A Generalized Autoregressive Model Fusing Both Linearity and Nonlinearity and Its Application

Fei Hao, Chengchong Gao, Ying Dong, Ruwen Chen, Tianqi Zhang

Abstract—**A generalized autoregressive (GNAR) model fusing both linearity and nonlinearity is proposed to solve the complex nonlinear time series modeling problem. First, the mathematical model of the GNAR model is established while the mathematical mechanism and physical meaning of the GNAR model are both expounded from the two aspects of Weierstrass theory and Volterra theory. Then, an improved least squares parameter estimation method, namely robust Residuals Adjusted Least Squares (RALS) method, is introduced and its process is successionally proposed to improve the anti-outlier performance of the GNAR model. Next, the mathematical model of the computational complexity for the GNAR model is established and the complexity model is introduced into the AIC to propose an improved AIC (*i*AIC). Finally, a dataset containing 23 records is established and experiments are carried out. The results show that the GNAR model with RALS estimator has high fitting accuracy. The Mean Square Error (MSE) of the series predicted by the GNAR model with Least Squares (LS) estimator is 325.3% higher than that of the series predicted by the GNAR model with RALS estimator at most. The effectiveness of order determination method reaches 82.61%. Therefore, the GNAR model together with its parameter estimation method and order determination method proposed in this paper are effective.**

Keywords—Autoregressive model; nonlinear time series; residual autocovariance; anti-outlier; AIC

I. Introduction

At present, while the energy situation is increasingly grim, the pace of new energy substitution is accelerating. By the end of 2020, China's Internet rental bicycles had reached 19.5 million, covering 360 cities nationwide, with more than 300 million registered users and 47 million daily orders [1]. According to the data of iiMedia Group, the number of shared electric bikes in China was nearly 4 million in 2021, and its revenue was 9.36 billion yuan. It is estimated that it will exceed 20 billion yuan in 2025. By June 2022, the number of new energy vehicles in China had reached 10.01 million, accounting for 3.23% of the total number of vehicles. New energy vehicles, shared bicycles and shared electric bicycles have become indispensable important means of transportation. The shared transport facilities have brought great convenience to people's travel, but also caused related problems. For example, the supply of shared charging stations for new energy vehicles cannot meet the increasing demand of users [2], and the imbalance between shared bicycle distribution and user demand leads to a serious waste or shortage of shared bicycles [3]. Based on the survey of more than 300 cities, simply increasing or reducing the number of shared bicycles according to the traditional demand function cannot fundamentally solve the growing problem of utilization because of the dynamic demand of shared transportation facilities [4]. It is well known that, the confirmation of demand for shared transport facilities is the primary basis for optimal allocation and efficient management of shared transport resources. Nevertheless, it is a time series modeling and prediction problem to forecast the future usages based on the historical usages of shared transport facilities as the historical trip data of shared transport facilities is complex and nonlinear. Furthermore, the traditional linear or nonlinear autoregressive (AR) models for time series were usually proposed to solve a certain type of problem or even a specific problem. Therefore, in this paper, we start with specific problem of demand prediction for the shared transport facilities and then explore a generalized autoregressive model fusing linearity and nonlinearity, which is applicable to solve various problems. The main contributions are as follows:

- The mathematical expression of the GNAR model is derived, and the mathematical mechanism and physical meaning of the GNAR model are expounded respectively based on Weierstrass theory and Volterra theory.
- A robust RALS parameter estimation method is introduced to raise the anti-outlier ability of the GNAR model, and the process of RALS method is subsequently proposed.
- During determining the order of the GNAR model, the problem is considered that the computational complexity of the GNAR model increases sharply due to the nonlinearity of the GNAR model, so as to present an iAIC method.

Manuscript received March 29, 2023; revised November 14, 2023.

This work was supported in part by the National Natural Science Foundation of China under grant number 51705238, the Major Natural Science Foundation for Colleges and Universities in Jiangsu Province under grant number 23KJA460009.

F. Hao is a professor at the School of Mechanical Engineering, Nanjing Institute of Technology, Nanjing, CO 211167 China (phone: 86-25-86118250; fax: 86-25-86118255; e-mail: feehao2012@163.com).

C. Gao is an associate professor at the School of Mechanical Engineering, Nanjing Institute of Technology, Nanjing, CO 211167 China (e-mail: gaocc@njit.edu.cn).

Y. Dong is a lecturer at the School of Mechanical Engineering, Nanjing Institute of Technology, Nanjing, CO 211167 China (e-mail: nlgdaisy@163.com).

R. Chen is a professor at the School of Automotive and Rail Transit, Nanjing Institute of Technology, Nanjing, CO 211167 China (e-mail: jscrw@163.com).

T. Zhang is a postgraduate student at the School of Mechanical Engineering, Nanjing Institute of Technology, Nanjing, CO 211167 China (e-mail: Y00450210343@njit.edu.cn).

• The experimental results indicate that the MSE of the series predicted by the GNAR model with LS estimator is 325.3% higher than that of the series predicted by the GNAR model with RALS estimator while the effectiveness of order determination method reaches 82.61%.

The remainder of this paper is organized as follows. The section II reviews the related works in the time series modeling. The section III describes the GNAR model. The section IV gives the experimental results and discussion. Finally, the conclusions are presented.

