

Research Article

A Hybrid Grey Prediction Model for Small Oscillation Sequence Based on Information Decomposition

Meng Zhou ¹, Bo Zeng ², and Wenhao Zhou ¹

¹National Research Base of Intelligent Manufacturing Service, Chongqing Technology and Business University, Chongqing, China

²College of Business Planning, Chongqing Technology and Business University, Chongqing, China

Correspondence should be addressed to Bo Zeng; bozeng@ctbu.edu.cn

Received 17 April 2019; Revised 8 July 2019; Accepted 11 August 2019; Published 16 January 2020

Guest Editor: Francisco G. Montoya

Copyright © 2020 Meng Zhou 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.

Grey prediction model has good performance in solving small data problem, and has been widely used in various research fields. However, when the data show oscillation characteristic, the effect of grey prediction model performs poor. To this end, a new method was proposed to solve the problem of modelling small data oscillation sequence with grey prediction model. Based on the idea of information decomposition, the new method employed grey prediction model to capture the trend characteristic of complex system, and ARMA model was applied to describe the random oscillation characteristic of the system. Crops disaster area in China was selected as a case study and the relevant historical eight-year data published by government department were substituted to the proposed model. The modelling results of the new model were compared with those of other traditional mainstream prediction models. The results showed that the new model had evidently superior performance. It indicated that the proposed model will contribute to solve small oscillation problems and have positive significance for improving the applicability of grey prediction model.

1. Introduction

Big data technology is a computational strategy and method for processing large data sets. It is based on large data and has gradually become a research hotspot in recent years. However, sometimes it is difficult to obtain large data. Due to technological capabilities or historical reasons, there are still many small data, such as unconventional energy production, short-term traffic flow, sulfur dioxide emissions, crops disaster area, and so on [1–4]. The above problems show that there are many grey systems in the real world, and the data of these grey systems are limited. Big data technology can not effectively describe the grey system from the small data.

Grey prediction model is a useful method to study uncertain systems with partly known information and partly unknown information [5, 6]. At present, there are mainly two kinds of sequences suitable for grey prediction model, one is monotone sequence [7–10], the other is a sequence with saturated “S” shape [11–13]. For other sequences, such as oscillation sequence or fluctuation sequence, the performance of grey prediction model is poor. However, the real world is complex. The monotonic sequence and the saturated S-shaped

sequence are only two special cases, and more sequences show oscillation characteristic [14–16]. Therefore, how to reasonably construct a grey prediction model to model with oscillation sequence has become a research trend.

Currently, grey prediction model has made some achievements in modelling with oscillation sequence. These studies are mainly manifested in the following three aspects: (a) increasing smoothness of oscillating sequences: the poor smoothness of the oscillation sequence is the main reason for the poor modelling accuracy of the oscillation sequence, so smoothing the oscillation sequence becomes a way to improve the modelling accuracy. At present, sequence smoothness is mainly improved by sequence transformation, such as smoothness operator and amplitude compression [17–20]; (b) modelling oscillation interval by envelope: from the perspective of scope, the oscillation sequence envelope is modelled. The envelope is modelled by grey prediction model, and the simulation and prediction of the oscillation sequence variation range are realized [21–23]; (c) improving the structure of grey prediction model by periodic operator: in order to adapt to periodic sequence, scholars have introduced periodic factor of triangular function and have established periodic grey

prediction model to match the periodic fluctuation of sequence and reduce modelling error [24, 25].

The above methods can improve the modelling ability of grey prediction model for oscillation sequence to a certain extent, but they still have some shortcomings. The sequence transformation method destroys the characteristics of the original sequence and can not make full use of the information transmitted by the sequence. The randomness of envelope design is too large and its generalization is weak. Grey periodic prediction model not only increases the complexity of the model structure, but also only works for periodic and regular fluctuation sequences. When the sequence has oscillation characteristic, the performance of grey periodic prediction model is poor.

The oscillation sequence is composed of different scales information, such as trend, randomness, periodicity, etc. It reflects the final result of the system under the influence of various uncertainties [26–28]. A single prediction method is suitable for modelling with a single time scale sequence. It can not simultaneously simulate and predict two or more time scale information of the oscillation sequence, which ultimately can not get intended effect.

However, preprocessing complex sequence into simpler mode, has often led to satisfactory predicting results. Empirical mode decomposition (EMD) algorithm is a multi-scale analysis method. It decomposes complex oscillation sequences into a set of sub-sequences, which contain the information of the original sequence in different time scales [29]. According to the characteristics of sub-sequences, appropriate models are selected to simulate and predict the corresponding sub-sequences. Integrating the simulated and predicted values of sub-sequences will obtain the simulated and predicted value of the original sequence.

Decomposed by EMD algorithm, the small sample oscillation sequence is usually decomposed into two sub-sequences. One part is short-time trend sub-sequence. The other is one or more random fluctuation sub-sequences. GM(1,1) model is the most classical model in grey prediction model, needs only a little data (not less than 4). It excavates the trend of system through grey generation processing, and then achieves the effect of simulating and prediction. Therefore, GM(1,1) has superior performance in modelling with small trend sub-sequence. Random fluctuation subsequence is usually modelled by ARMA model. Based on the above facts, we use GM(1,1) model and ARMA model to simulate and predict sub-sequences, respectively. According to the result of decomposition, there may be other kinds of sub-sequences, but trend and fluctuation subsequences are the most common cases. Therefore, we mainly study the general situation and specifically analyse the other situations.

In this paper, a hybrid grey model for predicting small oscillation sequence is proposed based on information decomposition. In order to verify the validity of the proposed model, we select the crops disaster area in China as the modelling object, which has small oscillation characteristics. Comparing the simulation accuracy of the new model with that of the traditional ARIMA and GM(1,1) models, the result shows that the new model is obviously superior to the traditional model, which proves the validity of the new model.

The remainder of this paper is organized as follows. In section 2, the principle of empirical mode decomposition is introduced. In Section 3, the EMD-ARMA-GM(1,1) prediction

model is proposed. In Section 4, modelling condition and testing method of model errors are studied. This is followed by comparisons of the proposed model with ARIMA and GM(1,1) model, and the proposed model is used to predict crops disaster area in China. Then, conclusions are drawn in Section 6.

A chart showing the structure of this paper is given as Figure 1.

2. Empirical Mode Decomposition Principle

Empirical mode decomposition (EMD) is a method of signal decomposition, which does not depend on prior data and completely relies on the intrinsic characteristic of the data itself. After EMD adaptively decomposed the original data according to its intrinsic characteristic, the obtained Intrinsic Mode Functions (IMFs) reflect the inherent characteristic of the data [30]. IMF satisfies the following two conditions at the same time: (i) in the whole data set, the number of extrema and the number of zero-crossings must either equal or differ at most by one; (ii) at any point, the mean value of the envelope defined by local maxima and the envelop defined by the local minima is zero [31]. The operation steps of the EMD algorithm for oscillation sequence $X(t)$ are as follows [32]:

Step 1. Recognize all the maximum points and minimum points in sequence $X(t)$, and use cubic spline interpolation function to fit all the maximum points to form the upper envelope, and then fit all the minimum points to form the lower envelope, which are marked as $E_{up}(t)$ and $E_{low}(t)$, respectively.

Step 2. In each time period t , the average of upper and lower envelopes of sequence $X(t)$ is denoted as $M_1(t)$, and is calculated as

$$M_1(t) = \frac{E_{up}(t) + E_{low}(t)}{2}. \quad (1)$$

Step 3. Minus the average envelope of sequence $X(t)$:

$$H_1(t) = X(t) - M_1(t). \quad (2)$$

If sequence $H_1(t)$ has negative local maxima and positive local minima, then $H_1(t)$ is regarded as a new original sequence $X(t)$. Repeat the above process until $H_1(t)$ satisfies the two conditions of IMF. It is denoted as $C_1(t)$, where $C_1(t) = H_1(t)$, which is called the first IMF component after decomposition of the original sequence $X(t)$.

