correlation between real data and prediction data. The closer R is to 1, the better the prediction performance of the model. Equations are as follows:

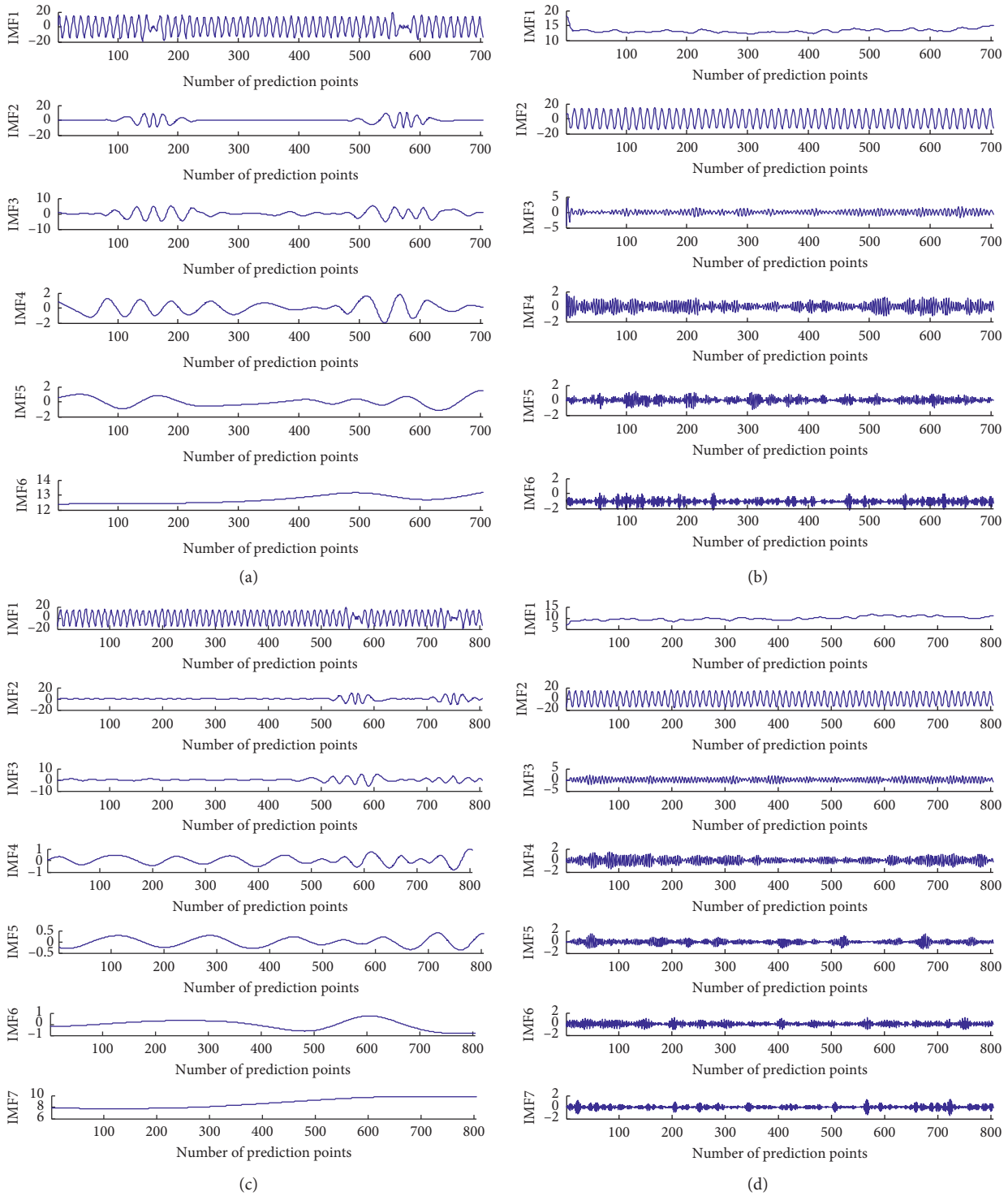

Figure 6: (a) The decomposition result of EMD in Xianyang area. (b) The decomposition result of VMD in Xianyang area. (c) The decomposition result of EMD in Yan'an area. (d) The decomposition result of VMD in Yan'an area.

Table 1: Recombined by the DSE value of VMD subsequence in Xianyang area.

Sub	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6
DSE value	0.4853	0.5983	0.7291	0.6286	0.8008	0.6916
Restructure	IMF1	IMF2&IMF4	IMF3&IMF6		IMF5	
New sub	DSE-IMF1	DSE-IMF2	DSE-IMF3		DSE-IMF4	