II. RELATED WORKS

A. Distribution Point Location and Demand Forecast of Shared Transport Facilities

The location of distribution points for shared transport facilities and the demand forecast of shared transport facilities are all attributed to the problem of facility location. According to the parameters of the location models, the location problems of shared bicycles were divided into demand certain type and demand uncertain type [5]. According to the characteristics of the areas, the location problems were divided into discrete type, continuous type and network type. According to the number of facilities, the location problems were divided into single-facility location problem and multi-facilities location problem. There are heuristic algorithm, projection algorithm, branch and bound algorithm for solving the multi-facilities location problems [6]. According to the charging characteristics of different types of electric vehicles, Zheng Chun et al. [7] built a demand forecasting model for electric vehicle charging facilities based on data fusion. Yang Zhenzhen et al. [8] proposed a data-driven location method for electric vehicle charging stations by calculating the traffic demand of each site of grid map. Tian Feng et al. proposed a bi-objective model on the basis of considering user preferences and the model was solved by using multi-objective particle swarm optimization algorithm to obtain the optimal node location and number of charging stations [9]. The mathematical models of P-center and P-median were selected by Suleyman et al. to allocate the alternative points and demand points of shared facilities [10]. The maximum coverage model was established by Frade et al. on the basis of considering the cost and service level of the lease point to obtain the location and capacity of each power distribution station [11]. Liu et al. proposed a seasonal grey Markov model to predict the demand for shared bicycles on the basis of the periodicity, nonlinearity and randomness of usages of shared bicycles [12]. Gao et al. proposed a hybrid method combining genetic algorithm based on fuzzy C-means and back-propagation network to predict the demand for shared bicycles [13].

Some models were established and solved for the location and layout of the shared transport facility distribution centers. All these models are basically based on the certain demands, but the demands are often unknown in the application. Therefore, the research on the location and layout of shared transport facility distribution centers is essentially to provide more accurate demand prediction.

B. Linear and nonlinear AR models

In 2020, the COVID-19 has brought a serious impact to the world, thus time series modeling plays an important role in forecasting the trend of epidemic development and assisting in accurate prevention and control [14-18]. Sunayana et al. used nonlinear autoregressive neural model to forecast the municipal solid waste generation [19]. Razzaq A. et al. analyzed and studied the dynamic and causality interrelationships from municipal solid waste recycling to economic growth, carbon emissions and energy efficiency using a novel bootstrapping autoregressive distributed lag model [20]. Rahmoune M. B. et al. studied the identification method of dynamic characteristics of rotating machinery using a nonlinear autoregressive with exogenous input (NARX) model [21]. Khraief N. et al. studied the movements of oil prices and exchange rates in China and India by the estimations of the wavelet-based, non-linear, autoregressive distributed lag model [22]. Berninger C. et al. proposed a bayesian time-varying autoregressive model for improved short- and long-term prediction of German exchange rate [23]. Zhang X. et al. refined the impact of adverse weather conditions such as rain, snow and wind on highway safety through the spatial polynomial Logit model [24]. Zhang F. et al. proposed a dynamic adaptive stepwise parameter estimation method of autoregressive model in real-time wind power forecasting [25]. Yano M. O. et al. used autoregressive model to predict the damage index of building structure [26]. Zhang Y. et al. adopted the spatial time series model to study and draw the conclusion that neighborhood socio-economic factors have the largest contribution to Toronto's house price [27]. To determine the best process control model thus to improve the yield of ester, a linear autoregressive with exogenous input (ARX) model and a NARX model were developed and compared by Zulkeflee S. A. et al. [28].

The time series modeling is widely used in various sectors of national economy, such as ecological protection and environmental governance, healthcare, manufacturing, finance, communications and transportation industry, power industry, construction, real estate industry, animal husbandry and so on. There are various kinds of time series models for the specific problems including both linear and nonlinear models.

Stationary linear time series models such as classical AR models are still widely used. However, in the face of increasingly complex economic and social problems and engineering problems, linear time series models show their own shortcomings. The main reason is that the classical stationary linear time series models are based on the normality and stationarity of the system and are expressed by linear difference equations which can only be applied to linear systems. When a system has strong nonlinear components, it is difficult to obtain good results by using linear time series model for system identification. Therefore, since the 1980s, experts in various fields have been committed to the research of nonlinear models. There are some classical nonlinear models including threshold AR model, exponential AR model, bilinear model, state-dependent model, etc. In recent years, some new nonlinear models have also emerged, such as nonlinear autoregressive (NAR) model and NARX model.

The output of solid waste can be expressed by a nonlinear function of some historical observations and the above nonlinear relationship can be implicitly expressed by neural network, namely the NAR model [19]. Li et al. also used the NAR model to solve the output optimization problem of nuclear power turbine units [29]. The nonlinear approximation property of neural network was employed into the NARX model to make the NARX model better adapt to industrial processes with strong nonlinear characteristics than ARX. Zulkeflee et al. applied the low-order NARX model with higher fitting accuracy to the control of ester production process [28]. A new dynamic soft measurement system based on the NARX model was established by Xiong et al. which solved the problem of relying on the real-time measurement of the dominant variables [30].

