

Research Article

Autoregressive Prediction with Rolling Mechanism for Time Series Forecasting with Small Sample Size

Zhihua Wang, Yongbo Zhang, and Huimin Fu

Research Center of Small Sample Technology, School of Aeronautical Science and Engineering, Beihang University, Beijing 100191, China

Correspondence should be addressed to Yongbo Zhang; zhang19840504@163.com

Received 20 February 2014; Revised 28 April 2014; Accepted 30 April 2014; Published 9 June 2014

Academic Editor: M. I. Herreros

Copyright © 2014 Zhihua Wang 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.

Reasonable prediction makes significant practical sense to stochastic and unstable time series analysis with small or limited sample size. Motivated by the rolling idea in grey theory and the practical relevance of very short-term forecasting or 1-step-ahead prediction, a novel autoregressive (AR) prediction approach with rolling mechanism is proposed. In the modeling procedure, a new developed AR equation, which can be used to model nonstationary time series, is constructed in each prediction step. Meanwhile, the data window, for the next step ahead forecasting, rolls on by adding the most recent derived prediction result while deleting the first value of the former used sample data set. This rolling mechanism is an efficient technique for its advantages of improved forecasting accuracy, applicability in the case of limited and unstable data situations, and requirement of little computational effort. The general performance, influence of sample size, nonlinearity dynamic mechanism, and significance of the observed trends, as well as innovation variance, are illustrated and verified with Monte Carlo simulations. The proposed methodology is then applied to several practical data sets, including multiple building settlement sequences and two economic series.

1. Introduction

Many planning activities require prediction of the behavior of variables, such as economic, financial, traffic, and physical ones [1, 2]. Makridakis et al. [3] concluded that predictions supported the strategic decisions of organizations, which in turn sustained a practical interest in forecasting methods. To obtain a reasonable prediction, certain laws governing the phenomena must be discovered based on either natural principles or real observations [4]. However, seeking available natural principles, in the real world, is extremely difficult either in the physical system or in the generalized system. Thus, forecasting the future system development directly from the past and current datum becomes a feasible means [5]. Pinson [6] further pointed out that statistical models based on historical measurements only, though taking advantage of physical expertise at hand, should be preferred for short lead time ahead forecasts. Since time series method can be used to forecast the future based on historical observations [3], it has been widely utilized to model forecasting systems when there is not much information about the generation

process of the underlying variable and when other variables provide no clear explanation about the studied variable [7].

In the literature, considerable efforts have been made for time series prediction, and special attention should be given to time series approaches [8, 9], regression models [10, 11], artificial intelligence method [12], and grey theory [13]. Time series methods mainly involve the basic models of autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), and Box-Jenkins. Most analyses are based on the assumption that the probabilistic properties of the underlying system are time-invariant; that is, the focused process is steady. Although this assumption is very useful to construct simple models, it does not seem to be the best strategy in practice. The reason is that systems with time-varying probabilistic properties are common in practical engineering. Although we can construct regression model with a few data points, accurate prediction cannot be achieved from the simplicity of linear model. Therefore, linear methodology is sometimes inadequate for situations where the relationships between the samples are not linear

with time, and then artificial intelligence techniques, such as expert system and neural network, have been developed [14]. However, abundant prediction rule and practical experience from specific experts and large historical data banks are requisite for precise forecasting. Meanwhile, grey theory constructs a grey differential equation to predict with as few as four data points by accumulated generating operation technique. Though the grey prediction model has been successfully applied in various fields and has demonstrated satisfactory results, its prediction performance still could be improved, because the grey forecasting model is constructed of exponential function. Consequently, it may derive worse prediction precise when more random data sets exist.

Furthermore, recent developments in Bayesian time series analysis have been facilitated by computational methods such as Markov chain Monte Carlo. Huerta and West [15, 16] developed a Markov chain Monte Carlo scheme based on the characteristic root structure of AR processes. McCoy and Stephens [17] extended this methodology by allowing for ARMA components and adopting a frequency domain approach. The fruitfully developed applications of predictability methods to physiological time series are also worth noticing. The K-nearest neighbor approach is conceptually simple to pattern recognition problems, where an unknown pattern is classified according to the majority of the class memberships of its K-nearest neighbors in the training set [18, 19]. Moreover, local prediction, proposed by Farmer and Sidorowich [4], derives forecasting based on a suitable statistic of the next values assigned L previous samples. Porta et al. [20] further established an improved method to allow one to construct a reliable prediction of short biological series. The method is especially suitable when applied to signals that are stationary only during short periods (around 300 samples) or to historical series the length of which cannot be enlarged.

We place ourselves in a parametric probabilistic forecasting framework under small sample size, for which simple linear models are recommended, such as AR model and grey prediction model, because these simple linear models are frequently found to produce smaller prediction errors than techniques with complicated model forms due to their parsimonious form [21]. However, two issues should be noticed. On one hand, linear approaches may output unsatisfactory forecasting accuracy when the focused system illustrates a nonlinear trend. This new arising problem indicates that model structure is instable. According to Clements and Hendry [22], this structural instability has become a key issue, which dominates the forecasting performance. Mass work on model structural change has been conducted [23, 24]. To settle this problem, similar with the basic idea of K-nearest neighbor and local prediction approaches, many scholars have recommended using only recent data to increase future forecasting accuracy if chaotic data exist. Based on this point of view, grey model GM(1,1) rolling model, called rolling check, was proposed by Wen [25]. In this approach, the GM(1,1) model is always built on the latest measurements. That is, on the basis of $x^{(0)}(k), x^{(0)}(k+1), x^{(0)}(k+2)$, and $x^{(0)}(k+3)$, the next data point $x^{(0)}(k+4)$ can be forecasted.