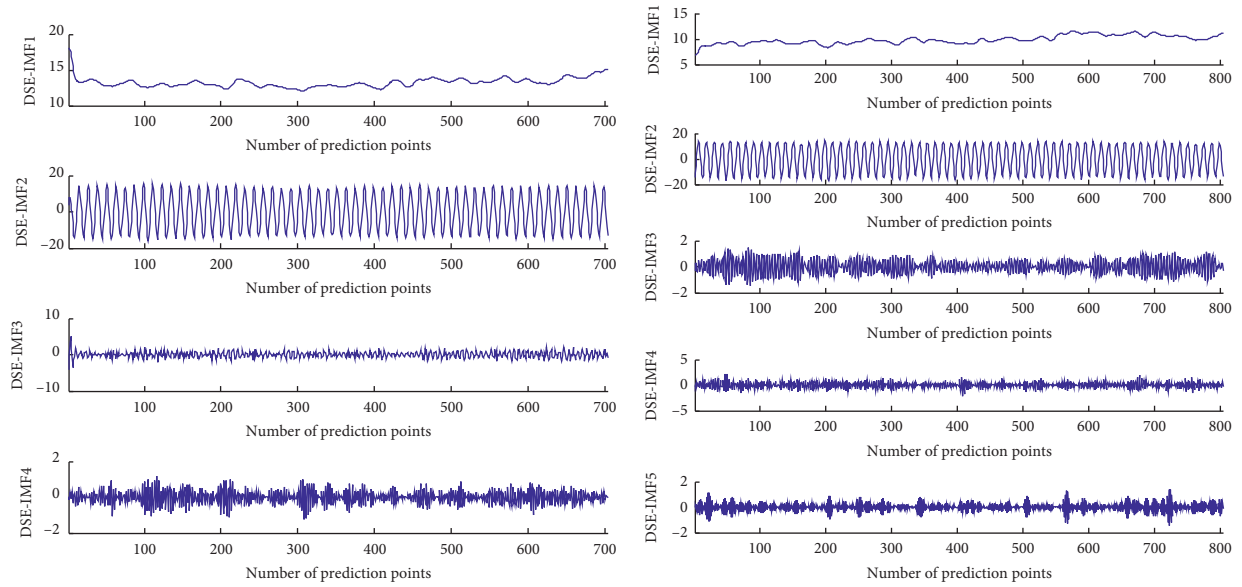

Figure 7: (a) VMD-DSE processing results of Xianyang temperature time series and (b) VMD-DSE processing results of Yan'an temperature time series.

Table 2: Recombined by the DSE value of VMD subsequence in Yan'an area.

Sub	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
DSE value	0.6032	0.7293	0.7496	0.7090	0.8347	0.8484	0.7832
Restructure	IMF1	IMF2&IMF3			IMF5&IMF6		
New sub	DSE-IMF1	DSE-IMF2		DSE-IMF3	DSE-IMF4		DSE-IMF5

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}|\hat{x}(t) - x(t)|, \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}[\hat{x}(t) - x(t)]^2}, \tag{20}$$

$$R = \frac{\sum_{t=1}^{N}\left(x(t) - x(t)'\right)\left(\hat{x}(t) - \hat{x}(t)'\right)}{\sqrt{\sum_{t=1}^{N}\left(x(t) - x(t)'\right)^2} - \sqrt{\sum_{n=1}^{N}\left(\hat{x}(t) - \hat{x}(t)'\right)^2}}, \tag{21}$$

where $\hat{x}(t)$ represents predictive data, $x(t)$ represents raw data, $\hat{x}(t)'$ represents mean value of predictive data, $x(t)'$ represents mean value of raw data, and N represents the length of data.

4.5. Model Prediction and Result Analysis in Xianyang. In order to verify the effectiveness of the method described in this paper, a total of 704 data of the monthly mean temperature in Xianyang from 1959 to 2017 are decomposed and merged into four subsequences by VMD-DSE, and then the third-order Volterra adaptive model is used to train and predict each component. The first 647 data are trained and the last 57 data are predicted.

According to the above settings, the new subsequences decomposed by VMD-DSE are predicted and reconstructed by the Volterra model. The prediction results of each component are shown in Figure 8(a). The red line in the picture represents the real value of the monthly mean temperature in Xianyang, and the blue line represents the predicted value of the combined forecasting model. From the forecasting effect of each component in Figure 8(a), the fitting degree of DSE-IMF2 to the original data is obviously better than others. When reconstructing subsequences, it can be seen from Table 1 that the DSE of IMF3, IMF5, and IMF6 is larger than IMF2 and IMF4. With the increase of DSE, the complexity of components increases. Because the complexity determines the difficulty of prediction, the prediction error will also increase. It can be seen in Table 1 that the new sequence DSE-IMF2 is generated by the merging of the original sequence IMF2 and IMF6, so the prediction error of the DSE-IMF2 is lower than others.

The final prediction result of VMD-DSE-Volterra is obtained by adding the prediction result of each component as shown in Figure 8(b). It can be seen that the real value represented by red line and the predicted value represented by blue line have a better fitting degree, which shows good prediction effect. It shows that the VMD-DSE-Volterra model has high prediction accuracy. In order to further measure the forecasting effect of this method, it is also compared with VMD-Volterra, VMD-DE-Volterra, EMD-