Step 4. Sequence $C_1(t)$ is separated from the original sequence $X(t)$ and the residual component is obtained, which is denoted as $R_1(t)$, that is

$$R_1(t) = X(t) - C_1(t). \quad (3)$$

Step 5. The residual component $R_1(t)$ is regarded as a new original sequence, and the “filtering” process of Step 1 is repeated until the new IMF component can not be separated. At this time, the original sequence $X(t)$ is “filtered” by EMD algorithm to get N IMFs and one residual component, where

$$X(t) = \sum_{i=1}^N C_i(t) + R_N(t). \quad (4)$$

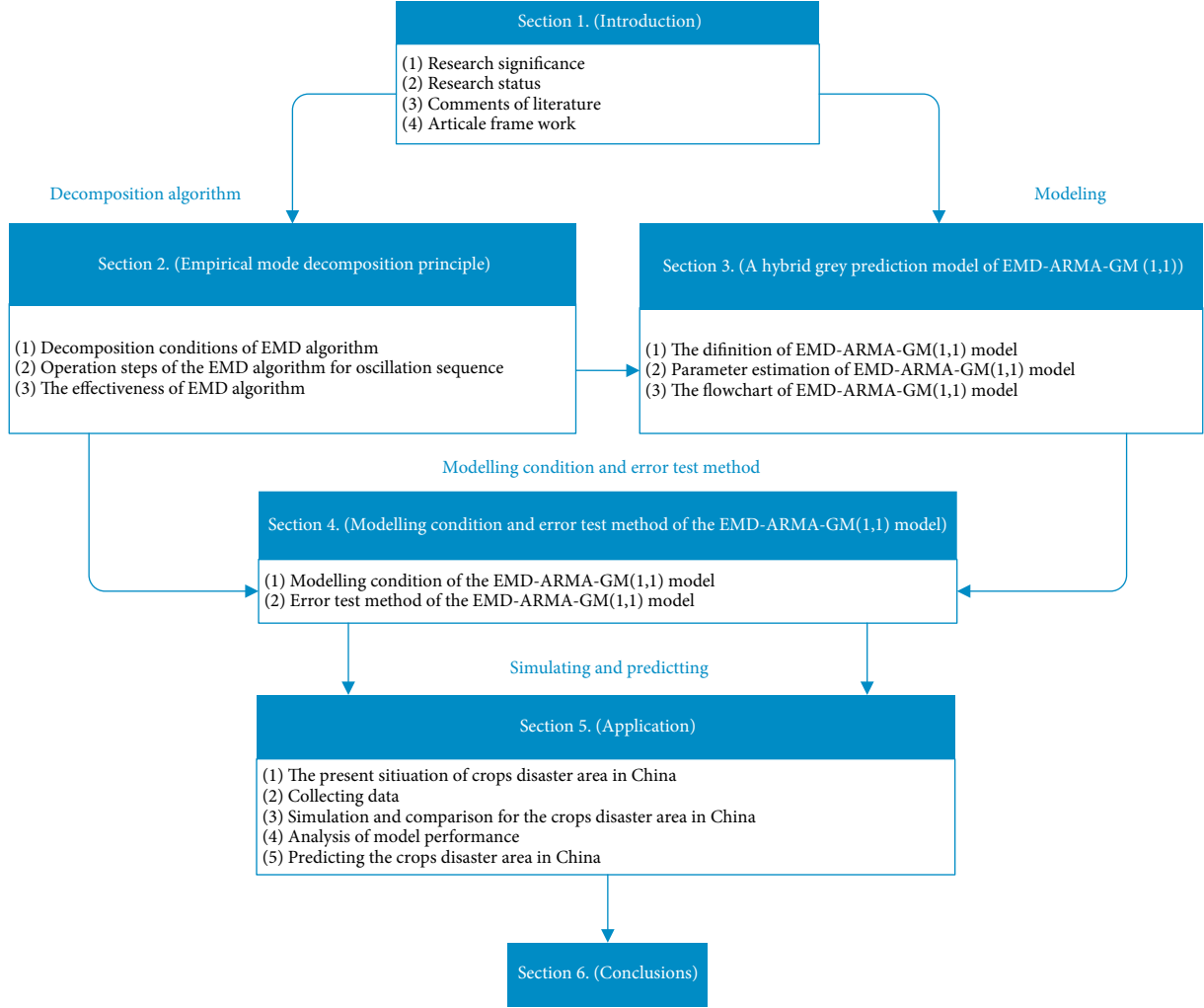


FIGURE 1: Structure of this paper.

An example of the empirical mode decomposition of an oscillating sequence is shown in Figure 2.

3. A Hybrid Grey Prediction Model of EMD-ARMA-GM(1,1)

Definition 1. Assume sequence $X = (x(1), x(2), \dots, x(n))$, $\forall k \in \{2, 3, \dots, t\}$, $\exists k' \in \{2, 3, \dots, t\}$, $x(k) - x(k-1) > 0$, $x(k') - x(k'-1) > 0$, X is called oscillation sequence.

Definition 2. Assume $C = (c(1), c(2), \dots, c(n))$ is an oscillation sequence and satisfies the following condition,

- (a) $\forall t \in N$, $E(C^2) < \infty$;
- (b) $\forall t \in N$, $E(C) = \mu$, μ is a constant;
- (c) $\forall t, s \in N$, $Cov(c(t), c(s)) = \gamma_{t-s}$, γ_{t-s} is independence with t ,

then C is a stationary oscillation sequence.

Definition 3. Assume $X = (x(1), x(2), \dots, x(n))$ is an oscillation sequence,

$$C_i = (c_i(1), c_i(2), \dots, c_i(n)), \quad i = 1, 2, \dots, N, \quad (5)$$

$$R_N = (r_N^{(0)}(1), r_N^{(0)}(2), \dots, r_N^{(0)}(n)) \quad (6)$$

are IMFs and residual components of X decomposed by EMD algorithm, respectively. Then

$$\hat{c}_i(t) = \sum_{k=1}^p \phi_i(k)c_i(t-k) + a_t - \sum_{N=1}^q \theta_i(k)a_i(t-k), \quad a \sim N(0, \sigma^2) \quad (7)$$

is called ARMA (p, q) model. When $t = 2, 3, \dots, n$, $\hat{c}_i^{(0)}(t)$ is called the simulated data. When $t = n+1, n+2, \dots$, $\hat{c}_i^{(0)}(t)$ is called the predicted data.

In Equation (7), C_i is a stationary oscillation sequence; p is the ACF tail order of sequence C_i and q is the PACF tail order of C_i ; $\phi_i(1), \dots, \phi_i(p), \theta_i(1), \dots, \theta_i(q)$ are real parameters and be estimated by identification function ARMAX.

Definition 4. Assume sequence R_N is stated as
Definition 3. $R_N^{(1)} = (r_N^{(1)}(1), r_N^{(1)}(2), \dots, r_N^{(1)}(n))$ is accumulating generation sequence with one order of

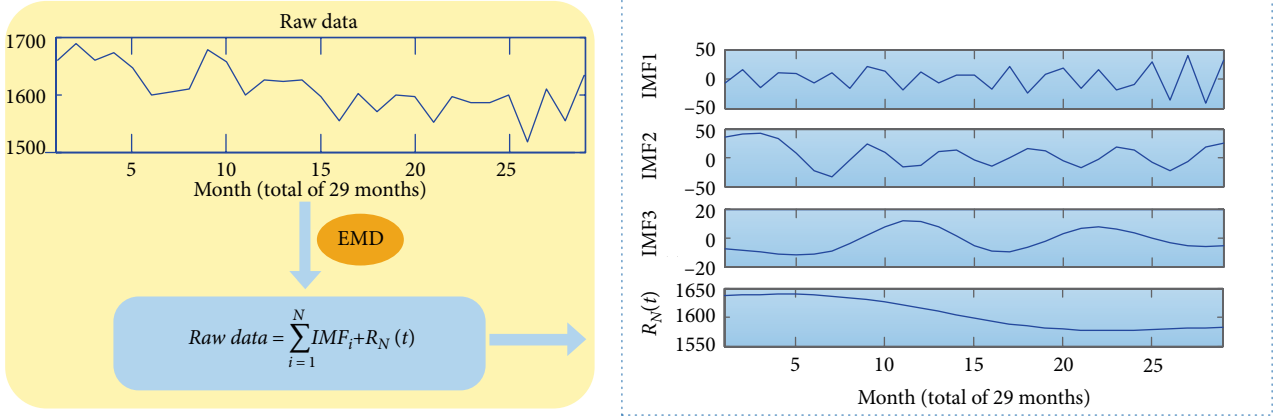


FIGURE 2: The effectiveness of EMD Algorithm.