The first datum is always shifted to the second in the following prediction step; that is, $x^{(0)}(k+1), x^{(0)}(k+2), x^{(0)}(k+3)$, and $x^{(0)}(k+4)$ are used to predict $x^{(0)}(k+5)$. The same technique called grey prediction with rolling mechanism (GPRM) is established for industrial electricity consumption forecasted by Akay and Atak [26]. However, this rolling grey model can only be utilized in one-step prediction. That is, the one-step-ahead value is always obtained by the observed data. On the other hand, an AR model can only be established for time series that satisfies the stationarity condition; that is, a stationary solution to the corresponding AR characteristic equation exists if and only if all roots exceed unity in absolute value (modulus) [27]. Consequently, AR models cannot be established for modeling nonstationary time series.

In addition, we should also notice that very short-term forecasting or 1-step-ahead prediction is of significant importance in various applications. This is clear in economy. Now we take wind power forecasting and heart period given systolic arterial pressure variations prediction as two further examples to elucidate the practical relevance of very short-term forecasting or 1-step-ahead prediction in different fields of science. Since transmission system operators require to operate reserves optimally for the continuous balance of the power system, very short-term predictions of the power fluctuations at typical time steps of 1, 10, and 30 min are recognized as a current challenge [28]. More specifically for the case of Denmark, 10-min lead time has been defined as most important, because power fluctuations at this timescale are those that most seriously affect the balance in the power system [29]. The characterization of the relation between spontaneous heart period (HP) and systolic arterial pressure (SAP) variabilities provides important information about one of the short-term neural reflexes essentially contributing to cardiovascular homeostasis, namely, baroreflex [30]. The difficulty in predicting HP given SAP variations depends on the forecasting time: the longer the forecasting time is, the more unreliable the prediction of HP given SAP changes is and the larger the information carried by HP given SAP modifications is [31]. More specifically, 1-step-ahead predictive information of heart period series given systolic arterial pressure dynamics better correlates with the activity of autonomic nervous system in cardiovascular physiology.

Motivated by the GPRM approach and the practical relevance of very short-term forecasting or 1-step-ahead prediction elucidated above, the first objective of this study is to construct a novel prediction model with the rolling mechanism to improve the forecasting precision. Therefore, the sample data set and model parameters are evolved in each prediction step. The second objective of this study is to develop an autoregressive model that can be used to model nonstationary time series. Consequently, this autoregression is different from the AR model in the time series analysis literature. We also call it autoregression, because the current value of the series is also a linear combination of several most recent past values of itself plus an "innovation" term that incorporates everything new in the series that is not explained by the past values, but it can be used to model nonstationary time series.

The remainder of the paper is organized as follows. We start from short introduction of AR method in Section 2. Section 3 involves the basic idea of ARPRM, model establishment, and parameter estimation. The best unbiased property of prediction is also discussed. In Section 4, simulation studies, including general verification, performance impact of sample size, system nonlinearity dynamic mechanism, and variance, are conducted to assess the validity of the approach, whilst Section 5 contains applications to real data sets, including building settlement sequences and economic series. In Section 6, we close this paper with a discussion and conclusion.

2. AR Model Introduction

Let x_1, x_2, \dots, x_n denote a time series with sample size n . Autoregressive models, in the literature, are created with the idea that the present value of the series x_t can be explained as a function of p past values $x_{t-1}, x_{t-2}, \dots, x_{t-p}$, where p determines the number of steps into the past needed to forecast the current value. The AR(p) model can be given as [8]

$$\Phi(L)x_t = \varphi_0 + \varepsilon_t \quad \varepsilon_t \sim \text{NID}[0, \sigma_\varepsilon^2], \quad (1)$$

where $\Phi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$ specifies the lag-polynomial with model order p and φ_0 is a constant relating to series mean. It is well known in the literature that a stationarity condition has to be satisfied for the AR(p) process; that is, subject to the restriction that ε_t is independent of x_{t-1}, x_{t-2}, \dots and that $\sigma_\varepsilon^2 > 0$, a stationary solution to (1) exists if and only if the root of the AR characteristic equation exceeds 1 in absolute value (modulus).

According to least-squares method, model parameters can be calculated by

$$\varphi_0 = \bar{x}_0 - \sum_{i=1}^p \varphi_i \bar{x}_i, \quad (2)$$

$$(\varphi_1, \varphi_2, \dots, \varphi_p)^T = (\mathbf{L}_{\eta\eta})_{p \times p}^{-1} (\mathbf{L}_\eta)_{p \times 1},$$

where mean $\bar{x}_i, i = 0, 1, \dots, p$ can be obtain by

$$\bar{x}_i = \frac{1}{n-p} \sum_{t=p+1}^n x_{t-p} \quad i = 0, 1, \dots, p \quad (3)$$

and $\mathbf{L}_{\eta\eta} = (S_{ij})_{p \times p}$ is a p -order matrix, and $\mathbf{L}_\eta = (S_1, S_2, \dots, S_p)^T$ is a p -order column vector, whose elements can be determined through

$$S_{ij} = \sum_{t=p+1}^n (x_{t-i} - \bar{x}_i)(x_{t-j} - \bar{x}_j) \quad i, j = 1, 2, \dots, p, \quad (4)$$

$$S_i = \sum_{t=p+1}^n (x_t - \bar{x}_0)(x_{t-i} - \bar{x}_i) \quad i = 1, 2, \dots, p.$$

Based on the estimated coefficients $\hat{\varphi}_0, \hat{\varphi}_1, \hat{\varphi}_2, \dots, \hat{\varphi}_p$, the AR(p) prediction equation can be determined as

$$\hat{x}_{n+l|n} = \hat{\varphi}_0 + \sum_{i=1}^p \hat{\varphi}_i x_{n+l-i|n}, \quad (5)$$

where $\hat{x}_{n+l|n}$ denotes the l -step-ahead prediction at time $n+l$.