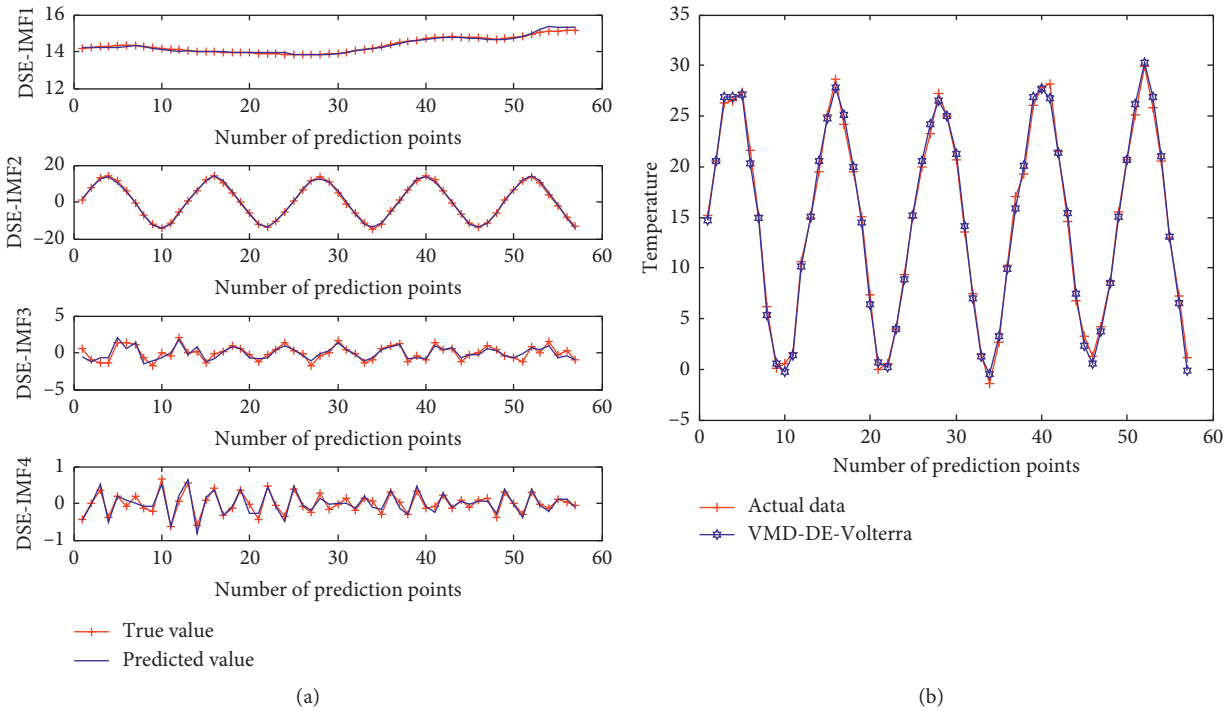

Figure 8: (a) Reconstructing subsequence prediction result; (b) VMD-DSE-Volterra prediction results.

Volterra, and other single forecasting methods such as Volterra and RBF. The results of different prediction methods are shown in Figure 9, and the local enlargement is shown in Figure 10. It is clear from the figures that the VMD-DSE-Volterra curve of blue is the closest to the original data curve of red. In other words, VMD-DSE-Volterra shows a better prediction effect. At the same time, Figure 11 and Table 3 list more intuitively the prediction error indicators MAE and RMSE of different prediction models.

Figures 9–11 and Table 3 verify that the combined model proposed in this paper has better prediction accuracy and error than other models. At the same time, it can be seen from Table 4 that VMD-DE-Volterra and VMD-DSE-Volterra have shorter running time than VMD-Volterra. It is shown that the model proposed in this paper can not only reduce the prediction error but also effectively reduce the operation scale. The RBF model and SARIMA model are used as benchmark models and compared with the Volterra model. The RBF model is a neural network with local approximation, and SARIMA is a seasonal ARIMA. According to the comparison of R-square and BIC criterion, the period of SARIMA is 12, and SARIMA $(1, 0, 1)\ (0, 1, 1)_{12}$ model is established to predict the monthly mean temperature. From the results, the prediction performance of the Volterra model and SARIMA model in Xianyang area is better than that of the RBF model, but the Volterra model has the best prediction effect. RMSE represents the mean square error between predicted and actual values relative to actual values. Therefore, when the index is smaller, the result is better. From Table 3, the prediction accuracy of the RBF model is less than that of the Volterra model. We can see that the RMSE value of VMD-Volterra is 1.0485 while the corresponding value of single prediction model Volterra is 2.1188. In addition, the MAE values of VMD-Volterra are also much smaller than those of the single model. It shows that the prediction model with the decomposition method can greatly improve the accuracy of prediction and reduce the prediction error.

From Table 3, it can be seen that adding VMD decomposition and EMD decomposition can improve the prediction accuracy. However, from the results of Table 3, it can be seen that the values of RMSE and MAE of VMD-Volterra are 1.0485 and 0.8181, respectively, while those of EMD-Volterra are 1.4036 and 1.1556, after EMD decomposition, the error exponent of the model is larger than that of VMD decomposition. The reason is that EMD has the mode mixing and endpoint effect in decomposition, so the decomposition effect is obviously worse than VMD and the prediction error is relatively higher. As a new tool to quantify the complexity of nonlinear time series, DSE greatly improves the flexibility of the detection of nonlinear complexity. At the same time, in order to verify the advantages of DSE, VMD-DSE-Volterra and VMD-DE-Volterra are compared on the basis of the same data preprocessing. The RMSE and MAE indices of VMD-DSE-Volterra are 0.5684 and 0.6703 and VMD-DE-Volterra are 0.7588 and 0.9348. According to the index value, the complexity of nonlinear signal detection by DSE is more flexible, which is beneficial to improve the sensitivity and accuracy of prediction.