$R_N, Z^{(1)} = (z_N^{(1)}(1), z_N^{(1)}(2), \dots, z_N^{(1)}(n))$ is mean sequence by consecutive neighbours of $R_N^{(1)}$, where

$$r_N^{(1)}(t) = \sum_{i=1}^t r_N^{(0)}(i), \quad t = 1, 2, \dots, n, \quad (8)$$

$$z_N^{(1)}(t) = 0.5 \times (r_N^{(1)}(t) + r_N^{(1)}(t-1)), \quad t = 1, 2, \dots, n. \quad (9)$$

Definition 5. Assume $R_N, R_N^{(1)}, Z^{(1)}$ are stated as Definition 3 and Definition 4. Then the following equation

$$r_N^{(0)}(t) + az_N^{(1)}(t) = b \quad (10)$$

is called GM (1,1) model, where a, b are real parameters.

Theorem 6. Assume that $P = [a, b]^T$ is parameter vector of Equation (10), where

$$Y = \begin{bmatrix} r_N^{(0)}(2) \\ r_N^{(0)}(3) \\ \vdots \\ r_N^{(0)}(n) \end{bmatrix}, \quad (11)$$

$$B = \begin{bmatrix} -z_N^{(1)}(2) & 1 \\ -z_N^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z_N^{(1)}(n) & 1 \end{bmatrix},$$

then the parameters of GM(1,1) model are identified as $\hat{a} = (B^T B)^{-1} B^T Y$.

Proof. GM(1,1) model is rewritten as follows,

$$\begin{aligned} r_N^{(0)}(2) + az_N^{(1)}(2) &= b, \\ r_N^{(0)}(3) + az_N^{(1)}(3) &= b, \\ &\vdots \\ r_N^{(0)}(n) + az_N^{(1)}(n) &= b. \end{aligned} \quad (12)$$

Matrix form of GM(1,1) model is as follows,

$$\begin{bmatrix} r_N^{(0)}(2) \\ r_N^{(0)}(3) \\ \vdots \\ r_N^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -z_N^{(1)}(2) & 1 \\ -z_N^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z_N^{(1)}(n) & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}, \quad (13)$$

that is

$$Y = BP. \quad (14)$$

Among the above equations, Y and B are known and P is sequence of undetermined parameters. The number of known equations is one, and the variables are two. Moreover, when the equations are incompatible, there is no solution, but the least square solution can be obtained by the least square method (LSM).

Assume ε is error sequence as follows,

$$\varepsilon = Y - BP. \quad (15)$$

Let

$$S = \min \|Y - BP\|^2 = \min (Y - BP)^T (Y - BP), \quad (16)$$

that is

$$\begin{aligned} S &= \min \sum_{t=2}^n (r_N^{(0)}(t) - (b - az_N^{(1)}(t)))^2 \\ &= \min \left(\sum_{t=2}^n r_N^{(0)}(t)^2 + (n-1)b^2 + a^2 \sum_{t=2}^n z_N^{(1)}(t)^2 - 2b \sum_{t=2}^n r_N^{(0)}(t) \right. \\ &\quad \left. + 2a \sum_{t=2}^n z_N^{(1)}(t)r_N^{(0)}(t) - 2ab \sum_{t=2}^n z_N^{(1)}(t) \right). \end{aligned} \quad (17)$$

According to LSM, S can be minimised with respect to parameters a, b to obtain

$$\frac{\partial S}{\partial a} = 2a \sum_{t=2}^n z_N^{(1)}(t)^2 + 2 \sum_{t=2}^n z_N^{(1)}(t)r_N^{(0)}(t) - 2b \sum_{t=2}^n z_N^{(1)}(t) = 0, \quad (18)$$

$$\frac{\partial S}{\partial b} = 2(n-1)b - 2 \sum_{t=2}^n r_N^{(0)}(t) - 2a \sum_{t=2}^n z_N^{(1)}(t) = 0. \quad (19)$$

Parameters a, b can be obtained, as follows,

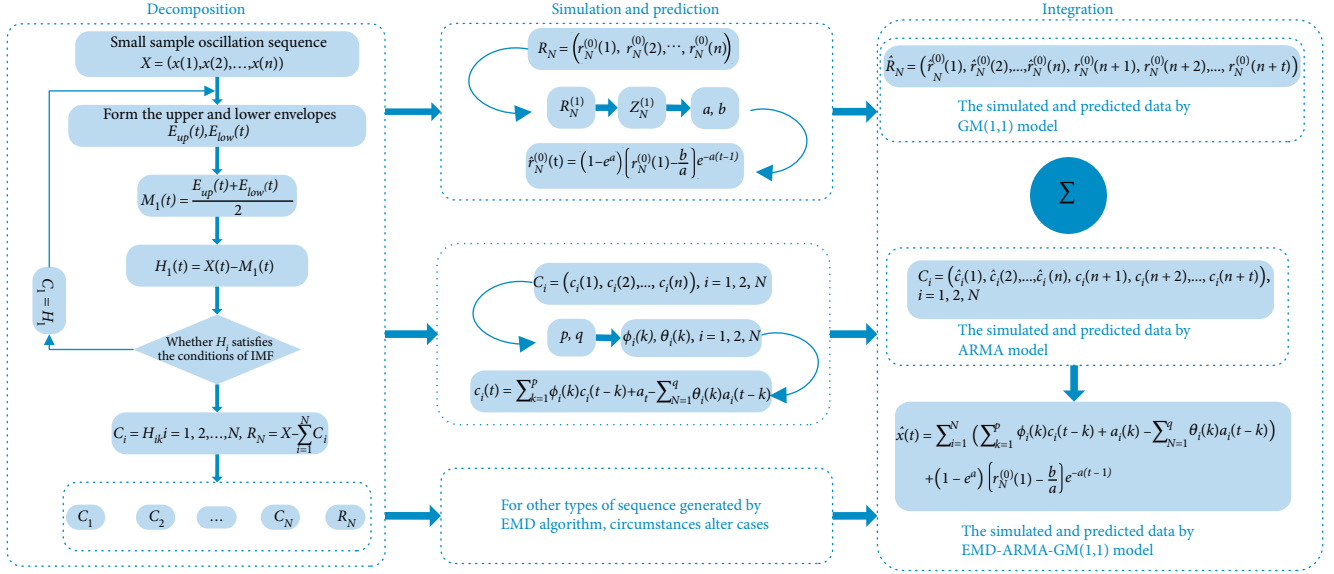


FIGURE 3: The flow chart of EMD-ARMA-GM(1,1) model.

TABLE 1: Data of crops disaster area in China from 2010 to 2017 (units: 1000 hectares).

