Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 5471748, 9 pages
http://dx.doi.org/10.1155/2016/5471748

Research Article

Research on a Novel Kernel Based Grey Prediction Model and Its Applications

Xin Ma

School of Science, Southwest University of Science and Technology, Mianyang, China

Correspondence should be addressed to Xin Ma; cauchy7203@gmail.com

Received 2 July 2016; Revised 23 October 2016; Accepted 14 November 2016

Academic Editor: Michele Betti

The discrete grey prediction models have attracted considerable interest of research due to its effectiveness to improve the modelling accuracy of the traditional grey prediction models. The autoregressive GM(1, 1) model, abbreviated as ARGM(1, 1), is a novel discrete grey model which is easy to use and accurate in prediction of approximate nonhomogeneous exponential time series. However, the ARGM(1, 1) is essentially a linear model; thus, its applicability is still limited. In this paper a novel kernel based ARGM(1, 1) model is proposed, abbreviated as KARGM(1, 1). The KARGM(1, 1) has a nonlinear function which can be expressed by a kernel function using the kernel method, and its modelling procedures are presented in details. Two case studies of predicting the monthly gas well production are carried out with the real world production data. The results of KARGM(1, 1) model are compared to the existing discrete univariate grey prediction models, including ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NGBMOP, and it is shown that the KARGM(1, 1) outperforms the other four models.

1. Introduction

The idea of the "Grey Box" modelling is trying to combine the advantages of the "White Box" and the "Black Box." Deng [1] has pioneered the Grey System Theory based on this idea. The grey prediction models play an important role in the Grey System Theory, and because of their effectiveness in time series prediction the grey prediction models have been widely adopted [2–5].

Over three decades of development, many new grey prediction models have been put forward, such as the FGM(1, 1) [6], DGMD(1, 1, 1) [7], NGM(1, 1, k) [8], and SAGM(1, 1) [9]. Along with these new models, some novel methodologies have also been proposed, and the discrete modelling technique is one of the most efficient methods to build the grey prediction models. The discrete modelling technique has been developed from the research of the DGM(1, 1) model [10], which is based on the basic GM(1, 1) model. This novel technique has also been used to build the NDGM(1, 1, k) model based on the NGM(1, 1, k) [11]. And in our previous works, this technique has been extended to build the discrete GM(1, n) models [12, 13]. In these works, the discrete modelling technique has been proved efficient to improve the accuracy of the grey prediction models. Some novel grey prediction models for the nonlinear sequences are developed in recent years. For the univariate regression problems, the nonlinear grey Bernoulli model (NGBM) has been proposed by Chen et al. [14], which is more flexible than the existing grey prediction models and efficient to predict various time series. The NGBM model has attracted considerable research, and some improved grey prediction models based on it have been proposed, such as the Nash NGBM [15], the NGBM with optimal parameter [16], and the optimized NGBM [17] model. As for the multivariate regression problems, the nonlinear GMC(1, n) model [18] has been proved to be more efficient to predict the nonlinear series than the existing models.

In recent researches, a novel grey prediction model directly built on the original series has been proposed, which is called the DDGM(1, 1) model [19] and is also called the autoregressive GM(1, 1) model (ARGM(1, 1)) [20] as it is essentially in the autoregressive formulation. It is unnecessary to use the 1-AGO when building the ARGM(1, 1) model; thus it is very easy to use. The ARGM(1, 1) model has been proved to coincidence with the nonhomogeneous exponential law [19], and it has also been presented to be more efficient than

the DGM(1, 1) model in some applications [19, 20]. However, the ARGM(1, 1) model is essentially a linear model; thus its applicability is limited.

In order to improve the applicability of the ARGM(1, 1) model, we use the kernel method to build a novel kernel based ARGM(1, 1) model, abbreviated as KARGM(1, 1). The kernel method has been developed from the Vapnik's Support Vector Machines (SVM) [21], and it has been proved to be very efficient to convert the classical linear models into nonlinear models in the previous researches [22–24]. The researches of Vapnik's SVM [21] are the initial works of the kernel method. But the formulation of the Vapnik's SVM is not easy to use as it involves in a quadric problem with inequivalent constraints. Suykens and Vandewalle [25] have proposed a simplified formulation of kernel method involving a quadric problem with equivalent constraints, which can be converted into a linear system. And the formulation by Suykens and Vandewalle has been proved to be available to extend to linear models into nonlinear models as efficient as the formulation of Vapnik's SVM [26, 27]. As for the time series regression, the recurrent LS-SVM [28] is a typical model for the nonlinear univariate time series, which can be more easily used to predict the chaotic time series than the recurrent neural networks [29]. In this work, the kernel method by Suykens and Vandewalle will be used to build the KARGM(1, 1) model.

The rest of this paper is organized as follows. Section 2 presents a brief overview of the existing ARGM(1, 1) model; Section 3 presents the modelling procedures of the KARGM(1, 1); two case studies of predicting the gas well production are presented in Section 4, and conclusions are drawn in Section 5.