3. ARPRM Model Construction

3.1. Basic Idea of ARPRM Method. Based on initial observational sample x_1, x_2, \dots, x_n , the 1-step-ahead ARPRM(p_1) model can be established as

$$x_t = \eta_{10} + \sum_{i=1}^{p_1} \eta_{1i} x_{t-i} + \varepsilon_{1t}, \quad (6)$$

where p_1 is the 1-step-ahead ARPRM model order that can be determined by Akaike Information criterion rule [32], $\{\varepsilon_{1t}\}$ is white noise series of 1-step-ahead ARPRM model, and η_{1i} ($i = 0, 1, \dots, p_1$) are autoregressive coefficients of the 1-step-ahead ARPRM equation. Then the 1-step-ahead prediction result can be calculated by

$$\hat{x}_{n+1|n} = \eta_{10} + \sum_{i=1}^{p_1} \eta_{1i} x_{n+1-i}. \quad (7)$$

Then for the 2-step-ahead forecasting, according to the basic idea described in Section 3.1, a new sample $x_2, x_3, \dots, x_n, \hat{x}_{n+1|n}$ is first constructed, and the 2-step-ahead ARPRM(p_2) model can be found as

$$x_t = \eta_{20} + \eta_{21} \hat{x}_{n+1|n} + \sum_{i=2}^{p_2} \eta_{2i} x_{t-i} + \varepsilon_{2t}, \quad (8)$$

where $p_2, \{\varepsilon_{2t}\}$, and η_{2i} ($i = 0, 1, \dots, p_2$) are model order, white noise series, and the autoregressive coefficients of the 2-step-ahead ARPRM equation, respectively.

Analogically, considering the l -step-ahead prediction, one can first form a new sample with general notations $x_l^*, x_{l+1}^*, \dots, x_{n+l-1}^*$ according to the rolling mechanism mentioned above, where

$$x_k^* = \begin{cases} x_k & k \leq n \\ \hat{x}_{k|n} & k > n. \end{cases} \quad (9)$$

It can be seen that x_k^* is an original observation data if $k \leq n$, while it will be a prediction result of a previous step when $k > n$. Accordingly, for the 1st-step-ahead prediction (i.e., $l = 1$), the sample will be $x_1^*, x_2^*, \dots, x_n^*$, that is x_1, x_2, \dots, x_n based on (9). For the 2nd-step-ahead prediction (i.e., $l = 2$), the sample will be $x_2^*, x_3^*, \dots, x_{n+1}^*$, that is $x_2, x_3, \dots, x_n, \hat{x}_{n+1|n}$ based on (9). This is consistent with the samples described above. Then based on this new sample $x_l^*, x_{l+1}^*, \dots, x_{n+l-1}^*$ ARPRM (p_l) equation can be established as

$$x_t = \eta_{l0} + \sum_{i=1}^{p_l} \eta_{li} x_{t-i}^* + \varepsilon_{lt}, \quad (10)$$

where p_l , $\{\varepsilon_{it}\}$, and η_{li} ($i = 0, 1, \dots, p_l$) are model order, white noise series, and autoregressive coefficients of the l -step-ahead ARPRM model, respectively. Accordingly, we can obtain the l -step-ahead prediction result by

$$\hat{x}_{n+l|n} = \eta_{l0} + \sum_{i=1}^{p_l} \eta_{li} x_{n+l-i}^* \quad (11)$$

From the above description, we can see that a new autoregressive prediction model, with its own model order and autoregressive parameters, is constructed in each step. We should note that the existing AR model in time series analysis literature can only model the series that satisfies the stationarity requirement. In this study, however, we only adopt the basic idea of autoregression in each prediction step. This autoregressive model in each forecasting step, can be used to model series that does not satisfy the stationarity condition. It regresses to the traditional AR model when the focused time series is a stationary one.

While aforementioned procedure is considered to add one forecasting result and delete one sample value in each prediction step, the adding and deleting number can be unequal; that is, one can add one and delete more sample values at each prediction step without departing from the spirit of the proposed method. Facts indicate that redefinition for each step can modify the ARPRM model coefficients in each prediction step according to the metabolic sample, and the prediction accuracy can consequently be effectively improved.

3.2. Parameter Estimation. First, model order p_l for the l -step-ahead ARPRM (p_l) (11) can be determined by Akaike information criterion rule [32]. When p_l increases from one, the calculated result should enable the flowing formula

$$\text{AIC}(p_l) = \ln \hat{\sigma}_{le}^2 + \frac{2p_l}{n+l-1} \quad (12)$$

to achieve its minimum.

Then, according to least-squares method, the autoregressive coefficient η_{l0} can be calculated by

$$\eta_{l0} = \bar{x}_{l0} - \sum_{i=1}^{p_l} \eta_{li} \bar{x}_{li}, \quad (13)$$

where mean \bar{x}_{li} ($i = 0, 1, \dots, p_l$) can be obtain by

$$\bar{x}_{li} = \frac{1}{n-p_l} \sum_{t=l+p_l}^{n+l-1} x_{t-i}^* \quad i = 0, 1, \dots, p_l, \quad (14)$$

where x_t^* is derived by (9). The autoregressive coefficients η_{li} , $i = 1, \dots, p_l$, can be derived by

$$(\eta_{l1}, \eta_{l2}, \dots, \eta_{lp_l})^T = (\mathbf{L}_{l\eta\eta})_{p_l \times p_l}^{-1} (\mathbf{L}_{l\eta})_{p_l \times 1}, \quad (15)$$

where $(\mathbf{L}_{l\eta\eta})_{p_l \times p_l} = (S_{lij})_{p_l \times p_l}$ is a p_l -order matrix and $(\mathbf{L}_{l\eta})_{p_l \times 1} = (S_{li1}, S_{li2}, \dots, S_{lip_l})^T$ is a p_l -order column vector. The elements can be determined through

$$S_{lij} = \sum_{t=p_l+l}^{n+l-1} (x_{t-i}^* - \bar{x}_{li}) (x_{t-j}^* - \bar{x}_{lj}) \quad i, j = 1, 2, \dots, p_l, \quad (16)$$

$$S_{li} = \sum_{t=p_l+l}^{n+l-1} (x_t^* - \bar{x}_{l0}) (x_{t-i}^* - \bar{x}_{li}) \quad i = 1, 2, \dots, p_l.$$