Year	2010	2011	2012	2013	2014	2015	2016	2017
Crops disaster area	37425.9	32470.5	24962	31349.8	24890.7	21769.8	26220.7	18478.1

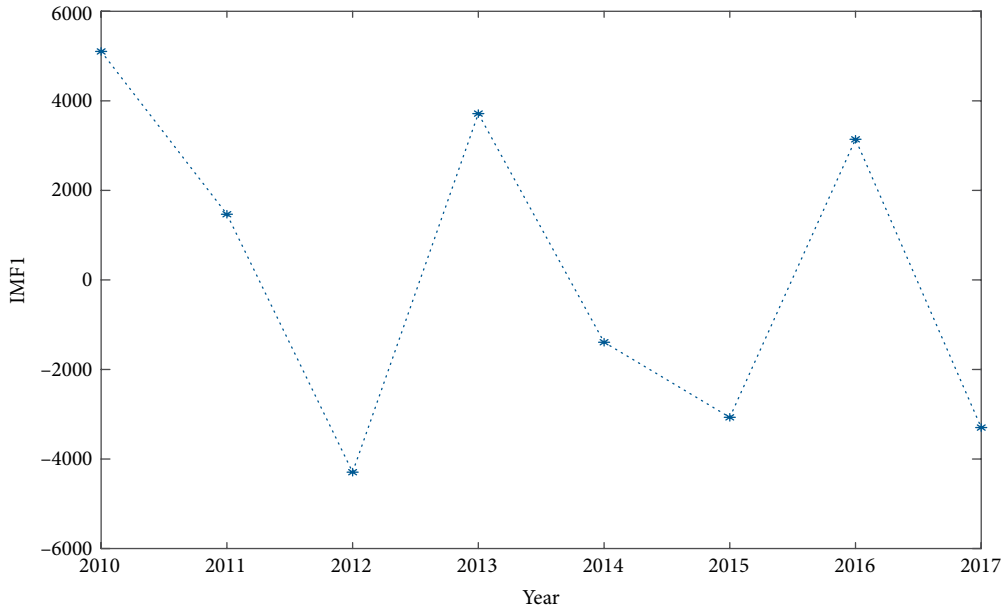
Data sources: <http://data.stats.gov.cn/easyquery.htm?cn=C01>.

FIGURE 4: The curve of IMF1.

$$a = \frac{\sum_{t=2}^n r_N^{(0)}(t) \sum_{t=2}^n z_N^{(1)}(t) - (n-1) \sum_{t=2}^n r_N^{(0)}(t) z_N^{(1)}(t)}{(n-1) \sum_{t=2}^n (z_N^{(1)}(t))^2 - \left(\sum_{t=2}^n z_N^{(1)}(t) \right)^2}, \quad (20)$$

$$b = \frac{\sum_{t=2}^n r_N^{(0)}(t) \sum_{t=2}^n (z_N^{(1)}(t))^2 - \sum_{t=2}^n z_N^{(1)}(t) \sum_{t=2}^n z_N^{(1)}(t) r_N^{(0)}(t)}{(n-1) \sum_{t=2}^n (z_N^{(1)}(t))^2 - \left(\sum_{t=2}^n z_N^{(1)}(t) \right)^2}. \quad (21)$$

Equations (20) and (21) are the expanded displays of parameter identification matrix. The proof is over. \square

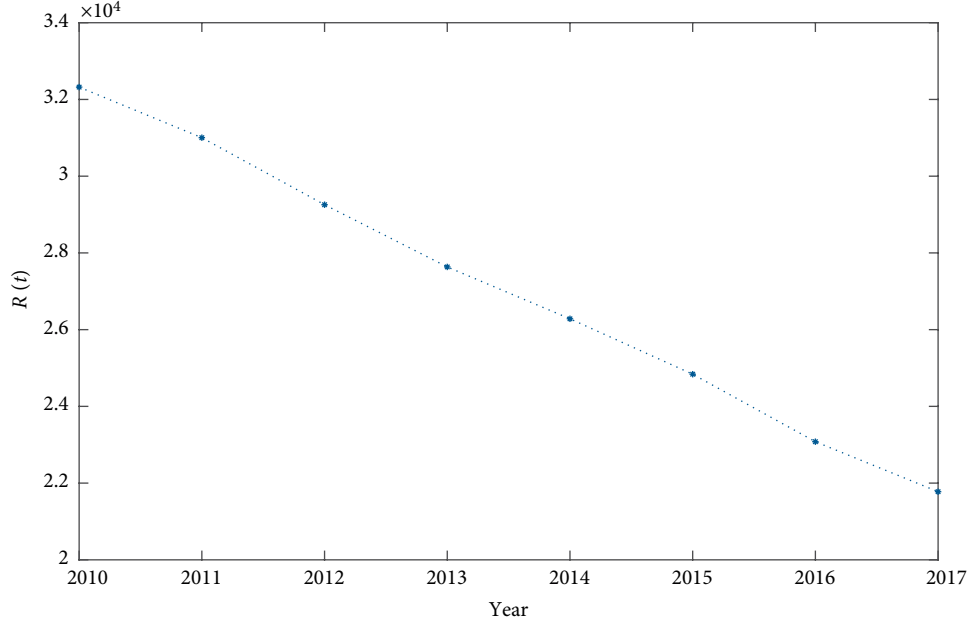
Definition 7. Assume sequence $R_N, R_N^{(1)}, Z_N^{(1)}$ and $\hat{a} = [a, b]^T$ are stated as Definition 4 and Theorem 1, then the following equation

$$\frac{dr^{(1)}}{dt} + ar^{(1)} = b \quad (22)$$

is named the whitization equation of GM(1,1) model.

The solution of the differential Equation (22) is as follows,

$$\hat{r}_N^{(0)}(t) = (1 - e^{-a}) \left(r_N^{(0)}(1) - \frac{b}{a} \right) e^{-a(t-1)}, \quad t = 1, 2, \dots, n. \quad (23)$$

FIGURE 5: The curve of $R(t)$.

Equation (23) is also called the time response function of the whitening differential equation. When $t = 2, 3, \dots, n$, $\hat{r}_N^{(0)}(t)$ is called the simulated data; When $t = n + 1, n + 2, \dots$, $\hat{r}_N^{(0)}(t)$ is called the predicted data.

Definition 8. Assume $\hat{c}_i(n)$, $\hat{r}_N^{(0)}(t)$ are stated as Definitions 3 and 7, then the following equation

$$\hat{x}(t) = \sum_{i=1}^N \left(\sum_{k=1}^p \phi_i(k) c_i(t-k)_{n-i} + a_i(k) - \sum_{k=1}^q \theta_i(k) a_i(t-k) \right) + (1 - e^a) \left(r_N^{(0)}(1) - \frac{b}{a} \right) e^{-a(t-1)} \quad (24)$$

is called EMD-ARMA-GM(1,1) model.

In the hybrid prediction model, EMD algorithm decomposes the original time series X into sequence C_i and sequence R_N to extract intrinsic characters of the complex system. sequence C_i is inputted into the ARMA model to describe the random changes and sequence R_N is substituted into GM(1,1) model to describe the trend. The value $\hat{x}(t)$ obtained by superposition $\hat{c}_i(n)$ and $\hat{r}_N^{(0)}(t)$ realizes the simulation or prediction of the original sequence. The flow of EMD-ARMA-GM(1,1) model is shown in Figure 3.

4. Modelling Condition and Error Checking Method for the EMD-ARMA-GM(1,1) Model

4.1. Modelling Condition of the EMD-ARMA-GM(1,1) Model. Each prediction model has a specific modelling condition and applicable rang. A model can be used for prediction only when the modelling condition is satisfied.

TABLE 2: The quasi-smooth condition of sequence $R(t)$.

Smoothness ratio	$\rho(3)$	$\rho(4)$	$\rho(5)$	$\rho(6)$	$\rho(7)$	$\rho(8)$
Value	0.462	0.298	0.218	0.169	0.134	0.112

TABLE 3: Parameter estimation of ARMA model.

Parameter of the ARMA model	p	q	ϕ_1	ϕ_2	θ_1
Estimated value	2	1	1.284	0.9363	0.9184

Definition 9. Assume that $T = (t(1), t(2), \dots, t(n))$, where $t(k) \geq 0$ for $k = 1, 2, \dots, n$, then the following is referred to as the smoothness ratio of sequence T :

$$\rho(k) = \frac{t(k)}{\sum_{i=1}^{k-1} t(i)}, \quad k = 2, 3, \dots, n. \quad (25)$$

The concept of smoothness ratio reflects the smoothness of a sequence. Obviously, the smoother the change of sequence T is, the smaller the smoothness ratio is.

Definition 10. If a sequence $T = (t(1), t(2), \dots, t(n))$, where $t(k) \geq 0$ for $k = 1, 2, \dots, n$ satisfies the following, then T is referred to as a quasi-smooth sequence:

- (1) $\rho(k+1)/\rho(k) < 1$, $k = 2, 3, \dots, n-1$,
- (2) $\rho(k) = [0, \varepsilon]$, $k = 3, 4, \dots, n$,
- (3) $\varepsilon < 0.5$.