However, in most cases, both linearity and nonlinearity exist simultaneously. To solve the coexistence of linearity and nonlinearity problems, the strategy of selecting different models separately was mainly adopted. Considering the characteristics of the coexistence of linearity and nonlinearity for daily tourist numbers, Yao L. [31] proposed a hybrid model, which combines the rescaled range method (R/S), the support vector regression (SVR) and the autoregressive integrated moving average model (ARIMA), in which the SVR performs well in nonlinear prediction, while ARIMA performs well in linear prediction.

The theoretical research and practical application of classical nonlinear time series models have reached a certain level, and the emerging nonlinear time series models have also been applied to certain problems. However, the above nonlinear time series models are proposed by researchers in combination with their own engineering background, aiming at solving a certain kind of problems, or even a specific problem, but not universal. Therefore, it is necessary to explore a form of time series model, which combines linear time series model and nonlinear time series model, and is suitable for solving different problems or various applications, that is, a universal and effective model expression, which has been pursued in the field of applied mathematics and engineering.

III. Autoregressive Model Fusing Both Linearity and Nonlinearity

A. GNAR Model Establishment

1) Analysis of GNAR Model Based on Weierstrass Function Approximation Theory

According to the theory of time series modeling, when the system inputs are white noise with zero mean value, the system outputs are the observed values $\{w_t\}$ satisfying the following relationship:

$$w(t)=f\left[w(t-1),w(t-2),\cdots w(t-\infty),\ a(t-1),a(t-2),\cdots a(t-\infty)\right]+a(t) \tag{1}$$

where $w(t\text{-}i)$ is the system observation value at time $t\text{-}i$, i = 0, 1, 2, 3, …; $a(t\text{-}j)$ is the white noise of the system at time $t\text{-}j$, j = 0, 1, 2, 3, ….

According to Weierstrass theory, a continuous function defined on a closed interval can be approximated arbitrarily and accurately by polynomials. Therefore, the nonlinear function $f(\cdot)$ can be approximately expanded on a closed interval as follows:

$$\begin{aligned}w(t)&=\varphi(1)w(t-1)+\varphi(2)w(t-2)+\cdots+\\&\gamma(1)a(t-1)+\gamma(2)a(t-2)+\cdots+\\&\varphi(1,1)w(t-1)w(t-1)+\varphi(1,2)w(t-1)w(t-2)+\cdots+\\&\gamma(1,1)a(t-1)a(t-1)+\gamma(1,2)a(t-1)a(t-2)+\cdots+\\&\eta(1,1)w(t-1)a(t-1)+\eta(1,2)w(t-1)a(t-2)+\cdots+\\&\varphi(1,1,1)w(t-1)w(t-1)w(t-1)+\\&\gamma(1,1,1)a(t-1)a(t-1)a(t-1)+\cdots+a(t)\\&=\sum_i\varphi(i)w(t-i)+\sum_i\sum_j\varphi(i,j)w(t-i)w(t-j)+\\&\sum_i\gamma(i)a(t-i)+\sum_i\sum_j\gamma(i,j)a(t-i)a(t-j)+\\&\sum_i\sum_j\eta(i,j)w(t-i)a(t-j)+\cdots+a(t)\end{aligned} \tag{2}$$

where $\varphi(i)$, $\varphi(i, j)$, $\gamma(i)$, $\gamma(i, j)$, $\eta(i, j)$,... are the expansion coefficients.

Under the condition that the system described in (2) has zero initial state, i.e., $w_0 = w_{-1} = w_{-2} = \cdots = 0$, and the system has no input before zero time, $w(t\text{-}i)$ is substituted into (2) to obtain (3). , where i = 1, 2, 3,

$$\begin{cases}\gamma(i)=0,\ i=1,2,3,\cdots\\ \gamma(i,j)=0,\ i,j=1,2,3,\cdots\\ \quad\vdots\\ \eta(i,j)=0,\ i,j=1,2,3,\cdots\\ \quad\vdots\end{cases} \tag{3}$$

Then (2) can be rewritten as (4):

$$w(t)=\sum_i\varphi(i)w(t-i)+\sum_i\sum_j\varphi(i,j)w(t-i)w(t-j)+\sum_i\sum_j\sum_k\varphi(i,j,k)w(t-i)w(t-j)w(t-k)+\cdots+a(t) \tag{4}$$

2) Analysis of GNAR Model Based on Volterra Functional Series Theory

Nonlinear analysis of mechanical systems or structures is a common. When using conventional numerical calculation methods to analyze the nonlinear dynamic response of structures, it is necessary to study the physical parameters of structures and their change rules to accurately analyze them. When using Volterra model to analyze the nonlinear dynamic response of structures, it is only necessary to identify the input and output data. Then, a mathematical model describing the nonlinear response characteristics of the structure can be established. Using Volterra functional series, the following mathematical models can be established for any nonlinear system:

$$y(k)=\sum_{d=1}^{N}\left(\sum_{r_1=0}^{N_d-1}\sum_{r_2=0}^{N_d-1}\cdots\sum_{r_d=0}^{N_d-1}h(r_1,r_2,\ldots,r_d)\prod_{j=1}^{d}x(k-r_j)\right) \tag{5}$$

where $h(r_1,r_2,\ldots,r_d)$ is the Volterra kernel of order d.