2. Overview of the Autoregressive GM(1, 1) Model

With a given time series $\{x(1), x(2), \ldots, x(n)\}$, the autoregressive GM(1, 1) (ARGM(1, 1)) model is represented as the following linear difference equation [20]:

$$x(k) = \alpha x(k-1) + \beta, \quad k = 2, 3, \ldots, n. \tag{1}$$

The parameters α and β can be obtained using the least squares method as follows:

$$[\alpha, \beta]^T = \left(B^T B\right)^{-1} B^T Y, \tag{2}$$

where

$$Y = \begin{pmatrix} x(2) \\ x(3) \\ \vdots \\ x(n) \end{pmatrix},$$

$$B = \begin{pmatrix} x(1) & 1 \\ x(2) & 1 \\ \vdots & \vdots \\ x(n-1) & 1 \end{pmatrix}. \tag{3}$$

The solution of the ARGM(1, 1) model can be obtained using the recursive method, which is given as the following discrete function:

$$x(k) = \alpha^k x(1) + \frac{1-\alpha^k}{1-\alpha}\beta. \tag{4}$$

The discrete function (4) is used to compute the predicted values.

3. The Proposed KARGM(1, 1) Model

In this section, the modelling procedures of the novel kernel based autoregressive GM(1, 1) model, abbreviated as KARGM(1, 1), will be presented.

3.1. Representation of the KARGM(1, 1) Model. With a given time series $\{x(1), x(2), \ldots, x(n)\}$, the KARGM(1, 1) model is represented as the following difference equation:

$$x(k) = \alpha x(k-1) + w^T \varphi(k) + \beta, \tag{5}$$

where φ is a nonlinear mapping, which is defined as

$$\varphi : R \longrightarrow \mathcal{F}, \tag{6}$$

and $\mathcal{F}$ is a higher dimensional feature space; w is a vector in $\mathcal{F}$.

The linear combination $w^T\varphi(k)$ is a nonlinear function of k. For example, if we consider a nonlinear function

$$g(k) = 0.3 \sin(k * \pi) + \sqrt{k}, \tag{7}$$

and define the nonlinear mapping as

$$\varphi(k) = \left[\sin(k * \pi), \sqrt{k}\right]^T, \tag{8}$$

and set $w = [0.3, 1]^T$, then we have

$$g(k) = w^T \varphi(k). \tag{9}$$

3.2. Parameters Estimation of the KARGM(1, 1) Model. Being different from the ARGM(1, 1) model, we cannot simply use the least squares method to estimate the parameters of the KARGM(1, 1) model, because it is not always computationally feasible to find the formulation of the nonlinear mapping φ [30]. Firstly, we consider the regularized problem as follows:

$$\begin{aligned} \min \quad & \mathcal{J}(\alpha, w, \mathbf{e}) = \frac{1}{2}\alpha^2 + \frac{1}{2}\|w\|^2 + \frac{\gamma}{2}\sum_{k=2}^{n} e_k^2 \\ \text{s.t.} \quad & x(k) = \alpha x(k-1) + w^T \varphi(k) + \beta + e_k, \end{aligned} \tag{10}$$

where γ is the regularized parameter to balance the flatness of the fitting curve and the error bound, and $\|\cdot\|$ is the 2-norm. The optimization problem (10) can be solved using the KKT conditions, which have been presented in Appendix A.

We firstly define the Lagrangian function as

$$\begin{aligned}&\mathscr{L}(\alpha, w, \mathbf{e}; \boldsymbol{\lambda})\\&\quad = \mathscr{J}(\alpha, w, \mathbf{e})\\&\qquad - \sum_{k=2}^{n} \lambda_k \left\{\alpha x(k-1) + w^T \varphi(k) + \beta + e_k - x(k)\right\},\end{aligned} \tag{11}$$

where λ_k is the Lagrangian multiplier. Then the KKT conditions can be given as

$$\begin{aligned}&\frac{\partial \mathscr{L}}{\partial \alpha} = 0 \Longrightarrow\\&\quad \alpha = \sum_{k=2}^{n} \lambda_k x(k-1)\\&\frac{\partial \mathscr{L}}{\partial w} = 0 \Longrightarrow\\&\quad w = \sum_{k=2}^{n} \lambda_k \varphi(k)\\&\frac{\partial \mathscr{L}}{\partial \beta} = 0 \Longrightarrow\\&\sum_{k=2}^{n} \lambda_k = 0\\&\frac{\partial \mathscr{L}}{\partial e_k} = 0 \Longrightarrow\\&\quad e_k = \lambda_k \gamma^{-1}\\&\frac{\partial \mathscr{L}}{\partial \lambda_k} = 0 \Longrightarrow\\&x(k) = \alpha x(k-1) + w^T \varphi(k) + \beta + e_k.\end{aligned} \tag{12}$$

By eliminating the α, w and e_k, the KKT conditions can be converted to the following linear system:

$$\begin{pmatrix} 0 & \vec{1}_{n-1}^T \\ \vec{1}_{n-1} & \Omega + Q + \gamma^{-1} I_{n-1} \end{pmatrix} \begin{pmatrix} \beta \\ \vec{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ Y \end{pmatrix}, \tag{13}$$

where

$$\begin{aligned}\vec{1}_{n-1} &= [1, 1, \ldots, 1]_{n-1}^T,\\ \vec{\lambda} &= [\lambda_2, \lambda_3, \ldots, \lambda_n]^T,\\ Y &= [x(2), x(3), \ldots, x(n)]^T,\\ Q_{ij} &= x(i)\, x(j),\\ \Omega_{ij} &= \varphi(i+1) \cdot \varphi(j+1),\end{aligned} \tag{14}$$

and I_{n-1} and $n-1$ dimensional identity matrix with all the diagonal elements are 1 and others are zero. The inner product Ω_{ij} can be expressed using a kernel function which satisfies the Mercer's condition, that is,

$$\Omega_{ij} = \varphi(i+1) \cdot \varphi(j+1) = K(i+1, j+1). \tag{15}$$

The Gaussian kernel is often employed, which is defined as

$$K(i, j) = \exp\left\{-\frac{(i-j)^2}{2\sigma^2}\right\}, \tag{16}$$

where σ is the kernel parameter.