In summary, by comparing the prediction performance of the other six benchmark models, it shows that the hybrid model proposed in this paper can effectively improve the prediction accuracy. In addition, the choice of the decomposition method and appropriate prediction model is very

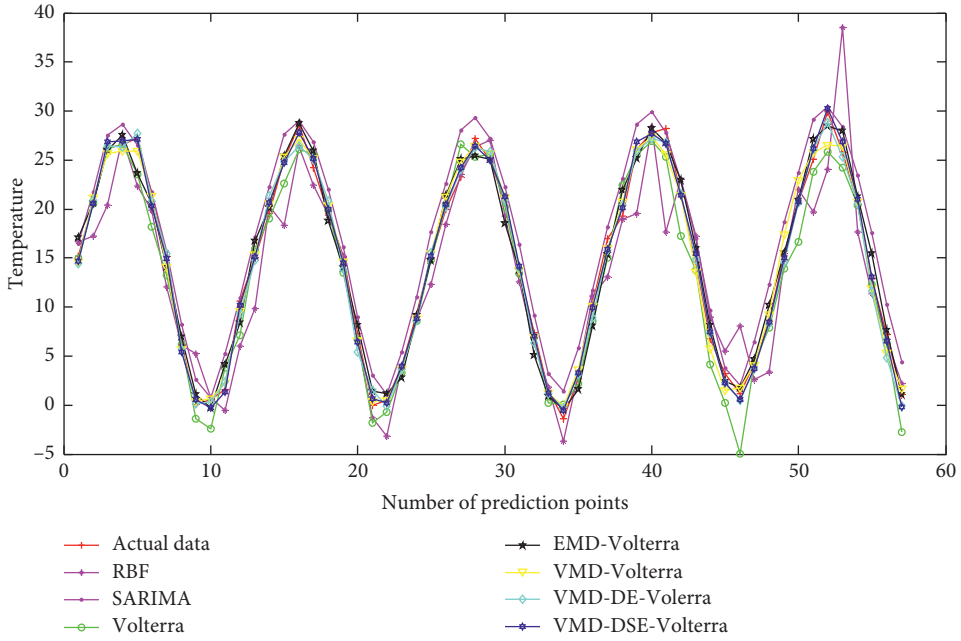

Figure 9: Prediction results for each model in Xianyang.

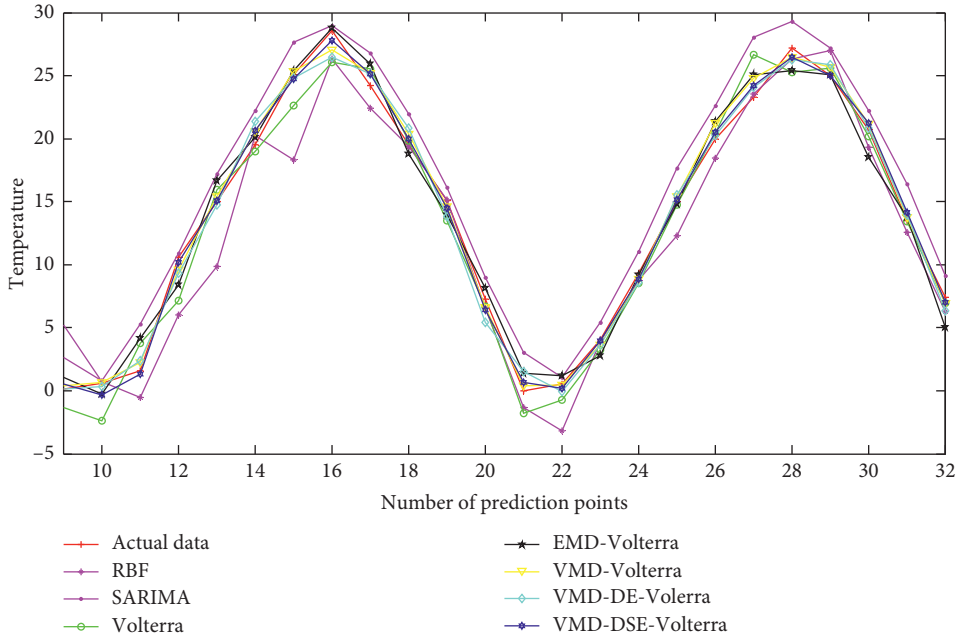

Figure 10: Local enlargement of prediction results for each model.

important for prediction. As can be seen from Table 3, the model proposed in this paper has the smallest error. In addition, it can be seen from Table 4 that the running time of VMD-DSE-Volterra is less than that of VMD-Volterra. It shows that the combined forecasting model proposed in this paper can reduce the modeling time and improve the forecasting accuracy, so it is an effective forecasting model.