The essence of time series modeling is output equivalence. The time series model is used to define the equivalent system of the original system. Under the action of white noise $\{a(t)\}$, the equivalent system outputs the observation sequence $\{w(t)\}$, while the original system outputs the observation sequence $\{w(t)\}$ under the action of $\{v(t)\}$. $\{w(t)\}$

and $\{a(t)\}$, which can be expressed by the polynomial of $w(t\text{-}i)$ according to (3), are substituted into (5) to obtain (4). The same time series model is obtained by Volterra theory and Weierstrass theory, and the physical meaning of GNAR model is introduced from two aspects of system identification and mathematical mechanism, respectively.

For the system described in (4), the more high-order items, i.e., the higher the order, the higher the model accuracy, but the system complexity rises sharply. The model order often takes a finite value in practice. Considering that the historical observations far away have little influence on the current observation, only some most adjacent observations need to be considered during modeling, so (4) can be rewritten as follows:

$$w(t)=\sum_{i_1=1}^{n_1}\varphi(i_1)w(t-i_1)+\sum_{i_1=1}^{n_2}\sum_{i_2=1}^{n_2}\varphi(i_1,i_2)w(t-i_1)w(t-i_2)+\cdots+\sum_{i_1=1}^{n_p}\cdots\sum_{i_p=1}^{n_p}\varphi(i_1,i_2,\cdots i_p)\prod_{k=1}^{p}w(t-i_k)+a(t) \tag{6}$$

Equation (6) is a one-dimensional GNAR model with order p, which fuses both linear and nonlinear autoregressive models into a uniform expression. The model is denoted as $G(p;\ n_1,\ n_2,\ \cdots;\ n_p)$ and $n_1,\ n_2,\ \cdots;\ n_p$ are the number of historical observations which need to be considered to predict the current observation. When p is 1, (6) is a linear model. When p is greater than 1, (6) is a nonlinear model.

B. Parameter Estimation Method

It is usually assumed that the time series data has Gaussian noise and thus the LS method is the most commonly used parameter estimation method. However, the Gaussian noise assumption is relatively ideal and time series data inevitably have two kinds of outliers, mainly including innovative outliers and additive outliers, resulting in poor robustness of the LS parameter estimation method [32]. To address the problems, some robust parameter estimation methods were studied, mainly including M estimation based on maximum likelihood estimation, L estimation based on linear combination and R estimation based on rank test. M estimation transforms the parameter estimation problem into the minimum value problem by establishing the maximum likelihood estimation function and its robustness is significantly better than the LS estimation [33]. However, the time series is typically contaminated by additive outliers and the robustness of M estimation still needs to be improved [34]. RA (Residual Autocovariances) estimator was first proposed by Bustos and Yohai. It has good noise resistance for additive outliers. For time series data with innovative outliers of symmetrical distribution, the RA estimation has Fisher consistency for a stationary and reversible Auto-Regressive and Moving Average (ARMA) model. The numerical results show that the RA estimation is more robust than LS estimation and M estimation for time series data with additive outliers modeling by Moving Average (MA) model [35]. RA estimation was extended to parameter estimation of two-dimensional time series model and its performance was further studied [36-39].

The main idea of the RA method is to adjust the residuals and their autocovariances by using a continuous bounded function, so as to suppress the "excessive values" in the residuals and make the parameter estimation more robust. Referring to the above idea that the residuals are adjusted by using a continuous bounded function, an improved recursive least squares parameter estimation method was proposed, which is the RALS method named previously.

Let $\boldsymbol{\lambda}=\left(\boldsymbol{\varphi}^{(1)},\boldsymbol{\varphi}^{(2)},\ldots\boldsymbol{\varphi}^{(p)}\right)$. The LS estimates of $\boldsymbol{\lambda}$ minimize

$$\sum r_t^2(\boldsymbol{\lambda}) \tag{7}$$

where $\boldsymbol{\varphi}^{(z)}=\varphi(i_1,i_2,\ldots,i_z)$, $z=1,2,\ldots,p$; the residuals $r_t(\boldsymbol{\lambda})$ are defined by

$$r_t(\boldsymbol{\lambda})=w(t)-\left(\sum_{i_1}\varphi(i_1)w(t-i_1)+\sum_{i_1}\sum_{i_2}\varphi(i_1,i_2)w(t-i_1)w(t-i_2)+\cdots+\sum_{i_1}\sum_{i_2}\cdots\sum_{i_p}\varphi(i_1,i_2,\ldots,i_p)\prod_{k=1}^{p}w(t-i_k)\right) \tag{8}$$

Consider the LS estimates obtained minimizing (7). Differentiating the expression we obtain the following system of equations for the LS estimates:

$$\sum r_t(\boldsymbol{\lambda})\cdot\left[\partial r_t(\boldsymbol{\lambda})/\partial\boldsymbol{\varphi}^{(z)}\right]=0 \tag{9}$$

It is easy to show from (8) that

$$\frac{\partial r_t(\boldsymbol{\lambda})}{\partial\boldsymbol{\varphi}^{(z)}}=-\prod_{k=1}^{z}w(t-i_k) \tag{10}$$