The linear system (13) can be solved within the inner product (15) expressed by the Gaussian kernel (16), and then the parameter β and the Lagrangian multipliers λ_k can be obtained. The parameter α can then be computed using the first equation in the KKT conditions (12).

3.3. The Solution of the KARGM(1, 1) Model. The KARGM(1, 1) model (5) can be easily solved using the recursive method as follows:

$$\begin{aligned}x(k) &= \alpha x(k-1) + w^T \varphi(k) + \beta\\&= \alpha\left[\alpha x(k-2) + w^T \varphi(k-1) + \beta\right] + w^T \varphi(k)\\&\quad + \beta = \alpha^2 x(k-2) + \sum_{\tau=k-1}^{k} \alpha^{k-\tau}\left[w^T \varphi(\tau) + \beta\right]\\&= \cdots = \alpha^{k-1} x(1) + \sum_{\tau=2}^{k} a^{k-\tau}\left[w^T \varphi(\tau) + \beta\right].\end{aligned} \tag{17}$$

Noticing the second equation in the KKT conditions (12), which can be rewritten as $w = \sum_{j=2}^{n} \lambda_j \varphi(j)$, we have

$$w^T \varphi(\tau) = \sum_{j=2}^{n} \lambda_j \varphi(j) \cdot \varphi(\tau). \tag{18}$$

Using the inner product (15), we can rewrite the nonlinear function (18) as follows:

$$w^T \varphi(\tau) = \sum_{j=2}^{n} \lambda_j K(j, \tau). \tag{19}$$

For convenience, we note the following discrete function:

$$\phi(\tau) = w^T \varphi(\tau) + \beta = \sum_{j=2}^{n} \lambda_j K(j, \tau) + \beta. \tag{20}$$

Then the solution (17) can be rewritten as

$$x(k) = \alpha^{k-1} x(1) + \sum_{\tau=2}^{k} a^{k-\tau} \phi(\tau). \tag{21}$$

The discrete function (21) can be used to compute the predicted series.

Table 1: Monthly gas production data of the B51 in Sichuan, China.

Month	Gas production	Month	Gas production	Month	Gas production	Month	Gas production
1	90.2837	6	63.9314	11	34.2398	16	53.2518
2	72.1744	7	30.8003	12	44.0861	17	50.3282
3	54.7787	8	60.5955	13	49.6398	18	37.8982
4	54.1369	9	58.2767	14	51.7017	19	36.4632
5	64.6249	10	7.5851	15	44.5744	20	48.0127

It should be noticed that we do not need to know the expression of the nonlinear mapping φ, because all the computational procedures of the KARGM(1, 1) model only involve in the inner product $\varphi(i)\cdot\varphi(j)$, and it can be expressed by a proper kernel function.

Actually, the KARGM(1, 1) represents a dynamical system which contains a "White" part and a "Black" part. The "White" part is the linear recursion which is known *a priori*, and the "Black" part is the linear combination $w^T\varphi(k)$, which is expressed by the kernel function and finally determined by the raw data. So we can say that the KARGM(1, 1) model is a real "Grey" model.

3.4. Summary of Computational Steps. The computational steps of the KARGM(1, 1) model are summarized as follows.

Step 1. Select an appropriate kernel function $K(\cdot,\cdot)$ and the regularized parameter γ in (10).

Step 2. Compute the parameters β and $\lambda_2, \lambda_3, \ldots, \lambda_n$ by solving the linear system (13), and then compute the parameter α using the first equation in the KKT conditions (12).

Step 3. Compute the values of the predicted series using the discrete function (21).

4. Applications

Two case studies of predicting the gas well production are carried out in this section to validate the effectiveness of the KARGM(1, 1) model. The monthly production data are collected from two gas wells in Sichuan, China. In order to compare the performance of the KARGM(1, 1) model and other existing discrete grey models, the ARGM(1, 1), NDGM(1, 1, k) [11], DGM(1, 1) [10], and the nonlinear grey Bernoulli model with optimal parameter i (NGBMOP) [16] are also applied in the case studies. The mean absolute percentage error is used as an overall measurement of accuracy of the prediction models, which is defined as

$$\text{MAPE} = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{\hat{x}(k) - x(k)}{x(k)}\right| \times 100\,(\%)\,, \qquad (22)$$

where $x(k)$ is the real value and $\hat{x}(k)$ is the predicted value, respectively.

4.1. Case Study 1. The raw data used in case study 1 are collected from the gas well B51 in Sichuan, China. Twenty points of monthly gas production ($10^4 m^3$) are listed in Table 1. The first 15 points are used to build the models, and the last 5 points are used for testing. The Gaussian kernel (16) is used to build the KARGM(1, 1) model.