4.6. Model Prediction and Result Analysis in Yan'an. In order to further illustrate the universality and validity of the VMD-DSE-Volterra hybrid prediction model in monthly mean temperature prediction, another case is obtained by using 804 monthly mean temperature data from 1951 to 2017 in Yan'an area. In this paper, the first 736 data are used for training, and the last 68 data are used for forecasting. The prediction results are as follows.

Figure 12 shows the monthly mean temperature prediction curves of seven models in Yan'an area, and the local enlargement is shown in Figure 13. Table 5 shows the values of three prediction evaluation indexes MAE, RMSE, and R of each model. It can be seen more clearly in Figure 14 that the prediction indexes of the VMD-DSE-Volterra model

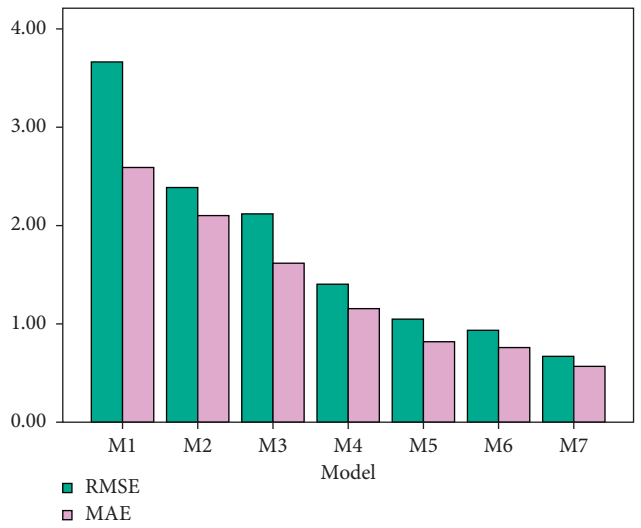

Figure 11: Bar chart of model error index. M1: RBF, M2: SARIMA, M3: Volterra, M4: EMD-Volterra, M5: VMD-Volterra, M6: VMD-DE-Volterra, and M7: VMD-DSE-Volterra.

Table 3: Error analysis of each prediction model in Xianyang area.

Model	RMSE	MAE	R
RBF	3.6643	2.5906	0.9303
SARIMA	2.3869	2.1014	0.9819
Volterra	2.1188	1.6170	0.9828
EMD-Volterra	1.4036	1.1556	0.9892
VMD-Volterra	1.0485	0.8181	0.9939
VMD-DE-Volterra	0.9348	0.7588	0.9955
VMD-DSE-Volterra	0.6703	0.5684	0.9977

Table 4: Running time of each prediction model.

Model	Volterra	EMD-Volterra	VMD-Volterra	VMD-DE-Volterra	VMD-DSE-Volterra
Time/s	0.5136	0.7110	0.7557	0.4016	0.4597

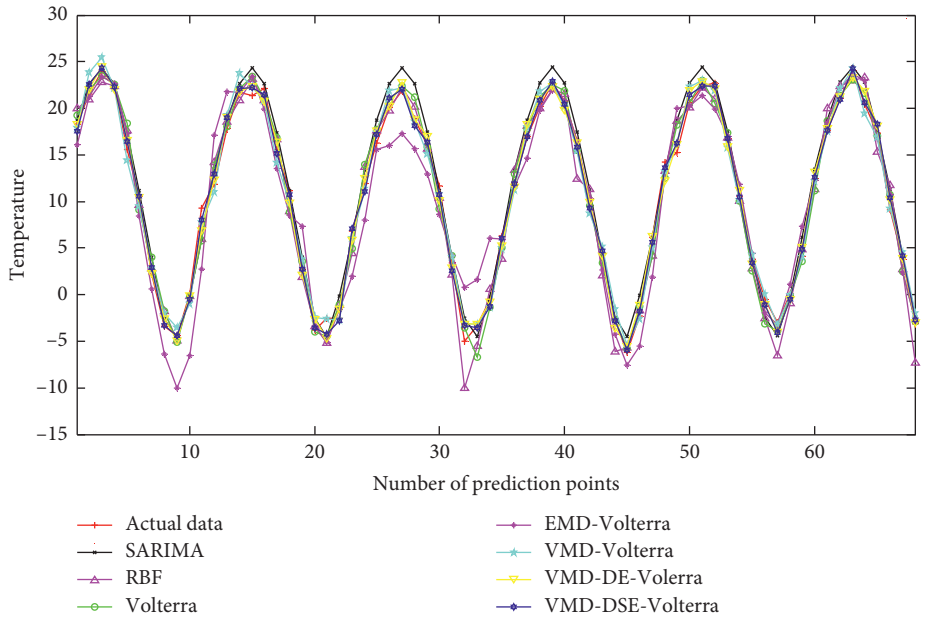

Figure 12: The prediction results for each model in Yan'an.

Table 5: Error analysis of each prediction model in Yan'an area.