FIGURE 6: The simulated curve of IMF1.

The quasi-smooth condition of residual component is used to act as the criteria to test whether an oscillation sequence can be used to establish EMD-ARMA-GM(1,1) model.

4.2. Error Checking Method for the EMD-ARMA-GM(1,1) Model. A model's performance can be judged by testing, and only the model that pass test can be meaningfully employed to make predictions.

Definition 11. Assume that a raw sequence $T^{(0)}$ is

$$T^{(0)} = (t^{(0)}(1), t^{(0)}(2), \dots, t^{(0)}(n)). \quad (26)$$

EMD-ARMA-GM(1,1) model is employed to simulate sequence $T^{(0)}$, and its corresponding simulation sequence is as follows,

$$\hat{T}^{(0)} = (\hat{t}^{(0)}(1), \hat{t}^{(0)}(2), \dots, \hat{t}^{(0)}(n)). \quad (27)$$

The residual sequence of $\hat{T}^{(0)}$ is ε_T , as follows,

$$\varepsilon_T = (\varepsilon_T(1), \varepsilon_T(2), \dots, \varepsilon_T(n)), \quad (28)$$

where

$$\varepsilon_T(u) = |t^{(0)}(u) - \hat{t}^{(0)}(u)|, \quad u = 1, 2, \dots, n. \quad (29)$$

The relative simulation percentage error (RSPE) of the simulation sequence is

$$\Delta_T = (\Delta_T(1), \Delta_T(2), \dots, \Delta_T(n)), \quad (30)$$

where

$$\Delta_T(u) = \left| \frac{\varepsilon_T(u)}{t^{(0)}(u)} \times 100\% \right|, \quad u = 1, 2, \dots, n. \quad (31)$$

The mean relative simulation percentage error (MRSPE) of simulation sequence Δ_T is as follows:

TABLE 4: Parameter estimation of GM(1,1) model.

Parameter of the GM(1,1) model	a	b
Estimated value	0.058197	33854.68073

$$\bar{\Delta}_T = \frac{1}{n} \sum_{u=1}^n \Delta_T(u). \quad (32)$$

For given threshold value ξ in which the threshold is set according to the specific situation of the system, when $\bar{\Delta}_T < \xi$ holds true, the grey model is said to be error-satisfactory.

5. Application

China is a large agricultural country, but its special geographical location and climate environment lead to natural disasters frequently, which cause a large number of crops disasters every year. Large-scale crops disaster has seriously affected the national grain security, the basic status of agriculture and the sustainable development of rural economy. A scientific prediction of crops disaster areas can provide reasonable reference for arranging agricultural production subsidy and disaster relief subsidy, which has positive significance for promoting the sustainable development of agriculture and China's economy.

The crops disaster in China has a long history. To prevent and mitigate disasters, Chinese government proposes and implements many significant policies since 2010. These policies have effectively improved the situation of crops disaster and profoundly influenced the crops disaster area in China. The data of crops disaster area in China from 2010 to 2017 are a small oscillation sequence.

The data of crops disaster area in China from 2010 to 2017 are shown in Table 1.

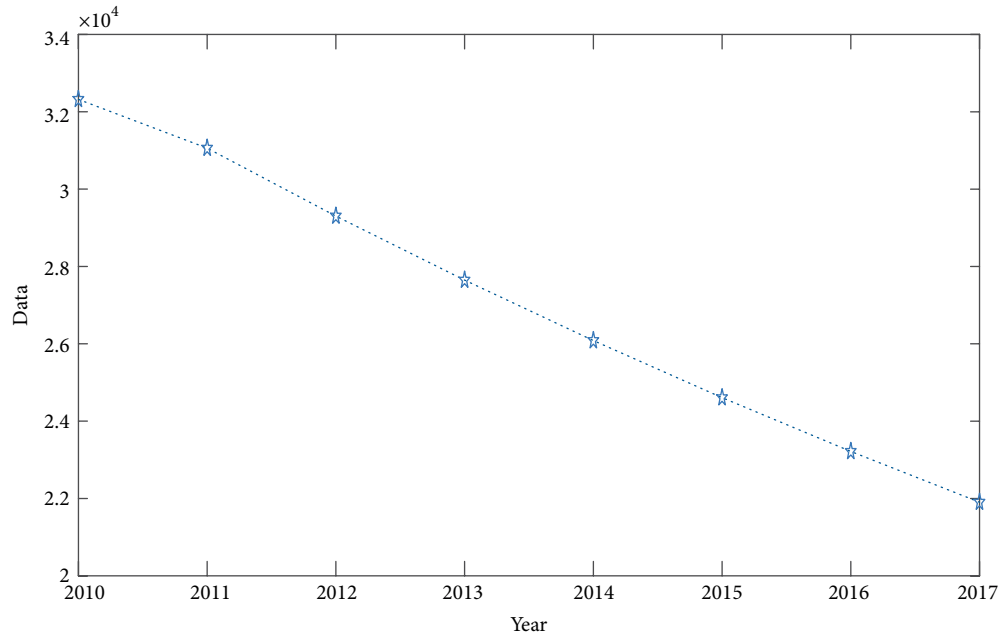


FIGURE 7: The simulated curve of $R(t)$.

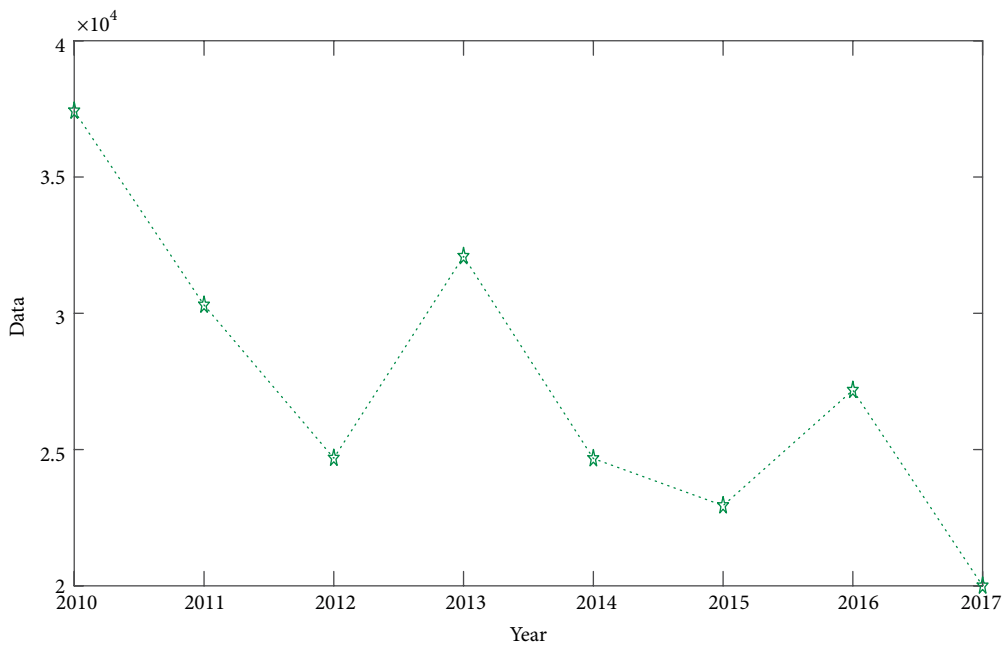


FIGURE 8: The simulated curve of crops disaster area in China.

5.1. *Data Decomposing.* EMD algorithm is applied to decompose the sequence of crops disaster area in China, and an IMF1 and a residual component $R(t)$ are obtained. The results are shown in Figures 4 and 5, respectively.

As can be seen from Figure 4, IMF1 is a curve of oscillations around the X -axis, showing linear and random characteristic of original sequence.

In Figure 5, $R(t)$ is a monotonic decreasing curve and shows the decreasing trend characteristic of the original sequence.