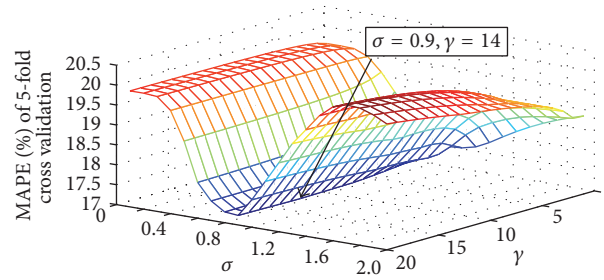

Figure 1: The MAPE of cross validation of KARGM(1, 1) with different σ and γ using the 5-fold cross validation in Case Study 1.

The kernel parameter σ and the regularized parameter γ are tuned using the cross validation, which is widely used in the other kernel methods, such as the Partially Linear LS-SVM [31]. The 5-fold cross validation is employed in this case. The MAPE of validation is used to assess the chosen kernel parameter and the regularized parameter. A brief summary of the 5-fold cross validation has been described in Appendix B. The results of the MAPE for each pair of σ and γ are plotted in Figure 1. It can be seen that the minimum MAPE is found at the point $\sigma = 0.9$ and $\gamma = 14$; thus this pair will be used to build the KARGM(1, 1) in this case.

The predicted values by the KARGM(1, 1), ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NGBMOP are listed in Table 2, along with the absolute percentage error of each point and the MAPE of each model. The results are also plotted in Figure 2.

It can be seen in Table 2 that the minimum MAPE for fitting of ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NBGMOP is more than three times larger than that of KARGM(1, 1). The MAPE for prediction of KARGM(1, 1) is much smaller than the NDGM(1, 1, k), DGM(1, 1), and NBGMOP, and it is slightly larger than that of the ARGM(1, 1). It is shown in Figure 2 that the predicted values of KARGM(1, 1) are quite close to the real values, and these values are presented to follow the overall trend of the monthly production of B51; but the distances between the predicted values of the other four models are very large, and they have failed in catching the overall trend of the gas production, which indicates that the modelling accuracy of these four models is not acceptable. In summary, the KARGM(1, 1) performs best in the presented models.

Table 2: Predicted values of monthly gas production of B51 by the ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), NGBMOP, and KARGM(1, 1).

Month	Gas production	ARGM(1, 1)	Error	NDGM(1, 1, k)	Error	DGM(1, 1)	Error	NGBMOP	Error	KAGM(1, 1)	Error
1	90.2837	90.2837	0	90.2837	0	90.2837	0	90.2837	0.00	90.2837	0.00
2	72.1744	59.1583	18.03	73.2911	1.55	63.5896	11.89	72.8510	0.94	70.0022	3.01
3	54.7787	51.0645	6.78	64.7782	18.25	61.0010	11.36	65.2404	19.10	55.9178	2.08
4	54.1369	48.9598	9.56	57.9739	7.09	58.5177	8.09	61.4602	13.53	53.5025	1.17
5	64.6249	48.4125	25.09	52.6425	18.54	56.1356	13.14	58.5062	9.47	65.0060	0.59
6	63.9314	48.2702	24.50	48.5808	24.01	53.8504	15.77	55.9700	12.45	61.2468	4.20
7	30.8003	48.2332	56.60	45.6134	48.09	51.6582	67.72	53.6995	74.35	34.3068	11.38
8	60.5955	48.2236	20.42	43.5893	28.07	49.5553	18.22	51.6200	14.81	59.2332	2.25
9	58.2767	48.2211	17.25	42.3785	27.28	47.5380	18.43	49.6887	14.74	55.7955	4.26
10	7.5851	48.2204	535.73	41.8686	451.98	45.6028	501.22	47.8783	531.21	12.5928	66.02
11	34.2398	48.2202	40.83	41.9629	22.56	43.7464	27.76	46.1701	34.84	32.4912	5.11
12	44.0861	48.2202	9.38	42.5780	3.42	41.9656	4.81	44.5509	1.05	45.1528	2.42
13	49.6398	48.2202	2.86	43.6422	12.08	40.2572	18.90	43.0106	13.35	49.4121	0.46
14	51.7017	48.2202	6.73	45.0933	12.78	38.6184	25.31	41.5412	19.65	51.3527	0.68
15	44.5744	48.2202	8.18	46.8780	5.17	37.0463	16.89	40.1364	9.96	45.2044	1.41
MAPE			52.13		45.39		50.63		51.30		7.00
16	53.2518	48.2202	9.45	48.9502	8.08	35.5382	33.26	38.7911	27.16	47.2950	11.19
17	50.3282	48.2202	4.19	51.2703	1.87	34.0915	32.26	37.5008	25.49	50.3266	0.00
18	37.8982	48.2202	27.24	53.8041	41.97	32.7037	13.71	36.2619	4.32	50.5842	33.47
19	36.4632	48.2202	32.24	56.5221	55.01	31.3724	13.96	35.0710	3.82	50.5825	38.72
20	48.0127	48.2202	0.43	59.3988	23.71	30.0953	37.32	33.9254	29.34	50.5832	5.35
MAPE			14.71		26.13		26.10		18.02		17.75