In addition, the estimator of $\sigma_{le}^2 = \text{Var}(\varepsilon_{lt})$ can be obtained by

$$\hat{\sigma}_{le}^2 = \frac{1}{n+l-p_l-1} \sum_{t=p_l+l}^{n+l-1} \left(x_t^* - \eta_{l0} - \sum_{i=1}^{p_l} \eta_{li} \bar{x}_{li} \right)^2. \quad (17)$$

Then, the l -step-ahead prediction value $\hat{x}_{n+l|n}$ can be determined by

$$\hat{x}_{n+l|n} = \eta_{l0} + \sum_{i=1}^{p_l} \eta_{li} x_{n+l-i|n}^*. \quad (18)$$

And its mean square error $\sigma_{n+l|n}^2$ can be derived by

$$\sigma_{n+l|n}^2 = E(x_{n+l} - \hat{x}_{n+l|n})^2 = \sum_{i=1}^l \alpha_{li}^2 \sigma_{(l+1-i)\varepsilon}^2, \quad (19)$$

where $\eta_{li} = 0$ when $i > p_l$ and

$$\alpha_{k1} = 1 \quad k = 1, 2, \dots, l,$$

$$\alpha_{ki} = \sum_{j=1}^{i-1} \eta_{kj} \alpha_{(k-j)(i-j)} \quad k = 2, 3, \dots, l; \quad i = 2, 3, \dots, k. \quad (20)$$

3.3. Best Unbiased Nature. The prediction result obtained by the proposed ARPRM method shows best unbiased property, which can be deduced from the least-square error forecasting method.

4. Simulation Study

ARPRM is a general approach for stochastic and unstable time series forecasting. In this section, we focus on the robustness of its performance. The general simulation study is firstly conducted to verify the general reasonableness of the ARPRM prediction method with 4 different models. ARPRM performance while varying sample size and innovation variance is also checked, because these two aspects are both of significant importance for small sample prediction problem. Meanwhile, statistical analysis has been performed to demonstrate the significance of the observed trends. Results show that the trends do affect the performance of the proposed method, and we further found that the trend nonlinearity plays a much more important role in forecasting performance than trend significance. Consequently, we illustrate the nonlinearity

TABLE 1: Model specifications by experiments for general verification.

Experiment	Model specifications	
1	$x_t = 0.3e^t + 0.2x_{t-1} + \varepsilon_t$	$\varepsilon_t \sim \text{NID}[0, 1]$
2	$x_t = t^2 + 2x_{t-1} + \varepsilon_t$	$\varepsilon_t \sim \text{NID}[0, 1]$
3	$x_t = 10 \ln t + 1.5x_{t-1} + \varepsilon_t$	$\varepsilon_t \sim \text{NID}[0, 1]$
4	$x_t = te^{0.6t} + 0.3x_{t-1} + \varepsilon_t$	$\varepsilon_t \sim \text{NID}[0, 1]$

dynamics monitoring in this simulation study. In addition, as ARPRM model has to be constructed independently based on the metabolism sample in each prediction step, it can be concluded that this is a nonlinear forecasting with a stepwise linear form. Accordingly, we also conduct a simulation study to monitor the effect of this rolling mechanism, which can also illustrate the ability of the model to follow nonlinear dynamics.

4.1. General Verification. The proposed ARPRM approach is quite general. We consider and study four simulation models to verify the general reasonableness of the ARPRM prediction method. The models assumed in each of the experiments are summarized in Table 1. We also considered an AR(p) specification with $p > 1$ to study the forecasting effect of the established model, and the results were very similar to those reported in this paper. Since we focus on the rationality and validity of ARPRM approach in small sample forecasting problems, we consider conducting $l = 5$ step ahead prediction based on a observational data set with sample size $n = 10$.

Simulation results are analyzed by the index of percent relative error $\text{err} = |x_t - \hat{x}_t| \times 100/x_t$, which is presented in Table 2. The results are based on 50000 Monte Carlo simulations with innovations drawn from an IID Gaussian distribution. We consider the mean percent relative error result of each forecasting step. It can be concluded, from Table 2, that the percent relative error is upward as forecasting step l increasing from 1 to 5. It can be seen that the 1-step-ahead prediction is so accurate that the minimum percent relative error is 0.041%. The maximum percent relative error for the 5-step-ahead forecasting is 7.90%, which is still considered to be good and shows a promise for future applications.

4.2. Sample Size and Nonlinearity Dynamics Monitoring. Simulation studies show that the sample size and class of nonlinear mechanism generating dynamics both play an important role in its prediction performance, especially in small sample problems. Consequently, the sample size checking and class of nonlinearity dynamics monitoring are both included in this section. Two simulation models with intuitive different nonlinear mechanism, listed in Table 3, are considered. To make the illustration more distinct, scale of the generated data set is controlled. The reason is that the same absolute prediction error will lead to a larger relative error for data of lower order of magnitude. Let $f(t) = e^{0.3t}$ be a deterministic function (dynamics generation mechanism of Experiment 1 in Table 3 without white noise); then it is

obvious that a linear relationship exists between current value $f(t)$ and the most recent past value $f(t-1)$; that is, $f(t) = e^{0.3}f(t-1)$. Considering Experiment 2, listed in Table 3, the data generating process is much complicated. An AR(1) series y_t (satisfying the stationarity condition) is firstly generated; then an accumulated sequence $d_t = d_{t-1} + y_t$ is derived by the accumulated generating operation process; finally data set x_t can be obtained by adding a trend $\sqrt{(e^{0.8t} - e^{-0.3t})/2}$ to d_t . Accordingly, it can be seen that the nonlinearity aptitude of Experiment 2 is much higher than that of Experiment 1. The nonlinearity dynamics can be monitored through the prediction performance of these two models.