To ensure robustness of the proposed estimator, $r_t(\boldsymbol{\lambda})$ is replaced by $\psi\left(r_t(\boldsymbol{\lambda})/\hat{\sigma}\right)$ where $\psi:R\to R$ is a continuous and bounded odd function and $\hat{\sigma}$ is a scale factor which is a robust estimate of the mean of a sequence, which is composed of the absolute values of residuals $r_t(\boldsymbol{\lambda})$. $\hat{\sigma}$ is estimated using, for example,

$$\hat{\sigma}^{(k)}=\text{Median}\left(\left\{\left|r_t(\boldsymbol{\lambda})\right|\right\}\right)/0.6475 \tag{11}$$

ψ may be a function from the Huber function family given by

$$\psi_{H,c}(u)=\text{sgn}(u)\cdot\min\left(|u|,c\right) \tag{12}$$

It can also be one of the bisquare function family proposed by Beaton and Tukey, which is defined by

$$\psi_{B,c}(u)=u\left(1-u^2/c^2\right)^2 \tag{13}$$

Subsequently, the parameters of a GNAR model can estimated with the RALS method according to process given by Algorithm 1.

Algorithm 1: RALS estimation method

1 Time series $\{w_t\}$

$$\{w_1,w_2,\ldots,w_{t-1},w_t,w_{t+1},\ldots,w_n\} \tag{14}$$

Modeling sequence $\{w_t\}$ using the GNAR model:

$$w_t=G\left(w_{t-1},w_{t-2},w_{t-3},\ldots,w_{t-s}\right)+\varepsilon_t \tag{15}$$

where G is the GNAR model of which the parameters are represented by the vector $\boldsymbol{\lambda}$. s is a constant determined by the order and neighbor structure of G.

2 The parameters of GNAR model are estimated by LS method which were denoted as $\hat{\boldsymbol{\lambda}}^{(0)}$. The initial GNAR model, the residual sequence and the median of the residual sequence were denoted as $\hat{G}^{(0)}$, $\left\{\boldsymbol{r}^{(0)}\right\}$ and $\hat{\sigma}^{(0)}$, respectively. Take the above parameters as the initial values and start iterative LS estimation.

3 In the k^{th} iteration, the parameters of GNAR model were estimated by LS method which were denoted as $\hat{\boldsymbol{\lambda}}^{(0)}$ and the GNAR model was denoted as $\hat{G}^{(k)}$. The residual error of the GNAR model were calculated:

$$r_t^{(k)} = w_t - \hat{G}^{(k)}\left(w_{t-1}, w_{t-2}, w_{t-3}, \ldots, w_{t-s}\right) \quad (16)$$

4 Residual sequence

$$\left\{\boldsymbol{r}^{(k)}\right\} = \left\{r_t^{(k)}, r_{t+1}^{(k)}, r_{t+2}^{(k)}, \ldots, r_n^{(k)}\right\} \quad (17)$$

5 The median of the residual sequence was estimated according to (18)

$$\hat{\sigma}^{(k)} = \text{Median}\left(\left\{\left|\boldsymbol{r}^{(k)}\right|\right\}\right) / 0.6475 \quad (18)$$

Where $\text{Median}(\cdot)$ is the median function.

6 Tuning the residual errors using a continuous bounded odd function

$$\hat{r}_t^{(k)} = \psi\left(r_t^{(k)} / \hat{\sigma}^{(k)}\right) \cdot \hat{\sigma}^{(k)} \quad (19)$$

where $\psi(\cdot)$ is a continuous and bounded odd function. A function from the Huber function family is commonly used:

$$\psi_H(u) = \text{sgn}(u) \cdot \min\left(|u|, c\right) \quad (20)$$

Where c is the tuning parameter between 1.5 and 2.0.

7 Update time series $\{w_t\}$ according to (21):

$$\hat{w}_t^{(k)} = \hat{r}_t^{(k)} + \hat{G}^{(k)}\left(\hat{w}_{t-1}^{(k)}, \hat{w}_{t-2}^{(k)}, \hat{w}_{t-3}^{(k)}, \ldots, \hat{w}_{t-s}^{(k)}\right) \quad (21)$$

Updated time series $\left\{\hat{w}_t^{(k)}\right\}$

$$\left\{\hat{w}_1^{(k)}, \hat{w}_2^{(k)}, \hat{w}_3^{(k)}, \ldots, \hat{w}_{t-1}^{(k)}, \hat{w}_t^{(k)}, \hat{w}_{t+1}^{(k)}, \ldots, \hat{w}_n^{(k)}\right\} \quad (22)$$

8 Let $k = k+1$. Repeat steps 3 to 7. Stop iteration when the following conditions are met.

$$\left|\hat{\sigma}^{(k)} - \hat{\sigma}^{(k+1)}\right| < \varepsilon_1, \quad \left\|\hat{\lambda}^{(k)} - \hat{\lambda}^{(k+1)}\right\| < \varepsilon_2 \quad (23)$$

where ε_1 and ε_2 are two relatively small positive constants, such as 10^{-3} or 10^{-6}. $\|\cdot\|$ is the matrix norm.