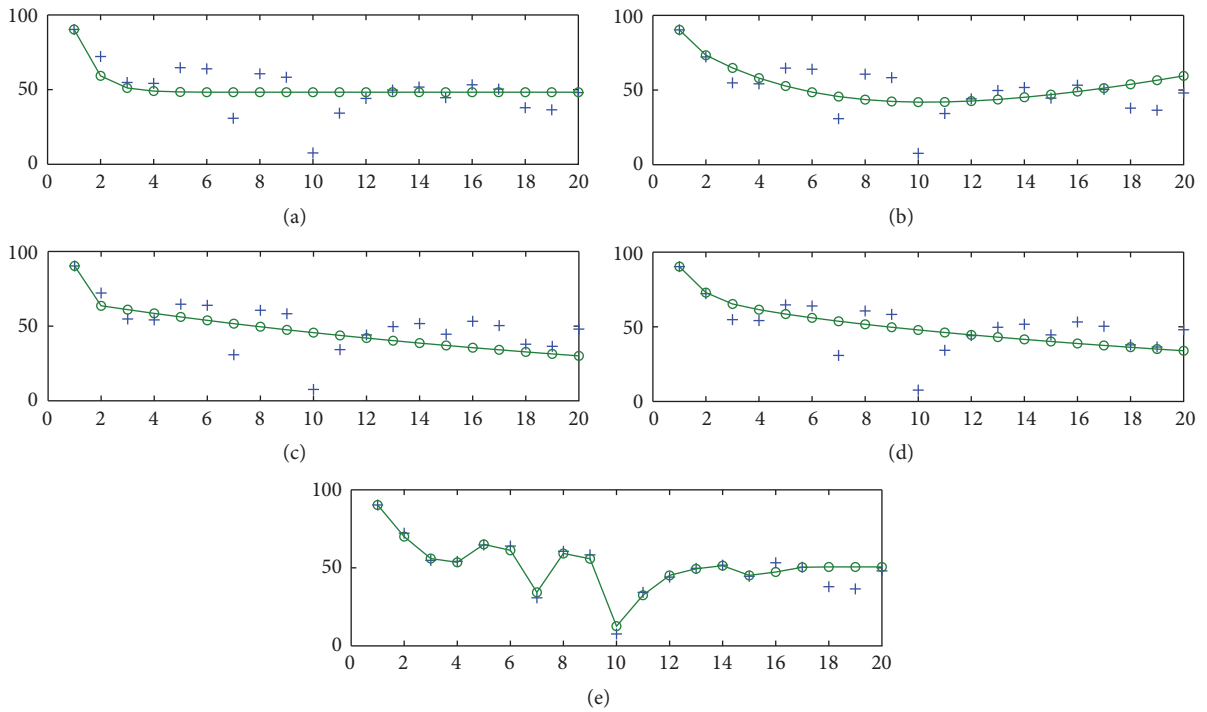

Figure 2: Prediction results of the monthly gas production of B51 by the prediction models. The raw data are plotted using the mark "+" and the predicted values are plotted using the solid line and the mark "o". (a) ARGM(1, 1); (b) NDGM(1, 1, k); (c) DGM(1, 1); (d) NGBMOP; (e) KARGM(1, 1).

Table 3: Monthly gas production data of the B41 in Sichuan, China.

Month	Gas production	Month	Gas production	Month	Gas production	Month	Gas production
1	12.8	6	9.3	11	12.2	16	7.5
2	9.2	7	10.3	12	11	17	6.8
3	8.8	8	9.4	13	9	18	7.5
4	9.2	9	9.3	14	8.2	19	7.7
5	8.5	10	11.2	15	6.4	20	8.5

Table 4: Predicted values of monthly gas production of B41 by the ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), NGBMOP, and KARGM(1, 1).

Month	Gas production	ARGM(1, 1)	Error	NDGM(1, 1, k)	Error	DGM(1, 1)	Error	NGBMOP	Error	KAGM(1, 1)	Error
1	12.8	12.8	0	12.8	0	12.8	0	12.8	0	12.8	0.00
2	9.2	10.83	17.75	7.95	13.60	9.60	4.35	9.09	1.19	9.22	0.21
3	8.8	9.89	12.33	8.56	2.70	9.57	8.79	9.41	6.91	8.88	0.96
4	9.2	9.43	2.48	9.11	0.99	9.55	3.77	9.50	3.23	8.80	4.33
5	8.5	9.21	8.33	9.58	12.72	9.52	12.00	9.53	12.15	9.00	5.89
6	9.3	9.10	2.13	9.97	7.20	9.49	2.08	9.54	2.61	9.41	1.22
7	10.3	9.05	12.13	10.26	0.38	9.47	8.08	9.54	7.40	9.60	6.81
8	9.4	9.03	3.98	10.44	11.11	9.44	0.44	9.52	1.31	9.54	1.52
9	9.3	9.01	3.07	10.50	12.96	9.41	1.23	9.50	2.17	9.93	6.76
10	11.2	9.01	19.57	10.43	6.91	9.39	16.17	9.47	15.41	10.95	2.21
11	12.2	9.01	26.18	10.19	16.48	9.36	23.26	9.44	22.59	11.64	4.60
12	11	9.00	18.14	9.77	11.14	9.34	15.12	9.41	14.46	11.04	0.35
13	9	9.00	0.04	9.16	1.74	9.31	3.45	9.37	4.14	9.41	4.59
14	8.2	9.00	9.80	8.31	1.34	9.28	13.22	9.33	13.83	7.76	5.42
15	6.4	9.00	40.68	7.20	12.55	9.26	44.67	9.29	45.21	6.78	5.94
MAPE			11.77		7.45		10.44		10.17		3.39
16	7.5	9.00	20.05	5.80	22.63	9.23	23.10	9.25	23.36	6.65	11.34
17	6.8	9.00	32.40	4.07	40.17	9.21	35.40	9.21	35.42	7.14	5.07
18	7.5	9.00	20.05	1.96	73.90	9.18	22.42	9.16	22.20	7.86	4.78
19	7.7	9.00	16.93	−0.58	107.55	9.16	18.91	9.12	18.44	8.44	9.61
20	8.5	9.00	5.92	−3.60	142.40	9.13	7.42	9.07	6.76	8.77	3.14
MAPE			19.07		77.33		21.45		21.24		6.79