5.2. *Checking the Quasi-Smooth Condition before Modelling.* According to Definition 9, a sequence can be used to build the new model only when its residual component satisfies the quasi-smooth condition. Therefore, we check the quasi-smooth condition of sequence $R(t) = (32321.95, 31003.12, 29256.20, 27634.73, 26282.51, 24838.58, 23079.35, 21775.8)$ before building the new model to predict the crops disaster area.

From Definition 9, we can obtain the smoothness ratio of sequence $R(t)$ and the values of smoothness ratio are shown in

TABLE 5: EMD-ARMA-GM(1,1) model, ARIMA model and GM(1,1) model simulated values (units:1000 hectares).

Year	Raw data $x(t)$	EMD-GM(1,1)-ARMA		Model ARIMA		Model GM(1,1)	
		$\hat{x}(t)$	$\Delta_T(t)$	$\hat{x}(t)$	$\Delta_T(t)$	$\hat{x}(t)$	$\Delta_T(t)$
2010	37425.9	37425.9	—	38226.23	2.1384%	37425.9	—
2011	32470.5	30303.79	6.6729%	26364.23	18.8056%	31220.772	3.849%
2012	24962	24684.6	1.1113%	27976.36	12.0758%	29183.227	16.911%
2013	31349.8	32088.6	2.3566%	21655.3	30.9236%	27278.657	12.986%
2014	24890.7	24668.63	0.8922%	19212.75	22.8116%	25498.385	2.441%
2015	21769.8	22941.88	5.384%	18678.01	14.2022%	23834.297	9.483%
2016	26220.7	27177.18	3.6478%	20250.14	22.7704%	22278.812	15.033%
2017	18478.1	19995.16	8.21%	18478.1	—	20824.842	12.7%
MRSPE			4.0393%		17.6754%		10.4863%

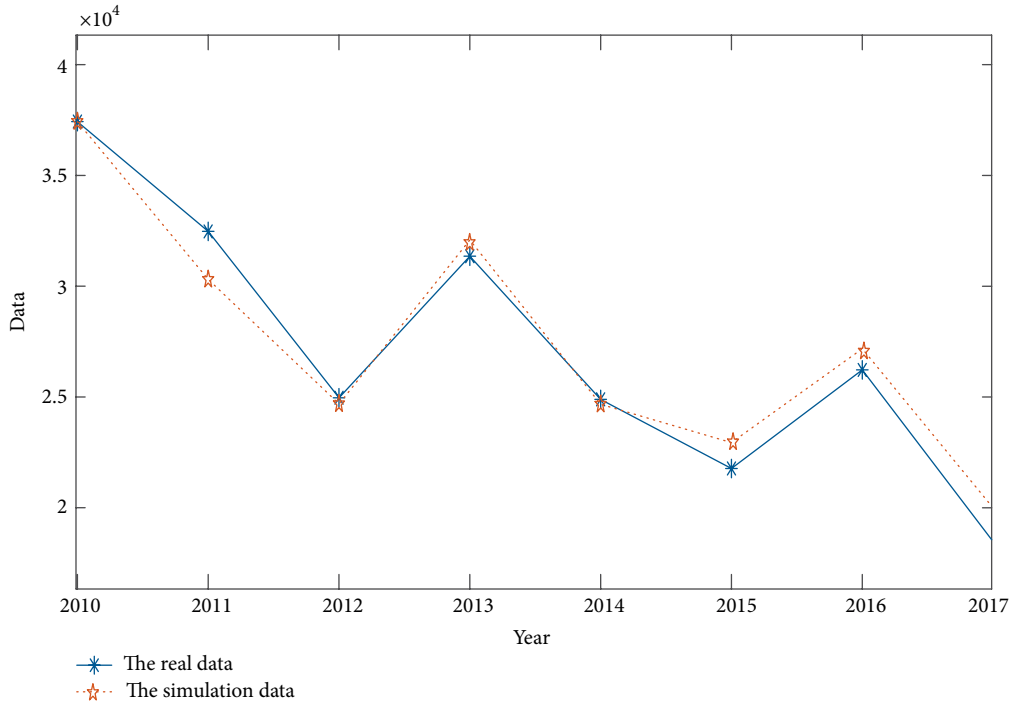


FIGURE 9: The simulated curve of the EMD-ARMA-GM(1,1) model.

Table 2. $\rho(8)/\rho(7) \approx 0.831513 < 1$, $\rho(6)/\rho(5) \approx 0.775513 < 1$, $\rho(5)/\rho(4) \approx 0.732441 < 1$ and $\rho(4)/\rho(3) \approx 0.646085 < 1$. Then, sequence of crops disaster area in China satisfies the quasi-smooth condition and can be used to build the new model. The modelling process is detailed in the next subsection.

5.3. Modelling. Firstly, IMF1 is introduced into ARMA (p, q) model. By increasing its order gradually, IMF1 is closer to the dependence of data. When fitting effect of the data is best, it stops and gets the value of p and q . Next, the parameter identification function ARMAX is used to estimate $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$. The optimal order and parameters are obtained as shown in Table 3.

As shown in Table 3, the proper value of p is 2 and q is 1. So we use ARMA(2, 1) model to simulate IMF1, and draw the simulated curve of this model based on IMF1, as shown in Figure 6.

Next, $R(t)$ is introduced into GM(1,1) and the parameters are estimated by least square method as shown in Table 4.

As provided in Table 4, We substitute the parameters into the whitening equation of GM(1,1) model, and get the simulated value of $R(t)$. The simulated curve of $R(t)$ is shown in Figure 7.

Finally, through integrating the simulated values of IMF1 and $R(t)$, we can get the simulated value of China's crops disaster area. The simulated curves of crops disaster area in China is shown in Figure 8.

5.4. Result and Analysis. To verify the performance of EMD-ARMA-GM(1,1) model, we compare the MRSPE of EMD-ARMA-GM(1,1) to that of traditional mainstream prediction models, including ARIMA model and GM(1,1) model. The simulated values $\hat{x}(t)$, $\Delta_r(t)$ and MRSPE of the three models are presented in Table 5.

As shown in Table 5, the proposed EMD-ARMA-GM(1,1) model has the lowest MRSPE among the three models and the MRSPE is 4.0393%; the MRSPEs of the other two models are

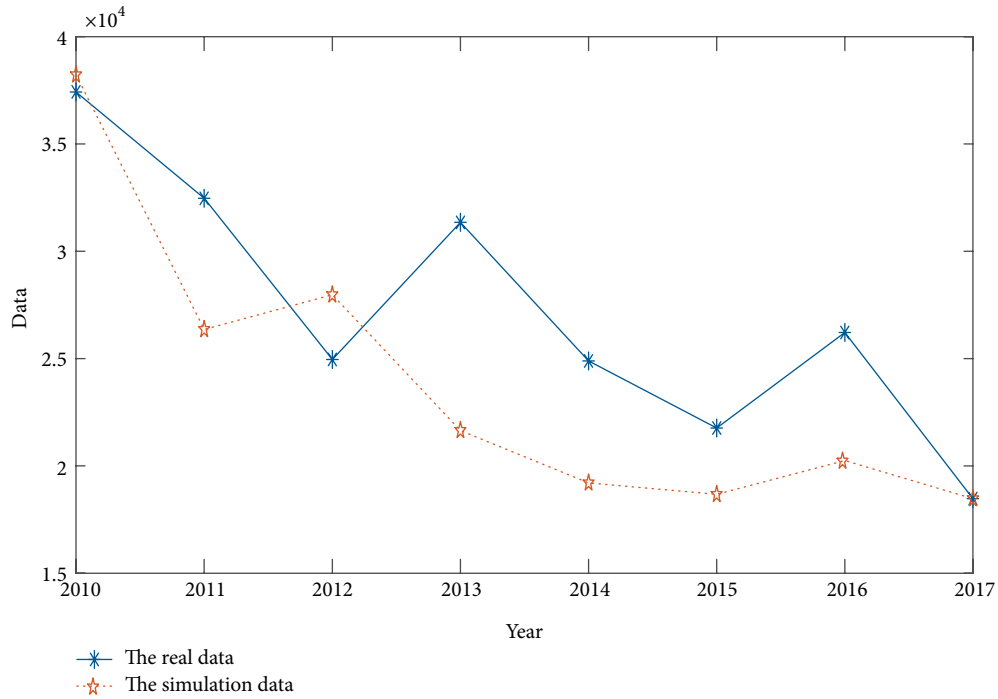


FIGURE 10: The simulated curve of the ARIMA model.