Model	RMSE	MAE	R
RBF	1.8565	1.5132	0.9850
SARIMA	1.6019	1.3093	0.9882
Volterra	1.4349	1.1668	0.9895
EMD-Volterra	2.8355	2.1598	0.9575
VMD-Volterra	1.0106	0.8677	0.9942
VMD-DE-Volterra	0.9248	0.7649	0.9953
VMD-DSE-Volterra	0.6728	0.5352	0.9976

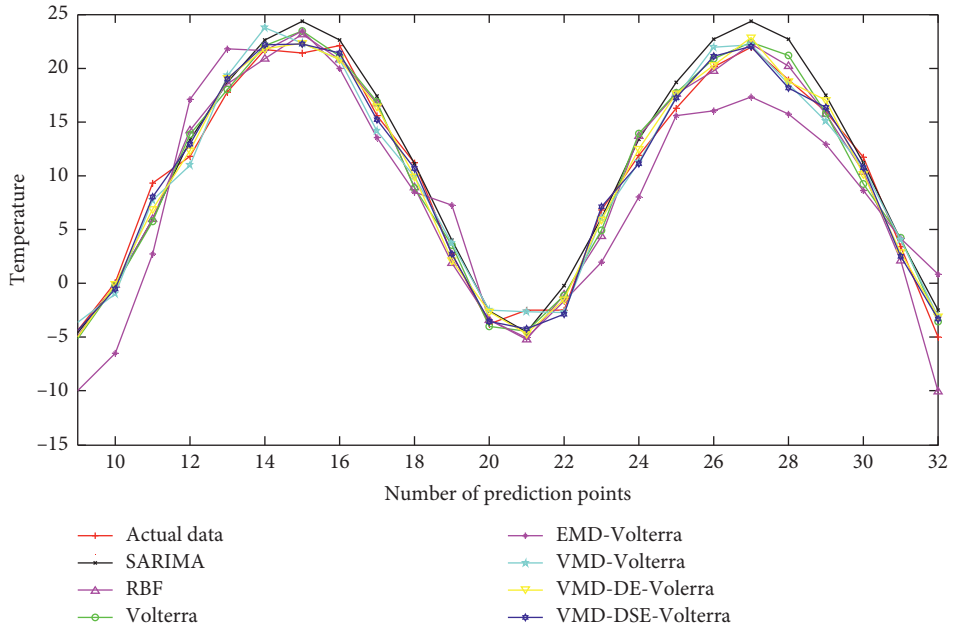

Figure 13: Local enlargement of prediction results for each model in Yan'an.

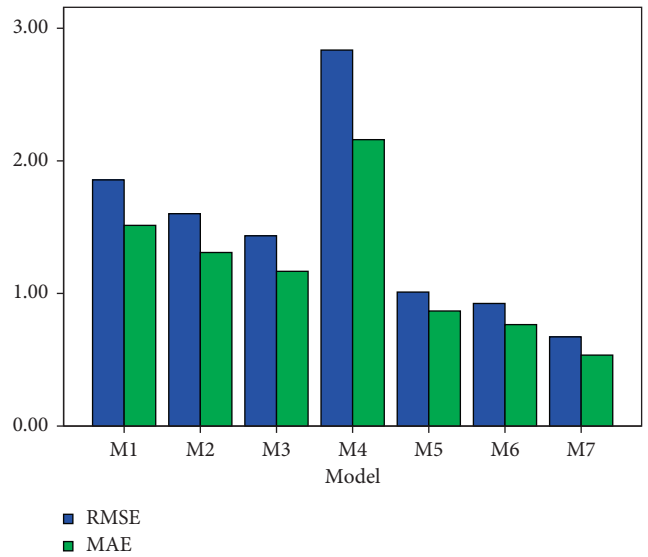

Figure 14: Bar chart of model error index. M1: RBF, M2: SARIMA, M3: Volterra, M4: EMD-Volterra, M5: VMD-Volterra, M6: VMD-DE-Volterra, and M7: VMD-DSE-Volterra.

Table 6: Running time of each prediction model.

Model	Volterra	EMD-Volterra	VMD-Volterra	VMD-DE-Volterra	VMD-DSE-Volterra
Time/s	0.4949	0.9812	0.9914	0.4071	0.6345

proposed in this paper are lowest. It can be seen in Figure 12 that the predicted value of the proposed model with the actual value has the highest fitting degree. In a word, the prediction accuracy of the hybrid model is higher than that of the single model, and the prediction effect of the model proposed in this paper is better than the other benchmark models. In addition, it can be seen from Table 6 that the operation scale can be reduced by combining components with approximate complexity. Similar to the prediction effect of monthly mean temperature in Xianyang area, the Volterra model is better than the RBF model and SARIMA model. In addition, compared with the decomposition effect of VMD and EMD, the sequence prediction error after EMD decomposition is large. It can be seen from Figure 6(c) in Section 4.3 that the temperature series in Yan'an area after EMD decomposition has mode mixing, and the position of mode mixing is in the second half of the sequence. So, the components cannot be separated effectively in the process of decomposition, which leads to a large error between the predicted value and the actual value. VMD can avoid mode mixing and produce better prediction results. From Table 5, the final prediction accuracy is basically the same as that of the Xianyang area. The best prediction accuracy model is still the VMD-DSE-Volterra model proposed in this paper. The error indexes RMSE, MAE, and R are 0.6728, 0.5352, and 0.9976, respectively.