4.2. Case Study 2. The raw data used in case study 2 are collected from the gas well B41 in Sichuan, China. Twenty points of monthly gas production ($10^4 m^3$) are listed in Table 3. The first 15 points are used to build the models, and the last 5 points are used for testing. The Gaussian kernel (16) is also used to build the KARGM(1, 1) model.

In this case, the kernel parameter σ and the regularized parameter γ are tuned using 5-fold cross validation, and the results of the MAPE for each pair of σ and γ are plotted in Figure 4. The minimum MAPE is found at the point $\sigma = 2.0$ and $\gamma = 11$; thus this pair will be used to build the KARGM(1, 1) in this case.

The predicted values by the KARGM(1, 1), ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NBGMOP are listed in Table 4, along with the absolute percentage error of each point and the MAPE of each model. The results are also plotted in Figure 3.

It can be seen in Table 4 that the minimum MAPE for fitting of ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NBGMOP is more than two times larger than that of KARGM(1, 1). The minimum MAPE for prediction of ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NBGMOP is more than three times larger than that of the KARGM(1, 1). In Figure 3 it is also shown that the predicted values of KARGM(1, 1) are very close to the real values and the KARGM(1, 1) can accurately catch the overall trend of the monthly production of B41; but the distances between the predicted values of the other four models are still very large. In this case study it is clearly that the KARGM(1, 1) has be best performance.

5. Conclusions

In this paper, a novel kernel based grey model (KARGM(1, 1)) has been proposed, and its effectiveness has been assessed in the two case studies of the gas well production forecasting in comparison to the existing discrete grey models, including the ARGM(1, 1), NDGM(1, 1, k), DGM(1, 1), and NBGMOP. The results of the case studies have shown that the novel KARGM(1, 1) model outperformed the other three discrete grey models.

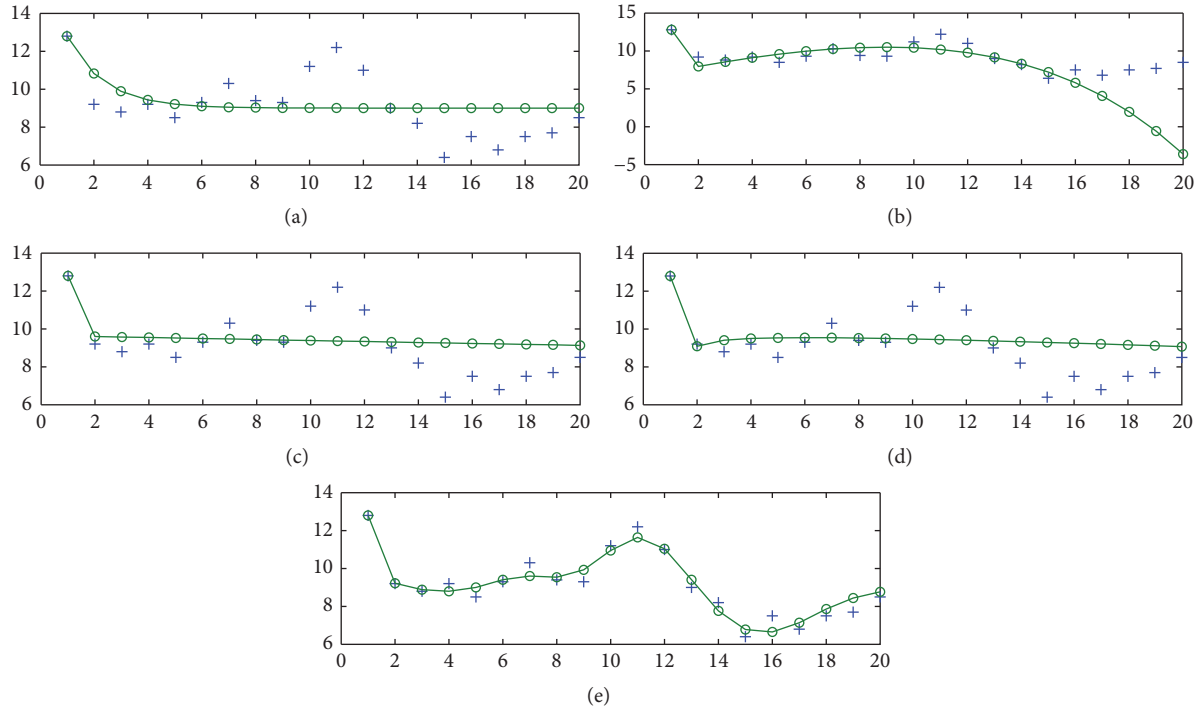

Figure 3: Prediction results of the monthly gas production of B41 by the prediction models. The raw data are plotted using the mark "+" and the predicted values are plotted using the solid line and the mark "o". (a) ARGM(1, 1); (b) NDGM(1, 1, k); (c) DGM(1, 1); (d) NGBMOP; (e) KARGM(1, 1).