Furthermore, the impact of sample size is also checked. For comparing, three samples of x_t , (1) $t = 1, 2, \dots, 10$, (2) $t = 2, 3, \dots, 10$, and (3) $t = 3, 4, \dots, 10$, are considered to conduct the same $l = 5$ step ahead prediction from $t = 11$ to $t = 15$. Index of percent relative error $\text{err} = |x_t - \hat{x}_t| \times 100/x_t$, $t = 11, 12, \dots, 15$, listed in Table 4, is also analyzed for clarifying the changes of the statistical properties while varying the sample size. The results are also based on 50000 Monte Carlo simulations with innovations drawn from an IID Gaussian distribution.

From the comparative results summarized in Table 4, the following can be seen. (1) The percent relative error is upward as forecasting step l increasing from 1 to 5 for one experiment model under the same sample size. (2) For either Experiment 1 or Experiment 2, forecasting performance decreases when sample size becomes smaller. The change is distinct, because this is a small sample problem where 10 or less data is used to conduct 5-step-ahead predictions. (3) Comparing percent relative error derived from the two classes of nonlinear mechanism generating dynamics, precision of Experiment 2 is significantly lower than that of Experiment 1. These conclusions are consistent with our general understanding and deduction.

4.3. Variance Check. In order to check whether the performance of the method depends on the variance of the innovation, model of Experiment 1, listed in Table 3 with four different innovation variances, (1) $\varepsilon_t \sim \text{NID}[0, 0.25]$, (2) $\varepsilon_t \sim \text{NID}[0, 2]$, (3) $\varepsilon_t \sim \text{NID}[0, 5]$, and (4) $\varepsilon_t \sim \text{NID}[0, 10]$, is considered for verification and the same $l = 3$ step ahead prediction from $t = 11$ to $t = 13$ is conducted with the same sample size x_1, x_2, \dots, x_{10} for comparison. Percent relative error $\text{err} = |x_t - \hat{x}_t| \times 100/x_t$, $t = 11, 12, 13$, under the four variances, is illustrated in Table 5. Likewise, the results are also based on 50000 Monte Carlo simulations with innovations drawn from an IID Gaussian distribution. Table 5 shows that performance of the ARPRM method decreases with the increasing of variance of the innovation. This is also consistent with our general understanding. The reason is that larger variance means greater uncertainty, which will lead to lower forecasting performance. There is an issue worth noticing. Expectation of generated sample x_1, x_2, \dots, x_{13} is not so large, where $E(x_1) = 1.35$, $E(x_2) = 1.82$, $E(x_3) = 2.46$, $E(x_4) = 3.32$, $E(x_5) = 4.48$, $E(x_6) = 6.05$, $E(x_7) = 8.17$, $E(x_8) = 11.0$, $E(x_9) = 14.9$, $E(x_{10}) = 20.1$, $E(x_{11}) = 27.1$, $E(x_{12}) = 36.6$, and $E(x_{13}) = 49.4$. It can be seen that

TABLE 2: Results of ARPRM forecasting percent relative error for general verification (%).

Percent error	Experiments 1	Experiments 2	Experiments 3	Experiments 4
Step $l = 1$	0.041	0.065	0.230	0.220
Step $l = 2$	0.077	0.176	0.510	0.670
Step $l = 3$	0.119	0.375	0.970	1.510
Step $l = 4$	0.177	0.986	1.750	2.970
Step $l = 5$	0.348	6.743	4.410	7.900

Note: Experiments 1–4 are defined in Table 1.

TABLE 3: Model specifications by experiments for sample size and nonlinearity dynamics monitoring.

Experiment	Model specifications	
1	$x_t = e^{0.3t} + \varepsilon_t$	$\varepsilon_t \sim \text{NID}[0, 1]$
	$y_t = 0.8y_{t-1} + \varepsilon_t$	$\varepsilon_t \sim \text{NID}[0, 1]$
2	$d_t = d_{t-1} + y_t$	$x_t = \sqrt{\frac{e^{0.8t} - e^{-0.3t}}{2}} + d_t$

TABLE 4: Results of ARPRM forecasting percent relative error for sample size and nonlinearity dynamics monitoring (%).

Percent error	Experiment 1 sample Size			Experiment 2 sample Size		
	$n = 10$	$n = 9$	$n = 8$	$n = 10$	$n = 9$	$n = 8$
Step $l = 1$	3.13	3.85	3.93	4.23	4.59	5.39
Step $l = 2$	4.40	5.83	7.44	9.73	11.3	13.5
Step $l = 3$	6.60	9.26	11.6	16.5	21.2	25.5
Step $l = 4$	8.80	12.0	16.8	24.4	34.2	41.9
Step $l = 5$	11.0	15.8	22.9	33.6	52.1	66.2

TABLE 5: Comparative results of ARPRM forecasting percent relative error for different variance ($\times 10^{-5}\%$).

Percent error	Variance of the innovation $\varepsilon_t \sim \text{NID}[0, \sigma^2]$			
	0.25	2	5	10
Step $l = 1$	1.64	4.65	7.43	10.8
Step $l = 2$	2.53	7.08	11.3	16.2
Step $l = 3$	3.70	10.3	16.4	23.2

the innovation $\varepsilon_t \sim \text{NID}[0, 1]$ does make a considerably great randomness for the first few sample values, and bigger innovation variances like 2, 5, 10 lead to lower performance because of the significant uncertainty of the sample data x_1, x_2, \dots, x_{10} . The objective of adopting Experiment 1 as the simulation model is to illustrate the variance impact in a more intuitive and obvious way. Difference between err indexes will not be that distinct when scale of sample data is enlarged; for example, model $x_t = e^{0.5t} + \varepsilon_t$ is used to generate data set.

5. Empirical Applications

In this section, we focus on the practical performance of the proposed ARPRM approach. Our experiments are presented for building settlement prediction illustrated in Section 5.1

and economic forecasting containing two different data sets shown in Section 5.2.

5.1. Building Settlement Prediction. The prediction of future building settlement, especially for high-rising building, is a hot topic in Structural Health Monitoring. However, it is very difficult to develop good mathematical models and thus yield accurate predictions for building settlement, which is caused by the problems of small sample size (settlement observations are relatively few), nonstationary (the statistical properties of the measurement change over time), and nonlinearity (it is difficult to use mathematical prediction models with linear structure) [33].