C. Order Determination of GNAR Model Based On Residual Characteristics and Time Complexity

Model order determination is very important which is intrinsically related to model applicability test and goodness judgment [40]. The classic criteria for judging the degree of excellence mainly include the residual whitening test-based criterion, the residual sum of squares-based criterion, the final prediction errors-based criterion and the information-based criterion. They are not universally applicable test criteria. For example, the residual whitening test-based criterion is only applicable to linear systems. Therefore, there are new order determination criteria constantly emerging, such as frequency ratio-based method, steady-state diagram-based method, AIC, improved F-test-based method and genetic algorithm-based method, etc [41, 42]. However, nonlinear models are quite different and so far there is no universally applicable method for model order determination.

The parameters of GNAR model increase rapidly and the operation cost increases sharply with nonlinear terms in the model, which should be considered when determining the model order. Therefore, combining the time complexity model with the now available criteria, a comprehensive order determination method for linear and nonlinear fused autoregressive model based on the residual characteristics and the time complexity was proposed.

1) Complexity of GNAR Model

The dimensions of GNAR model parameters can be calculated as follows:

$$m = \sum_{i=1}^{p} C_{n_i+i-1}^{i} \quad (24)$$

Where m is the dimension of GNAR model parameters, p is the order of GNAR model, and n_i is the number of historical items of each order.

According to (24), the calculation amount S of GNAR model parameters is defined as follows:

$$S = \sum_{i=1}^{p}\left(1+2^i\right) \cdot C_{n_i+i-1}^{i} \quad (25)$$

2) Improved AIC

An iAIC criteria is defined by considering the computational complexity of GNAR model in the AIC criteria. The mathematical model of iAIC is as follows:

$$i\text{AIC}(S) = \ln \sigma_a^2 + \tau \cdot S \quad (26)$$

Where τ is a tuning factor and its value can be $\ln(N)$. N is the number of historical observations.

IV. Experiments and Results

A. Experimental Implementation

1) Dataset

A dataset including 23 records is employed to validate the performances of GNAR model and RALS estimator, as shown in Table I. Among the 23 records, the first 9 records are from a book titled Time Series Analysis in Engineering Application by Yang Shuzi, Wu Ya and Xuan Jianping. The 10th record is the trip history data from Citi Bike which was downloaded through the link https://ride.citibikenyc.com/system-data. The trip data from Nov. 1, 2022 to Nov. 30, 2022 of the site named 2 Ave&E 29 St and numbered 6122.09 are analyzed. A day is split up into six intervals. 0:00-06:00 is the first interval, 06:00-10:00 is the second interval, 10:00-14:00 is the third interval, 14:00-18:00 is the fourth interval, 18:00-22:00 is the fifth interval and 22:00-24:00 is the sixth interval. The 11th record, the 12th record and the 13th record are trip history data from Divvy in Chicago which are downloaded through the link https://ride.divvybikes.com/system-data. The trip data from Nov. 1, 2022 to Nov. 30, 2022 of three sites respectively named Ellis Ave & 60th St, University Ave & 57th St and Streeter Dr & Grand Ave and numbered KA1503000014, KA1503000071 and 13022 are analyzed. The 14th to 23rd

records are from a dataset named Boston Hubway Data Visualization Challenge Dataset established by Massachusetts Department of Transportation. The trip history data from July 28, 2011 to Nov. 27, 2013 of ten sites which are respectively numbered 3, 4, 5, 6, 7, 9, 10, 11, 12 and 23 are analyzed. The download link can be obtained by searching the dataset name through a search engine.

Each record is divided into two parts, i.e., modeling part (training set) and verification part (test set), accounting for 80% and 20% respectively.

TABLE I
DATASET FOR VALIDATION OF GNAR MODEL AND RALS ESTIMATOR

No.	Observation Data (History Data) Title	Length
1	Errors of Lead Screw System	490
2	Stocks of ※※ Company	369
3	Vibration Acceleration of Lathe Tip (Stable State)	130
4	Vibration Acceleration of Lathe Tip (Transition State)	130
5	Vibration Acceleration of Lathe Tip (Flutter State)	130
6	Dynamic Cutting Force	150
7	Vibration Displacement	150
8	Tickets of ※※ International Airline per Month	144
9	Yearly Averaged Sunspot Numbers from 1700 to 1987	288
10	Trip Data of Station 6122.09 (Citi Bike)	171
11	Trip Data of Station KA1503000014 (Divvy)	170
12	Trip Data of Station KA1503000071 (Divvy)	164
13	Trip Data of Station 13022 (Divvy)	139
14	Trip Data of Station 3 (Hubway)	616
15	Trip Data of Station 4 (Hubway)	584
16	Trip Data of Station 5 (Hubway)	587
17	Trip Data of Station 6 (Hubway)	608
18	Trip Data of Station 7 (Hubway)	564
19	Trip Data of Station 9 (Hubway)	591
20	Trip Data of Station 10 (Hubway)	589
21	Trip Data of Station 11 (Hubway)	619
22	Trip Data of Station 12 (Hubway)	620
23	Trip Data of Station 23 (Hubway)	324

2) GNAR Models

As previously explained, the GNAR model can be written as $G(p;n_1,n_2,\cdots,n_p)$. In the performance verification of GNAR model, the maximum value of p is 5, that is, the highest order of GNAR model is 5. When p = 1, n_1 ranges from 1 to 20. When p = 2, n_1 ranges from 1 to 10 and n_2 ranges from 1 to 5. When p = 3, n_1 ranges from 1 to 10 and n_2 and n_3 range from 1 to 5. When p = 4, n_1 ranges from 1 to 10, n_2 and n_3 range from 1 to 5 and n_4 ranges from 1 to 2. When p = 5, n_1 ranges from 1 to 10, n_2 and n_3 range from 1 to 5, and n_4 and n_5 range from 1 to 2. There are 770 GNAR models.