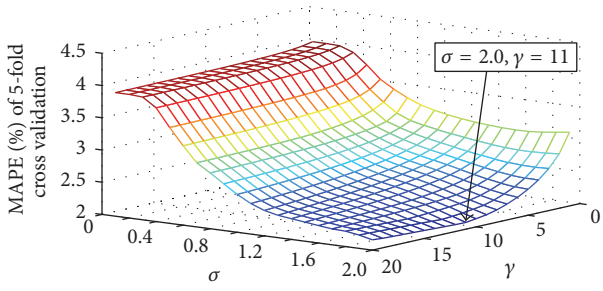

Figure 4: The MAPE of cross validation of KARGM(1, 1) with different σ and γ using the 5-fold cross validation in Case Study 2.

Essentially, the existing ARGM(1, 1), NDGM(1, 1, k), and DGM(1, 1) are linear model, and their solutions are all exponential functions. And this is also the reason why they only produce simple curves in the case studies, which cannot be used to express complex time series. The NBGMOP is a nonlinear model, but its applicability is still limited according to the results shown in the case studies. On the other hand, the KARGM(1, 1) model contains a general nonlinear function $w^T\varphi$, which is expressed by a proper kernel function as shown in (19) and determined by the raw data. Thus the KARGM(1, 1) model can be more effective to deal with complex time series.

What is more, this paper has illustrated a new way of building the discrete grey model using the kernel method and also proved the effectiveness of this new methodology. Thus, the relative researches can be carried out based on this research in the future.

Appendix

A. The KKT Conditions for Equality Constrained Convex Minimization

Consider the equality constrained optimization problem

$$\begin{aligned} &\min \quad f(x) \\ &\text{s.t.} \quad h_j(x) = 0, \quad j = 1, 2, \ldots, m, \end{aligned} \tag{A.1}$$

where the objective function f and the constraint functions $h_j(x) = 0$ $(j = 1, 2, \ldots, m)$ are supposed to be continuously differentiable at point x^*. One defines the Lagrangian function as

$$\mathscr{L}(x; \boldsymbol{\lambda}) = f(x) - \sum_{j=1}^{m} \lambda_j h_j(x), \tag{A.2}$$

where the $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_m]^T$ are the Lagrangian multipliers. The point x^* and the Lagrangian multipliers must satisfy the KKT conditions if the point x^* is a local minimum. In this case, the KarushCKuhnCTucker (KKT) conditions can be interpreted using the Lagrangian function as

$$\begin{aligned} &\frac{\partial \mathscr{L}(x; \boldsymbol{\lambda})}{\partial x} = 0 \\ &\frac{\partial \mathscr{L}(x; \boldsymbol{\lambda})}{\partial \lambda_j} = 0, \quad j = 1, 2, \ldots, m. \end{aligned} \tag{A.3}$$

The KKT conditions are sufficient and necessary to the optimization problem (A.1) if the objective function $f(x)$ is convex and the constrained functions $h_j(x) = 0$ ($j = 1, 2, \ldots, m$) are linear functions. In this case the local minimum x^* which satisfies the KKT conditions is a global minimum; that is, the solution of the minimization problem (A.1) is equivalent to the solution of the KKT conditions (A.3). (More details of the KKT conditions can be seen in [32], pages: 243–245.)

B. A Brief Summary of the 5-Fold Cross Validation Used in the Case Studies

With a given time series $\{x(1), x(2), \ldots, x(n)\}$, the 5-fold cross validation can be described as follows.

Step 1. Divide the original time series into 5 subsets randomly, and mark them as $S_1, S_2, \ldots, S_5$.

Step 2. Build the KARGM(1, 1) model using four subsets and the rest subset is used for validation. (For example, the first time, the KARGM(1, 1) model is built on the subset S_1, S_2, S_3, S_4, and the subset S_5 is used for validation. This procedure will be repeated for 5 times, and all the subset will be used for validation.) Compute the MAPE for all the validation procedures.

Step 3. For $\sigma = 0.1, 0.2, \ldots, 2$ and $\sigma = 1, 2, \ldots, 20$, repeat Step 2 and store the values of the MAPE with all the pairs of σ and γ.

Step 4. Output the values of optimal σ and γ corresponding to the minimum MAPE.

Competing Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by the Doctoral Research Foundation of Southwest University of Science and Technology (no. 16zx7140).

References

[1] J.-L. Deng, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.

[2] X. Ma, D. Luo, X.-F. Ding, and J.-J. Zhou, "An algorithm based on the GM(1,1) model on increasing oil production of measures operation for a single well," in *Proceedings of the 24th IEEE International Conference on Grey Systems and Intelligent Services (GSIS '13)*, pp. 158–160, Macao, China, November 2013.

[3] X. Ma and Z.-B. Liu, "Predicting the cumulative oil field production using the novel grey ENGM model," *Journal of Computational and Theoretical Nanoscience*, vol. 13, no. 1, pp. 89–95, 2016.

[4] D. Chenyan, Z. Songtao, and D. Julong, "Numerical mapping in DNA sequences and analysis of the genetic information by GM (1, N)," *Journal of Grey System*, vol. 24, no. 3, 2012.

[5] Q.-F. Li, Y.-G. Dang, and Z.-X. Wang, "Analysis of the regional coordination development systems based on GRA and GM(1,N)," *Journal of Grey System*, vol. 24, no. 1, pp. 95–100, 2012.