The application illustrated in this section is the future settlement forecasting for Changchun branch's business building of China Everbright Bank. The building has a steel structure form, with 28 floors above ground, and a base area of 2176 square meters. There are ten observation points marked $K1, K2, \dots, K10$. Observation settlement series $x_t^{(k)}$ ($k = 1, 2, \dots, 10, t = 1, 2, \dots, 18$), corresponding to each observation point $K1, K2, \dots, K10$, was got [34]. See Figure 1 for the data sets. It can be seen that the problems of small sample size, nonlinear, and nonstationary all exist.

The last 3 settlement values are predicted based on the former 15 data. The GM(1,1) model, GM(1,1) rolling model, and AR model with linear interpolation [34] are adopted as reference methods for comparisons. Before AR model construction, the sample size of each data set is increased from 18 to 35 through linear interpolation; that is, a new data $(x_i^{(k)} + x_{i+1}^{(k)})/2$ is inserted between $x_i^{(k)}$ and $x_{i+1}^{(k)}$, $i = 1, 2, \dots, 17, k = 1, 2, \dots, 10$. And this is so called AR model with linear interpolation.

To demonstrate the effectiveness and reasonability, the prediction results are analyzed by the index of average relative percentage error $\text{ARPE} = (1/10) \sum_{k=1}^{10} |x_t^{(k)} - \hat{x}_t^{(k)}| \times 100/x_t^{(k)}$, $t = 16, 17, 18$, where $\hat{x}_t^{(k)}$ is the prediction result corresponding to observation point $k = 1, 2, \dots, 10$. The comparative results are listed in Table 6, which show that the most precise forecast is given by ARPRM, the following one is obtained by AR model with linear interpolation, and a considerably unreasonable accuracy is got by the GM(1,1) model and GM(1,1) rolling model. One problem worth noticing is that, in the AR model with linear interpolation processing procedure, the former 32 values of the interpolation series are utilized to conduct a 5-step-ahead prediction. And the forecasting result of the last 3 settlement values at $t = 16, 17, 18$, which correspond to $t^* = 33, 35, 37$ in the interpolated sequence,

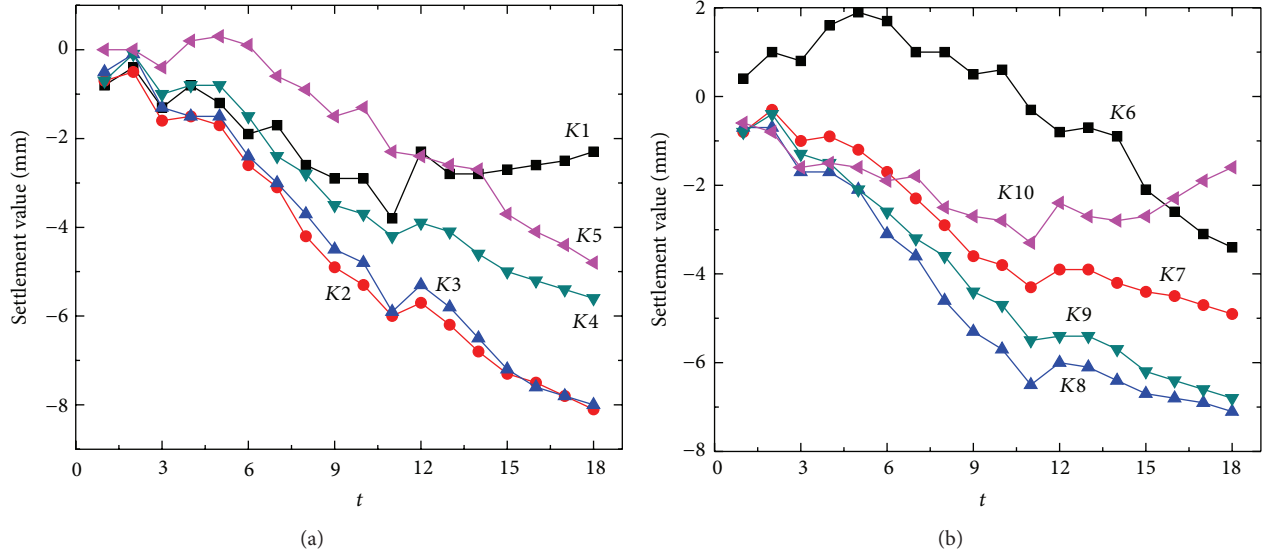


FIGURE 1: (a) Settlement sequence for observational point K1 to K5. (b) Settlement sequence for observational point K6 to K10.

can be derived. In can be seen that the 32nd value, obtained interpolation $x_{32}^{*(k)} = (x_{15}^{(k)} + x_{16}^{(k)})/2$, contains the information of $x_{16}^{(k)}$ that we want to predict. Actually, this interpolation value $x_{32}^{*(k)}$ cannot be utilized in the forecasting.

5.2. Economic Forecasting. In this section, our empirical study is also devoted to the comparison of our forecasting procedure with GM(1,1) model, GM(1,1) rolling model, and AR model with linear interpolation. The economic data sets include Chinese annual power load data [35] and Chinese annual gross domestic product (GDP) data [36] from 1987 to 2002, seen in Figures 2 and 3.

We predict the latest two data with the former 14 for Chinese annual power load and GDP. The index of relative percentage error $RPE = |x_t - \hat{x}_t| \times 100/x_t$ is used to demonstrate the accuracy. The comparative results for each economic data set are listed in Tables 7 and 8. Table 7 shows that ARPRM derives the best prediction accuracy, and GM(1,1) model and GM(1,1) rolling model also give good forecast results for the Chinese annual power load forecast. Although AR model with linear interpolation provides the worst prediction, its accuracy is still acceptable. A similar conclusion with Section 5.1 can be obtained for the Chinese annual GDP prediction results.