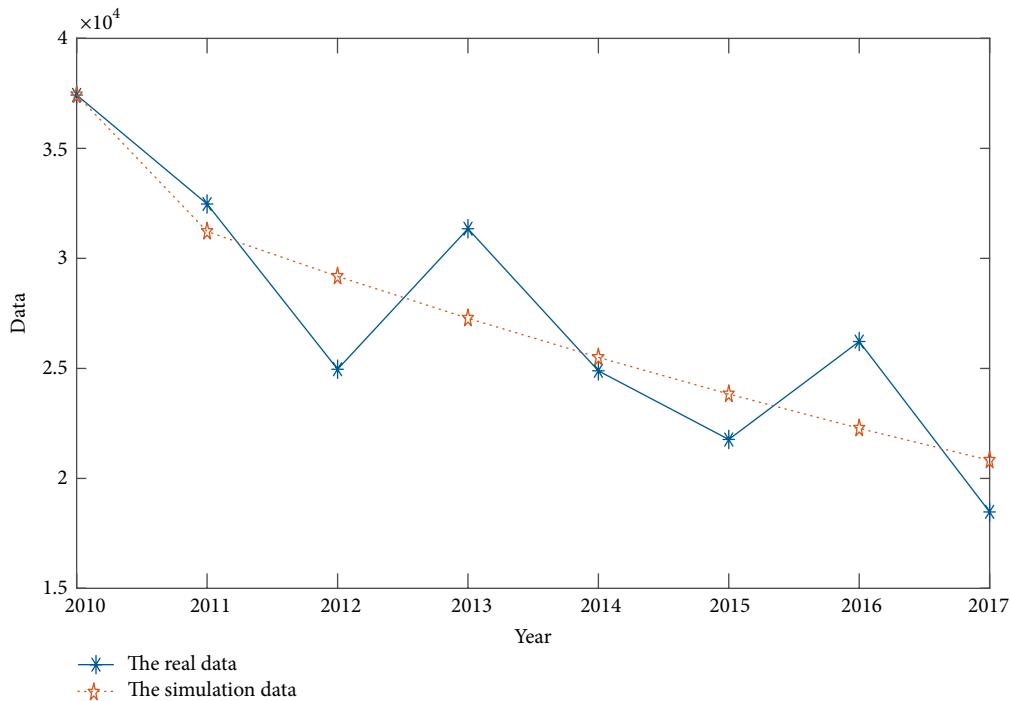


FIGURE 11: The simulated curve of the GM(1,1) model.

more than 10%. Comparatively, the performance of the GM(1,1) model is second to that of EMD-ARMA-GM(1,1) model because it does not consider the effect of random oscillation characteristic; the performance of the ARMA model is the worst among the three model because it does not consider the influence of trend characteristic. In order to clearly

illustrate the simulation effects of the three models for China's crops disaster area, we draw the simulated curves and errors of the three models based on the data in Table 5 in MATLAB as shown in Figures 9–12.

According to Figures 9–12, the performance of IEMD-ARMA-GM(1,1) model is best among the above three models.

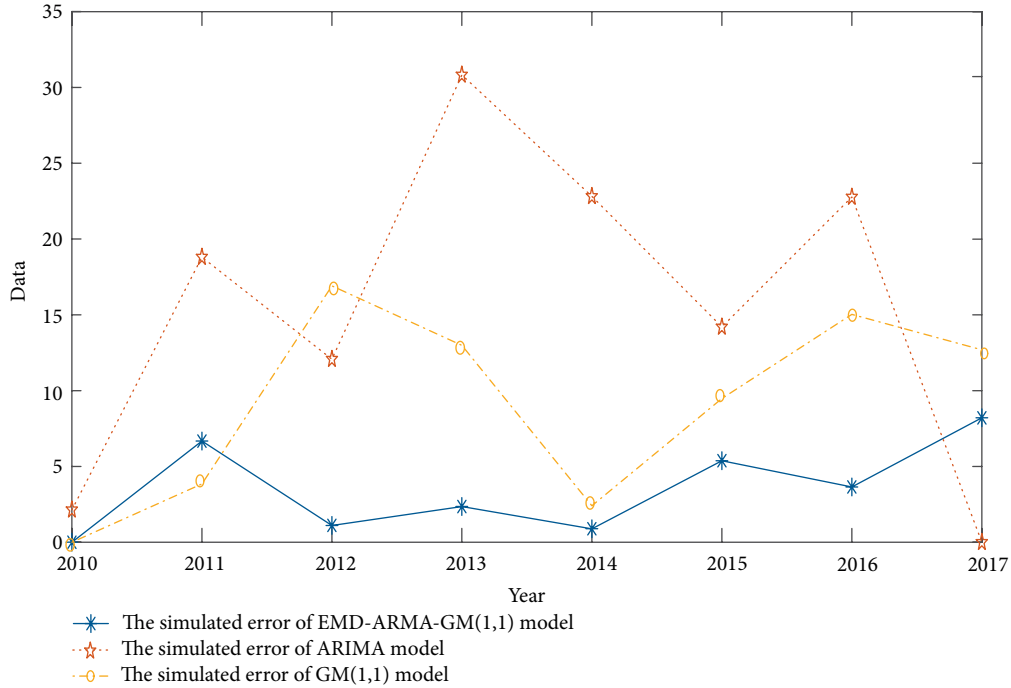


FIGURE 12: The simulated errors of EMD-ARMA-GM(1,1), ARIMA and GM(1,1).

TABLE 6: Prediction results of crops disaster area in China by EMD-ARMA-GM(1,1) model (units: 1000 hectares).

Year	2018	2019	2020	2021
Crops disaster area	20857.76	22335.99	14606.27	19633.39

Thus it is evident that the performance of EMD-ARMA-GM(1,1) model is better than that of traditional mainstream prediction models.

5.5. Prediction of Crops Disaster Area in China. The EMD-ARMA-GM(1,1) model is used to predict the crops disaster area in China from 2018 to 2021, and the results are shown in Table 6.

Table 6 shows that the overall trend of crops disaster area in China is decreasing in the next four years, but the crops disaster area is still very large. By 2021, it will reach 1963390 hectares. The large of crops disaster area may cause shortage of grain and inhibit rural economic. Therefore, in order to maintain the sustainable development of agriculture and national economy, the Chinese government needs to develop policies for production subsidies and disaster relief subsidies, and set aside sufficient funds to deal with the problems of crops failures caused by future natural disasters.

6. Conclusion

In this paper, the shortcomings of grey prediction model in modelling small oscillation sequence are analysed, and then we find out the reason why grey prediction model is

not effective in predicting oscillation sequence by analyzing the intrinsic characteristics of oscillation sequence: the system of oscillation sequence is complex, and the trend and random oscillation are often combined. Therefore, based on information decomposition and aiming at extracting the intrinsic characteristics of the sequence, a hybrid grey prediction model is established in this paper. The results of case analysis show that the proposed model considers the complexity of system information, effectively describes the operation behavior and rules of the system, and the effect is higher than that of a single classical prediction model.

The new grey hybrid prediction model provides a new idea and method for small oscillation sequence. However, when the size of oscillation sequence is big, the big data methods can be used to simulate and predict the oscillation sequence, such as neural network and support vector machine. At this time, the performance of the new hybrid grey prediction model needs to be compared with that of the big data method, and the simulation and prediction errors can be used to determine the performance of those methods, and then the superior one is selected for study the oscillation sequence.

In the following work, we will further consider the other characteristics of the sub-sequence generated by EMD algorithm, and establish suitable methods to study the oscillation sequence.