3) Evaluation Index

To compare the performance of the LS estimator and the RALS estimator, $\mathrm{ReRes}(i)$, i.e., percentage of relative residual, is defined as follows:

$$\mathrm{ReRes}(i)=\frac{\mathrm{LsRes}(i)-\mathrm{RaRes}(i)}{\mathrm{LsRes}(i)}\times 100\% \qquad (27)$$

where $\mathrm{LsRes}(i)$ is the residual of forecasting the i^{th} observation point by using the LS method to estimate the model parameters. $\mathrm{RaRes}(i)$ is the residual of forecasting the i^{th} observation point by using the RALS method to estimate the model parameters. $\mathrm{ReRes}(i)$ is the percentage of relative residual of $\mathrm{RaRes}(i)$ and $\mathrm{LsRes}(i)$. If $\mathrm{ReRes}(i) > 0$, it means that the prediction residual of the model whose parameters are estimated by the LS method is larger than that of the model whose parameters are estimated by the RALS method. Otherwise, it means that the prediction residual of the former is smaller than that of the latter. An evaluation index r was defined as follows:

$$r=\frac{N_2}{N_1}\times 100\% \qquad (28)$$

where N_2 is the number of elements less than zero in the relative residual sequence $\{\mathrm{ReRes}(i)\}$ and N_1 is the length of the relative residual sequence $\{\mathrm{ReRes}(i)\}$.

Furthermore, in order to evaluate the overall performance of GNAR model using the above two parameter estimation methods on the test sets, the MESs of the prediction residual sequences are calculated.

B. Results and Analysis

1) Fitting Characteristic Analysis of GNAR Model

As shown in Fig. 1 (a), the yearly averaged sunspot numbers are fitted by using $G(4; 7, 5, 1, 1)$. The iAIC is 6.9005 when LS method is used to estimate the parameters of the model and the iAIC is 5.5360 when RALS method is used. The MSE of the sequence predicted by the model of which the parameters are estimated by LS method is 1108.40 and the MSE of the sequence predicted by the model of which the parameters are estimated by RALS method is 260.61. Compared with the above GNAR model, $G(4; 9, 4, 2, 1)$ can also better fit the data, as shown in Fig. 1 (b). The two iAICs are 6.3302 and 5.5768, respectively. The two MSEs are 645.30 and 274.21, respectively.

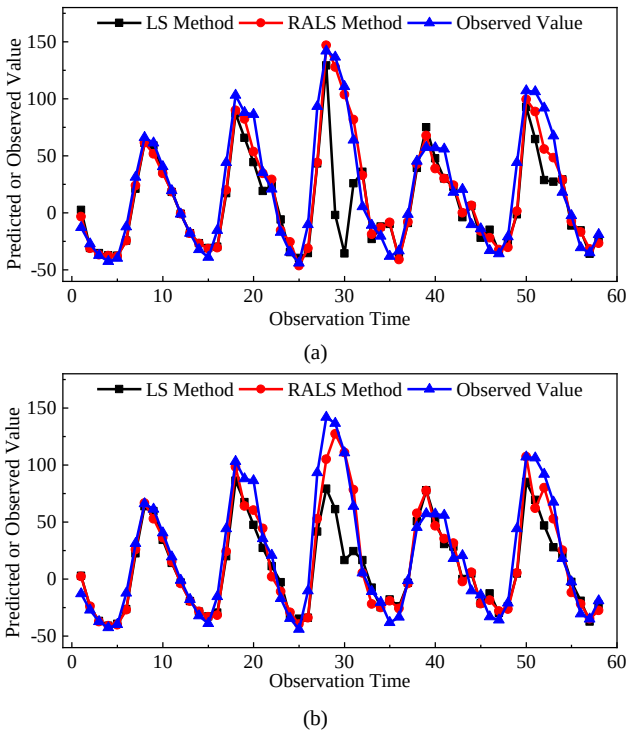

Fig. 1. Fitting Yearly Averaged Sunspot Numbers with GNAR Models.

To further quantify the prediction error at each point, the differences between observed value and two predicted values in Fig. 1 are calculated and the results are plotted in Fig. 2.

Fig. 2 shows that there are some abnormal points in both Fig. 2 (a) and Fig. 2 (b) when LS method is employed to estimate the parameters of the GNAR model. However, the prediction errors at these points are relatively small when

RALS method is adopted. Among these 58 points, the maximum relative difference of prediction errors is up to 2054.50% at the 30th point. Consequently, the 9th record (Yearly Averaged Sunspot Numbers) can be well fitted by the GNAR model with the RALS estimator.