Based on the empirical study discussed above, it can be concluded that prediction accuracy of ARPRM and AR model with linear interpolation method is relatively stable. Precision of GM(1,1) model and GM(1,1) rolling model is considerably good when data set show an exponential trend, such as Chinese annual power load. Otherwise, the prediction accuracy of GM(1,1) model and GM(1,1) rolling model cannot be satisfied when nonexponential data exist. In addition, GM(1,1) rolling approach derives a better prediction precise compared with GM(1,1) model, and this is the superiority of the rolling mechanism.

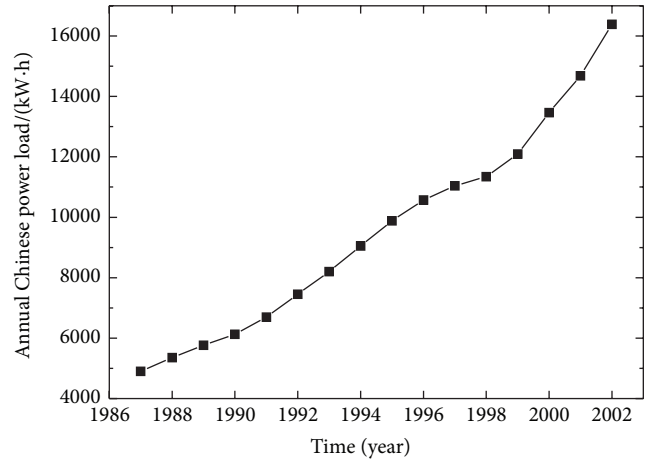


FIGURE 2: Chinese annual power load data.

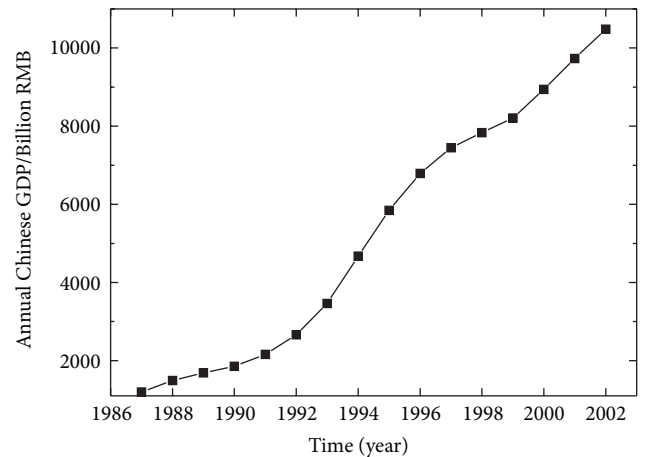


FIGURE 3: Chinese annual GDP.

TABLE 6: ARPE of the prediction results (%).

ARPE	ARPRM	GM(1,1) model	GM(1,1) rolling model	AR model with linear interpolation
$t = 16$	3.87	30.96	30.96	3.69
$t = 17$	7.46	48.75	44.80	11.66
$t = 18$	12.15	71.23	65.98	18.17

TABLE 7: ARPE of the Chinese annual power load prediction results (%).

ARPE	ARPRM	GM(1,1) model	GM(1,1) rolling model	AR model with linear interpolation
$t = 15$	3.08	0.68	0.68	0.08
$t = 16$	0.14	4.17	4.42	7.37

TABLE 8: ARPE of the Chinese annual GDP prediction results (%).

ARPE	ARPRM	GM(1,1) model	GM(1,1) rolling model	AR model with linear interpolation
$t = 15$	5.77	23.22	23.22	1.83
$t = 16$	7.61	31.53	32.50	10.18

6. Conclusions

This study presents a novel approach to settle the small sample time series prediction problem in the presence of unstable. One should construct an ARPRM model independently based on the metabolism sample in every prediction step. In other words, the model order and parameters are modified in each new forecasting step. This modification is much helpful to improve the prediction accuracy in small sample circumstance. We should note that the autoregressive in this study, which can model nonstationary sequences, is different from the AR method in the literature, which has to satisfy the stationarity requirement.

Simulation study first conducts a general verification, then monitors the performance of the ARPRM method while varying the sample size and nonlinearity dynamic mechanism, and finally checks whether the reasonability depends on the variance of the innovation. We also performed statistical analysis to demonstrate the significance of the observed trends. Results show that factors including sample size, nonlinearity dynamic mechanism, significance of trends, and innovation variance do impacts on the performance of the proposed methodology. Specifically, we found that the nonlinearity plays a much more important role in forecasting performance than trend significance. For the same simulation model, precision will be enhanced while increasing sample size or reducing the innovation variance. Meanwhile, the established approach will illustrate a better property for the dynamic mechanisms that show a stronger linearity. Comparing the results in simulation study in Section 4, it can be seen that innovation variance (the uncertainty or randomness) has a greater effect on the performance.

Empirical applications with building settlement prediction and economic forecasting are also included. The results show that the proposed method outperforms the Grey prediction model GM(1,1) and GM(1,1) rolling model, and AR

model with linear interpolation, and also illustrate a promise for future applications in small sample time series prediction.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors are grateful to the anonymous reviewers and the editor for their critical and constructive comments on this paper. This study was cosupported by the National Natural Science Foundation of China (Grant no. 11202011), National Basic Research Program of China (973 Program) (Grant no. 2012CB720000), and Fundamental Research Funds for the Central Universities (Grant no. YWK13HK11).