[6] T. L. Tien, "A new grey prediction model FGM (1, 1)," *Mathematical and Computer Modeling*, vol. 49, no. 7-8, pp. 1416–1426, 2009.

[7] C.-K. Chen and T.-L. Tien, "The indirect measurement of tensile strength by the deterministic grey dynamic model DGDM(1, 1, 1)," *International Journal of Systems Science*, vol. 28, no. 7, pp. 683–690, 1997.

[8] J. Cui, S.-F. Liu, B. Zeng, and N.-M. Xie, "A novel grey forecasting model and its optimization," *Applied Mathematical Modelling*, vol. 37, no. 6, pp. 4399–4406, 2013.

[9] D. Q. Truong and K. K. Ahn, "An accurate signal estimator using a novel smart adaptive grey model SAGM(1,1)," *Expert Systems with Applications*, vol. 39, no. 9, pp. 7611–7620, 2012.

[10] N.-M. Xie and S.-F. Liu, "Discrete grey forecasting model and its optimization," *Applied Mathematical Modelling*, vol. 33, no. 2, pp. 1173–1186, 2009.

[11] N.-M. Xie, S.-F. Liu, Y.-J. Yang, and C.-Q. Yuan, "On novel grey forecasting model based on non-homogeneous index sequence," *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 5059–5068, 2013.

[12] X. Ma and Z. Liu, "Predicting the oil field production using the novel discrete GM(1,N) model," *Journal of Grey System*, vol. 27, no. 4, pp. 63–73, 2015.

[13] X. Ma and Z.-B. Liu, "Research on the novel recursive discrete multivariate grey prediction model and its applications," *Applied Mathematical Modelling*, vol. 40, no. 7-8, pp. 4876–4890, 2016.

[14] C.-I. Chen, H. L. Chen, and S.-P. Chen, "Forecasting of foreign exchange rates of Taiwan's major trading partners by novel nonlinear Grey Bernoulli model NGBM(1, 1)," *Communications in Nonlinear Science and Numerical Simulation*, vol. 13, no. 6, pp. 1194–1204, 2008.

[15] H.-T. Pao, H.-C. Fu, and C.-L. Tseng, "Forecasting Taiwans major stock indices by the Nash nonlinear grey Bernoulli model," *Energy*, vol. 40, no. 1, pp. 400–409, 2012.

[16] H.-T. Pao, H.-C. Fu, and C.-L. Tseng, "Forecasting of CO_2 emissions, energy consumption and economic growth in China using an improved grey model," *Energy*, vol. 40, no. 1, pp. 400–409, 2012.

[17] Z.-X. Wang, K. W. Hipel, Q. Wang, and S.-W. He, "An optimized NGBM(1,1) model for forecasting the qualified discharge rate of industrial wastewater in China," *Applied Mathematical Modelling*, vol. 35, no. 12, pp. 5524–5532, 2011.

[18] Z.-X. Wang, "Nonlinear grey prediction model with convolution integral NGMC (1, n) and its application to the forecasting of China's industrial SO_2 emissions," *Journal of Applied Mathematics*, vol. 2014, Article ID 580161, 9 pages, 2014.

[19] B. Zeng and S.-F. Liu, "Direct modeling approach of DGM (1, 1) with approximate non-homogeneous exponential sequence," *System Engineering Theory and Practice*, vol. 31, no. 2, pp. 297–301, 2011.

[20] L. Wu, S. Liu, H. Chen, and N. Zhang, "Using a novel grey system model to forecast natural gas consumption in China," *Mathematical Problems in Engineering*, vol. 2015, Article ID 686501, 7 pages, 2015.

[21] V. N. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, 1998.

[22] B. Schölkopf, A. Smola, and K. Müller, "Kernel principal component analysis," in *Proceedings of the 7th International Conference on Artificial Neural Networks (ICANN '97)*, pp. 583–588, Springer, Lausanne, Switzerland, 1997.

[23] B. Scholkopft and K.-R. Mullert, "Fisher discriminant analysis with kernels," in *Proceedings of the Neural Networks for Signal Processing IX*, vol. 1, pp. 41–48, August 1999.

[24] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *The Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.

[25] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.

[26] T. Van Gestel, J. A. K. Suykens, G. Lanckriet, A. Lambrechts, B. De Moor, and J. Vandewalle, "Bayesian framework for least-squares support vector machine classifiers, gaussian processes, and kernel fisher discriminant analysis," *Neural Computation*, vol. 14, no. 5, pp. 1115–1147, 2002.

[27] J. A. K. Suykens, T. Van Gestel, J. Vandewalle, and B. De Moor, "A support vector machine formulation to PCA analysis and its kernel version," *IEEE Transactions on Neural Networks*, vol. 14, no. 2, pp. 447–450, 2003.

[28] J. A. K. Suykens and J. Vandewalle, "Recurrent least squares support vector machines," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, no. 7, pp. 1109–1114, 2000.

[29] L. R. Medsker and L. C. Jain, *Recurrent Neural Networks: Design and Applications*, CRC Press, Boca Raton, Fla, USA, 2001.

[30] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.

[31] M. Espinoza, J. A. K. Suykens, and B. De Moor, "Partially linear models and least squares support vector machines," in *Proceedings of the 43rd IEEE Conference on Decision and Control (CDC '04)*, vol. 4, pp. 3388–3393, IEEE, December 2004.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.