5. Conclusion

In order to improve the prediction accuracy of monthly mean temperature, a combined prediction model based on VMD, DSE, and Volterra is proposed and applied to the prediction of monthly mean temperature in Xianyang and Yan'an area. The main conclusions are as follows:

(1) The VMD can effectively overcome the defect of mode mixing in EMD. The simulation results show that the decomposition effect of VMD is clearer and is beneficial to prediction.

(2) In this paper, DSE is used to calculate the entropy of the mode component of VMD, and the approximate components of the entropy value are merged and recombined. The simulation results of two groups of real data show that this method can effectively reduce the computational complexity and improve the prediction accuracy.

(3) This paper combines DSE with the prediction of nonlinear signals for the first time. The experimental results show that the application of DSE in the prediction of nonlinear signals can effectively improve the prediction accuracy and reduce the scale of operation.

(4) The method proposed in this paper is tested by actual monthly mean temperature data and compared with six other models, including RBF, SARIMA, Volterra, EMD-Volterra, VMD-Volterra, and VMD-DE-Volterra. The experimental results show that the combined forecasting model VMD-DSE-Volterra proposed in this paper can effectively predict the monthly mean temperature. Compared with other models, the combined forecasting model in this paper improves the forecasting accuracy and reduces the error. The modeling method proposed in this paper provides a new way to solve the uncertainty of the prediction model. In addition, the proposed model also has better reference value for hydrological time series prediction.

Data Availability

The temperature data used in this paper come from the China Meteorological Data Network (http://data.cma.cn/).

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51709228).

References

[1] L. V. Alexander, X. Zhang, and T. C. Peterson, "Global observed changes in daily climate extremes of temperature and precipitation," *Journal of Geophysical Research*, vol. 111, Article ID D05109, , 2006.

[2] T. Chen, T. Q. Ao, X. Zhang, X. D. Li, and K. B. Yang, "Climate change characteristics of extreme temperature in the Minjiang river basin," *Advances in Meteorology*, vol. 2019, Article ID 1935719, 15 pages, 2019.

[3] L. Gao, J. Chen, J. W. Zheng, and Q. L. Chen, "Progress in researches on ensemble forecasting of extreme weather based on numerical models," *Advances in Earth Science*, vol. 34, no. 7, pp. 706–716, 2019.

[4] C. X. Liu, H. Q. Tang, Y. Y. Yin, and C. X. Xu, "Regionalization of integrated environmental risk of China under future climate change," *Scientia Geographica Sinica*, vol. 38, no. 4, pp. 636–644, 2018.

[5] O. Kisi and J. Shiri, "Prediction of long-term monthly air temperature using geographical inputs," *International Journal of Climatology*, vol. 34, no. 1, pp. 179–186, 2014.

[6] G. Blöschl, M. FP. Bierkens, A. Chambel et al., "Twenty–three unsolved problems in hydrology (UPH)–a community perspective," *Hydrological Sciences Journal*, vol. 64, no. 10, pp. 1141–1158, 2019.

[7] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "One–step ahead forecasting of geophysical processes within a purely statistical framework," *Geoscience Letters*, vol. 5, no. 12, pp. 1–19, 2018.

[8] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes," *Stochastic Environmental Research and Risk Assessment*, vol. 33, no. 2, pp. 481–514, 2019.

[9] L. Ma, R. L. Yu, J. Li, and Z. L. Xia, "Research on local environmental temperature change and forecast in the three gorges reservoir area based on time series analysis," *Chinese Journal of Agrometeorology*, vol. 39, no. 1, pp. 9–17, 2018.

[10] K. V. Murthy, R. Saravana, and K. V. Kumar, "Stochastic modelling of the monthly average maximum and minimum temperature patterns in India 1981–2015," *Meteorology and Atmospheric Physics*, vol. 131, no. 4, pp. 775–787, 2019.

[11] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "Predictability of monthly temperature and precipitation using automatic time series forecasting methods," *Acta Geophysica*, vol. 66, no. 4, pp. 807–831, 2018.

[12] H. Tyralis and G. Papacharalampous, "Variable eelection in time series forecasting using random forests," *Algorithms*, vol. 10, no. 4, p. 114, 2017.

[13] S. Suzulmus, "Prdiction of average temperatures using artificial neural netwark methods: the case of Gaziantep Provnice, Turkey," *Fresenius Environmental Bulletin*, vol. 28, no. 2A, pp. 1494–1502, 2019.

[14] B. Huang, G. Qin, R. Zhao, Q. Wu, and A. Shahriari, "Recursive Bayesian echo state network with an adaptive inflation factor for temperature prediction," *Neural Computing and Applications*, vol. 29, no. 12, pp. 1535–1543, 2016.

[15] Y. Luo, J. WU, and X. J. Xu, "Application of non–stationary time series analysis in dynamic change of average temperature in Hefei city," *Journal of Anhui Agricultural University*, vol. 40, no. 6, pp. 1059–1062, 2013.

[16] Z. Y. Wang, C. X. Tian, Q. B. Zhu, and M. Huang, "Hourly solar radiation forecasting using a volterra least squares support vector machine model combined with signal decomposition," *Energies*, vol. 11, no. 1, Article ID en11010068, 2018.