References

- [1] F. Busetti and J. Marcucci, "Comparing forecast accuracy: a Monte Carlo investigation," *International Journal of Forecasting*, vol. 29, no. 1, pp. 13–27, 2013.
- [2] W. K. Wong, M. Xia, and W. C. Chu, "Adaptive neural network model for time-series forecasting," *European Journal of Operational Research*, vol. 207, no. 2, pp. 807–816, 2010.
- [3] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting-Methods and Applications*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1998.
- [4] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," *Physical Review Letters*, vol. 59, no. 8, pp. 845–848, 1987.
- [5] Y.-H. Lin and P.-C. Lee, "Novel high-precision grey forecasting model," *Automation in Construction*, vol. 16, no. 6, pp. 771–777, 2007.
- [6] P. Pinson, "Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions," *Journal of*

- the Royal Statistical Society C: Applied Statistics*, vol. 61, no. 4, pp. 555–576, 2012.
- [7] P. G. Zhang, “Time series forecasting using a hybrid ARIMA and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003.
 - [8] J. D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, NJ, USA, 1994.
 - [9] S. Liu, E. A. Maharaj, and B. Inder, “Polarization of forecast densities: a new approach to time series classification,” *Computational Statistics & Data Analysis*, vol. 70, pp. 345–361, 2014.
 - [10] A. Gelman and J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, Cambridge, UK, 2007.
 - [11] M. Hofmann, C. Gatu, and E. J. Kontoghiorghe, “Efficient algorithms for computing the best subset regression models for large-scale problems,” *Computational Statistics & Data Analysis*, vol. 52, no. 1, pp. 16–29, 2007.
 - [12] R. Ghazali, A. Jaafar Hussain, N. Mohd Nawi, and B. Mohamad, “Non-stationary and stationary prediction of financial time series using dynamic ridge polynomial neural network,” *Neurocomputing*, vol. 72, no. 10–12, pp. 2359–2367, 2009.
 - [13] E. Kayacan, B. Ulutas, and O. Kaynak, “Grey system theory-based models in time series prediction,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1784–1789, 2010.
 - [14] G. Rubio, H. Pomares, I. Rojas, and L. J. Herrera, “A heuristic method for parameter selection in LS-SVM: application to time series prediction,” *International Journal of Forecasting*, vol. 27, no. 3, pp. 725–739, 2011.
 - [15] G. Huerta and M. West, “Priors and component structures in autoregressive time series models,” *Journal of the Royal Statistical Society B: Statistical Methodology*, vol. 61, no. 4, pp. 881–899, 1999.
 - [16] G. Huerta and M. West, “Bayesian inference on periodicities and component spectral structure in time series,” *Journal of Time Series Analysis*, vol. 20, no. 4, pp. 401–416, 1999.
 - [17] E. J. McCoy and D. A. Stephens, “Bayesian time series analysis of periodic behaviour and spectral structure,” *International Journal of Forecasting*, vol. 20, no. 4, pp. 713–730, 2004.
 - [18] M. A. Sharaf, D. L. Illman, and B. R. Kowalski, *Chemometrics*, John Wiley & Sons, New York, NY, USA, 1986.
 - [19] W. Zheng and A. Tropsha, “Novel variable selection quantitative structure-property relationship approach based on the K-nearest-neighbor principle,” *Journal of Chemical Information and Computer Sciences*, vol. 40, no. 1, pp. 185–194, 2000.
 - [20] A. Porta, G. Baselli, S. Guzzetti, M. Pagani, A. Malliani, and S. Cerutti, “Prediction of short cardiovascular variability signals based on conditional distribution,” *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 12, pp. 1555–1564, 2000.
 - [21] J. H. Stock and M. W. Watson, “A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series,” in *Cointegration, Causality and Forecasting*, R. F. Engle and H. White, Eds., Oxford University Press, Oxford, UK, 1999.
 - [22] M. P. Clements and D. F. Hendry, *Forecasting Economic Time Series*, Cambridge University Press, Cambridge, UK, 1998.
 - [23] J. Bai and P. Perron, “Estimating and testing linear models with multiple structural changes,” *Econometrica*, vol. 66, no. 1, pp. 47–78, 1998.
 - [24] J. Bai and P. Perron, “Computation and analysis of multiple structural change models,” *Journal of Applied Econometrics*, vol. 18, no. 1, pp. 1–22, 2003.
 - [25] K. L. Wen, *Grey Systems*, Yang’s Scientific Press, Tucson, Ariz, USA, 2004.
 - [26] D. Akay and M. Atak, “Grey prediction with rolling mechanism for electricity demand forecasting of Turkey,” *Energy*, vol. 32, no. 9, pp. 1670–1675, 2007.
 - [27] W. W. S. Wei, *Time Series Analysis*, Addison Wesley, Redwood City, Calif, USA, 2nd edition, 2006.
 - [28] P. Sørensen, N. A. Cutululis, A. Viguera-Rodríguez et al., “Power fluctuations from large wind farms,” *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 958–965, 2007.
 - [29] V. Akhmatov, “Influence of wind direction on intense power fluctuations in large offshore windfarms in the North Sea,” *Wind Engineering*, vol. 31, no. 1, pp. 59–64, 2007.
 - [30] M. di Rienzo, G. Parati, G. Mancina, A. Pedotti, and P. Castiglioni, “Investigating baroreflex control of circulation using signal processing techniques,” *Engineering in Medicine and Biology Magazine*, vol. 16, no. 5, pp. 86–95, 1997.
 - [31] A. Porta, P. Castiglioni, M. D. Rienzo et al., “Information domain analysis of the spontaneous baroreflex during pharmacological challenges,” *Autonomic Neuroscience: Basic and Clinical*, vol. 178, pp. 67–75, 2013.
 - [32] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, 1974, System identification and time-series analysis.
 - [33] H. Sohn, C. R. Farrar, F. M. Hemez et al., “A review of structural health monitoring literature: 1996–2001,” Tech. Rep. LA-13976-MS, Los Alamos National Laboratory, Los Alamos, NM, USA, 2004.
 - [34] X. Lin, *Application of Time Series Analysis to the Buildings Deformation Monitor*, Ji Lin University, Changchun, China, 2005.
 - [35] China Electric Power Committee, *China Electric Power Yearbook 2003*, China Electric Power Publishing House, Beijing, China, 2003.
 - [36] National Bureau of Statistics of China, *China Statistical Yearbook 2003*, China Statistics Press, Beijing, China, 2003.