Data Availability

The China's crop disaster area data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

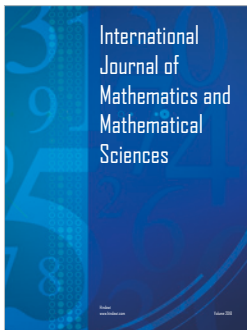
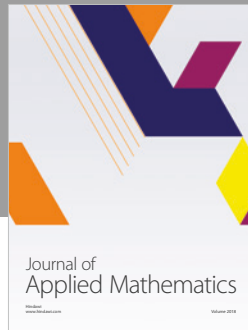
Acknowledgments

This work was supported by the National Natural Science Foundation of China (71771033).

References

- [1] H. M. Duan and X. P. Xiao, "A multimode dynamic short-term traffic flow grey prediction model of high-dimension tensors," *Complexity*, vol. 2019, Article ID 9162163, 18 pages, 2019.
- [2] M. Wei, D. L. Yang, and H. Huang, "Prediction of China's sulfur dioxide emissions by discrete grey model with fractional order generation operators," *Complexity*, vol. 2018, Article ID 8610679, 13 pages, 2018.
- [3] S. H. Mao, M. Y. Gao, J. H. Wen, and X. W. Nie, "Generalized admissible region of class ratio for fractional accumulated GM(1,1) model," *The Journal of Grey System*, vol. 26, no. 3, pp. 55–68, 2014.
- [4] B. Zeng and C. Li, "Forecasting the natural gas demand in China using a self-adapting intelligent grey model," *Energy*, vol. 112, no. 1, pp. 810–825, 2016.
- [5] X. Ma and Z. B. Liu, "The GMC (1, n) model with optimized parameters and its application," *The Journal of Grey System*, vol. 29, no. 4, pp. 122–138, 2017.
- [6] S. H. Mao and M. Y. Gao, "A novel fractional grey system model and its application," *Applied Mathematical Modelling*, vol. 40, no. 7–8, pp. 5063–5076, 2016.
- [7] B. Zeng and C. Li, "Improved multi-variable grey forecasting model with a dynamic background-value coefficient and its application," *Computers & Industrial Engineering*, vol. 118, pp. 278–290, 2018.
- [8] X. Ma, M. Xie, W. Q. Wu, and B. Zeng, "A novel fractional time delayed grey model with grey wolf optimizer and its applications in forecasting the natural gas and coal consumption in chongqing China," *Energy*, vol. 178, pp. 487–507, 2019.
- [9] B. Zeng, H. M. Duan, and Y. F. Zhou, "A new multivariable grey prediction model with structure compatibility," *Applied Mathematical Modelling*, vol. 75, pp. 385–397, 2019.
- [10] S. H. Mao, X. P. Xiao, M. Y. Gao, and X. Wang, "Nonlinear fractional order grey model of urban traffic flow short term prediction," *The Journal of Grey System*, vol. 30, no. 4, pp. 1–17, 2018.
- [11] J. Cui, H. Y. Ma, C. Q. Yuan, and S. Liu, "Novel grey verhulst model and its prediction accuracy," *The Journal of Grey System*, vol. 27, no. 2, pp. 122–138, 2017.
- [12] C. Puripat and S. Sarikavanij, "The comparison of grey system and the verhulst model for rainfall and water in dam prediction," *Advances in Meteorology*, vol. 2018, Article ID 7169130, 11 pages, 2018.
- [13] Z. X. Wang and Q. Li, "Modelling the nonlinear relationship between CO₂ emissions and economic growth using a PSO algorithm-based grey verhulst model," *Journal of Cleaner Production*, vol. 207, pp. 214–224, 2019.
- [14] D. Y. Chen and W. T. Han, "Prediction of multivariate chaotic time series via radial basis function neural network," *Complexity*, vol. 18, pp. 55–66, 2013.
- [15] J. Yang, J. F. Su, and L. J. Song, "Selection of manufacturing enterprise innovation design project based on consumer's green preferences," *Sustainability*, vol. 11, no. 5, p. 1375, 2019.
- [16] P. Jiang, X. Liu, J. Zhang, and X. Yuan, "A framework based on hidden markov model with adaptive weighting for microcystin forecasting and early-warning," *Decision Support Systems*, vol. 84, pp. 89–103, 2016.
- [17] X. Lu, J. Zhang, G. R. Wu, and X. Yuan, "On translation operator-based DGM(1,1) model and its properties," *The Journal of Grey System*, vol. 31, no. 1, pp. 13–22, 2019.
- [18] J. Zhang, X. Lu, M. F. Ran, and G. Han, "DGM model based on anti-cotangent function and its application," *The Journal of Grey System*, vol. 28, no. 3, pp. 63–74, 2016.
- [19] B. Zeng and S. F. Liu, "Prediction model of stochastic oscillation sequence based on amplitude compression," *Systems Engineering—Theory & Practical*, vol. 32, no. 11, pp. 2493–2497, 2012.
- [20] W. Y. Qian and Y. G. Dang, "GM(1,1) model based on oscillation sequences," *Systems Engineering—Theory & Practic*, vol. 29, no. 3, pp. 149–154, 2009.
- [21] A. Dejamkhooy, A. Dastfan, and A. Ahmadyfard, "Modelling and forecasting nonstationary voltage fluctuation based on grey system theory," *IEEE Transactions on Power Delivery*, vol. 32, no. 3, pp. 1212–1219, 2017.
- [22] X. Ma, W. Q. Wu, B. Zeng, Y. Wang, and X. Wu, "The conformable fractional grey system model," *ISA Transactions*, 2019.
- [23] F. M. Liu and W. C. Chuang, "A precaution diagnosis of financial distress via grey situation decision," *Journal of Grey System*, vol. 22, no. 3, pp. 395–403, 2010.
- [24] H. Zhou, J. J. Huang, Y. B. Yuan, and B. Tang, "Prediction of water consumption in hospitals based on a modified grey GM (0,1|sin) model of oscillation sequence: the example of Wuhan city," *Journal of Applied Mathematics*, vol. 2014, Article ID 521973, 7 pages, 2014.
- [25] J. L. Deng, "Undulating grey model (UGM) GM (1,1 tan(k-τ) p, sin(k-τ)p)," *The Journal of Grey System*, vol. 13, no. 3, pp. 201–205, 2001.
- [26] X. Du, Y. Wu, S. D. Gu, P. Sun, Y. Ji, and H.-M. Tai, "Power oscillation analysis and control of three-phase grid-connected voltage source converters under unbalanced grid faults," *IET Power Electronics*, vol. 9, no. 11, pp. 2162–2173, 2016.
- [27] X. F. Zhang and J. F. Su, "An integrated QFD and 2-tuple linguistic method for solution selection in crowdsourcing contests for innovative tasks," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 6, pp. 6329–6342, 2018.
- [28] J. Z. Wang, P. Du, T. Niu, and W. Yang, "A novel hybrid system based on a new proposed algorithm-multi-objective whale optimization algorithm for wind speed forecasting," *Applied Energy*, vol. 208, pp. 344–360, 2017.
- [29] W. Ghezaiel, A. Ben Slimane, and E. Braiek Ben, "Nonlinear multi-scale decomposition by EMD for co-channel speaker identification," *Multimedia Tool and Applications*, vol. 76, no. 20, pp. 20973–20988, 2017.
- [30] F. T. Wang, G. Deng, C. X. Liu, W. Su, Q. Han, and H. Li, "A deep feature extraction method for bearing fault diagnosis based on empirical mode decomposition and kernel function," *Advances in Mechanical Engineering*, vol. 10, no. 9, pp. 1–12, 2018.

- [31] X. Tang, Z. F. Ma, X. X. Niu, and Y. X. Yang, "Robust audio watermarking algorithm based on empirical mode decomposition," *Chinese Journal of Electronics*, vol. 25, no. 6, pp. 1005–1010, 2016.
- [32] Y. Zhang, C. Zhang, J. Sun, and J. Guo, "Improved wind speed prediction using empirical mode decomposition," *Advances in Electrical and Computer Engineering*, vol. 18, no. 2, pp. 3–10, 2018.




Hindawi

Submit your manuscripts at
www.hindawi.com