[17] F. M. Huang and Y. G. Tian, "WA–Volterra coupling model based on chaos theory for monthly precipitation forecasting," *Earth Science*, vol. 39, no. 3, pp. 368–374, 2014.

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

[19] J. H. Cai, S. J. Zhang, and L. Yang, "Nanjing monthly average temperature prediction base on empirical mode decomposition and support vector machine," *Mathematics in Practice and Theory*, vol. 44, no. 22, pp. 103–111, 2014.

[20] Y. H. Jiang, W. D. Jiao, R. Q. Li et al., "A study on the method for eliminating mode mixing in B–spline empirical mode decomposition based on adaptive bandwidth constrained signal," *Journal of Vibration and Shock*, vol. 37, no. 16, pp. 83–90, 2018.

[21] Z. P. He, Z. Q. Zhu, H. C. Xie et al., "Restrain boundary effect of EMD based on least square fitting," *Journal of System Simulation*, vol. 30, no. 9, pp. 3377–3385+3398, 2018.

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

[23] Y. Wang, R. Markert, J. Xiang, and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 243–251, 2015.

[24] C. Aneesh, S. Kumar, P. M. Hisham, and K. P. Soman, "Performance comparison of variational mode decomposition over empirical wavelet transform for the classification of power quality disturbances using support vector machine," *Procedia Computer Science*, vol. 46, pp. 372–380, 2015.

[25] M. Ali, A. Khan, and N. U. Rehman, "Hybrid multiscale wind speed forecasting based on variational mode decomposition," *International Transactions on Electrical Energy Systems*, vol. 28, Article ID e2466, , 2017.

[26] Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," *Science of the Total Environment*, vol. 683, pp. 808–821, 2019.

[27] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[28] S. Wei and Y. Wang, "Short–term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back–propagation neural network," *Energy Conversion & Management*, vol. 157, pp. 1–12, 2018.

[29] W. Cao, L. C. Liu, Z. Q. Wang, and J. H. Li, "Research on ultra–short–term wind power prediction based on MEEMD and permutation entropy," *Renewable Energy Resources*, vol. 37, no. 3, pp. 439–444, 2019.

[30] Y. Chen, Z. Wei, and L. Cheng, "Combinatorial prediction of chaotic network traffic based on AVMD–DE and IBSA–KELM," *Computer Applications and Software*, vol. 35, no. 6, pp. 117–121, 2018.

[31] W. Yao and J. Wang, "Differential symbolic entropy in nonlinear dynamics complexity analysis," 2019, http://arxiv.org/abs/1801.08416v2.

[32] W. Yao and J. Wang, "Double symbolic joint entropy in nonlinear dynamic complexity analysis," *AIP Advances*, vol. 7, no. 7, Article ID 075313, 2017.

[33] G. Li, Z. Yang, and H. Yang, ""Feature extraction of ship–radiated noise based on regenerated phase–shifted sinusoid-assisted EMD, mutual information, and differential symbolic entropy," *Entropy*, vol. 21, no. 2, 76, 2019.

[34] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.

[35] F. Takens, D. A. Rand, and L. S. Young, "Detecting strange attractors in turbulence," *Lecture Notes in Mathematics*, vol. 898, pp. 366–381, 1981.

[36] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring, "The analysis of observed chaotic data in physical systems," *Reviews of Modern Physics*, vol. 65, no. 4, pp. 1331–1392, 1993.

[37] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.

[38] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, vol. 9, no. 1-2, pp. 189–208, 1983.

[39] X. Q. Lv, B. Cao, M. Zeng, S. S. Huang, and X. G. Liu, "An algorithm of selecting delay time in the mutual information method," *Chinese Journal of Computational Physics*, vol. 23, no. 2, pp. 184–188, 2006.

[40] G. Li, X. Ma, and H. Yang, "A hybrid model for forecasting sunspots time series based on variational mode decomposition and backpropagation neural network improved by firefly algorithm," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 3713410, 9 pages, 2018.

[41] R. D. Nowak and B. D. Van Veen, "Random and pseudo-random inputs for volterra filter identification," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2124–2135, 1994.

[42] Z. B. Lu, Z. M. Cai, and K. Y. Jiang, "Nonlinear adaptive prediction of chaotic time series based on sparse volterra filter," *Systems Engineering Electronics*, vol. 29, no. 9, pp. 1428–1431, 2007.

[43] L. Ding, Z. C. Pan, and W. Cong, "Design and simulation of digital fiter based on signal processing toolbox of MATLAB,"

Power System Protection and Control, vol. 31, no. 9, pp. 49–51, 2003.

[44] D. S. Xu, W. Yang, Z. Wei, and G. X. Liu, "Application of SPSS in characteristic of short load and its forecasting," *Power System Protection and Control*, vol. 37, no. 21, pp. 147–151, 2009.

[45] C. L. Liu, Y. J. Wu, and C. G. Zhen, "Rolling bearing fault diagnosis based on variational mode decomposition and fuzzy C means clustering," *Proceedings of the CSEE*, vol. 35, no. 13, pp. 3358–3365, 2015.