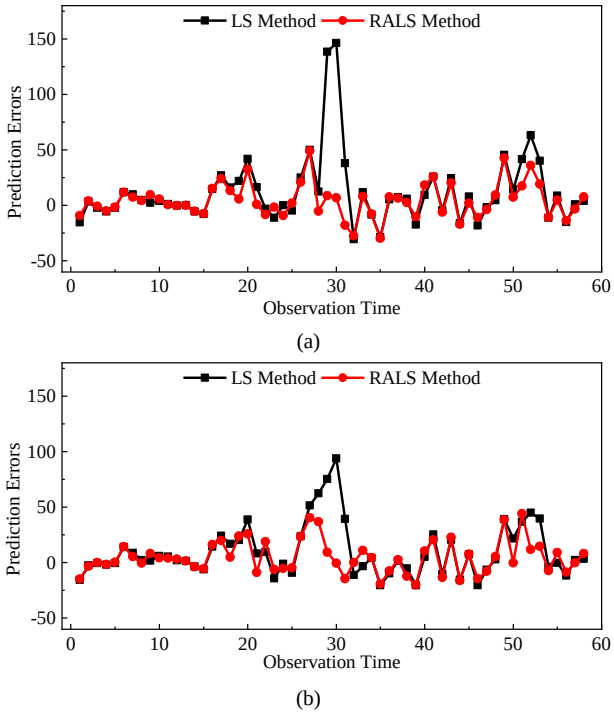

Fig. 2. Prediction Errors for Yearly Averaged Sunspot Numbers.

The trip historical data of station KA1503000014, i.e., the 11th record in Table I, is fitted by $G(5; 4, 3, 3, 2, 2)$, as illustrated in Fig. 3.

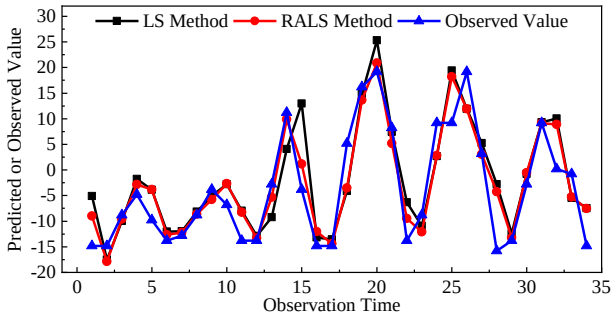

Fig. 3. Fitting Trip Historical Data of Station KA1503000014

The series shown in Fig. 1 is relatively stable. There is little difference at both ends, but a slight increase in the middle. Although the series shown in Fig. 3 is also generally relatively stable, its trend is different from that shown in Fig. 1. When observation time is less than 10, the observation values are relatively small and when observation time is greater than 12, the observation values have a relatively significant increase.

$G(5; 4, 3, 3, 2, 2)$ can well fit the trip historical data of station KA1503000014 with the *i*AIC of 3.5508 and the MSE of 33.9517 when RALS method was adopted to estimate the parameters of GNAR model. When LS method is adopted, the *i*AIC and MSE are 3.6434 and 38.8320, respectively.

The historical data of station 13022, i.e., the 13th record in Table I, is fitted by $G(3; 10, 4, 2)$, as shown in Fig. 4.

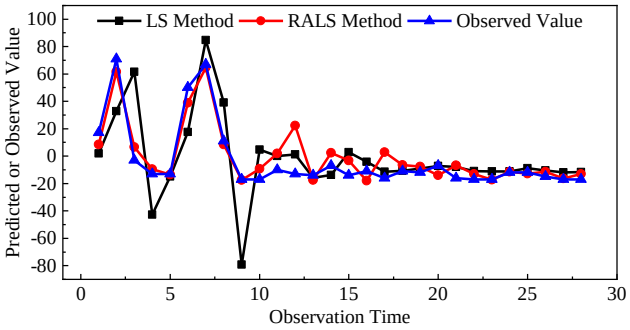

Fig. 4. Fitting Trip Historical Data of Station 13022

The variation of the trip historical data of station 13022 is different from the previous two series. The trip historical data of the station should be clearly divided into two parts. When observation time is less than 10, the observations is relatively large as well as its variation rate. However, when observation time is greater than 10, the observations decreases significantly and its variation rate is relatively gentle.

$G(3; 10, 4, 2)$ can well fit the trip historical data of station 13022 with the *i*AIC of 5.0227 and the MSE of 148.4276 when RALS method is adopted to estimate the parameters of GNAR model. When LS method is adopted, the *i*AIC and MSE are 6.2727 and 506.9140, respectively. Compared with LS method, the trip historical data of station 13022 is better fitted by using $G(3; 10, 4, 2)$ when RALS method is adopted to estimate the parameters of GNAR model. Especially when the observation time is less than 10, the fitting accuracy of the GNAR model with parameters estimated by RALS method is significantly better than that of GNAR model with parameters estimated by LS method. The maximum relative difference of prediction errors is up to 11944.81% at the 9th point.

The trip historical data of station 7, i.e., the 18th record in Table I, is fitted by $G(1; 14)$, as illustrated in Fig. 5.

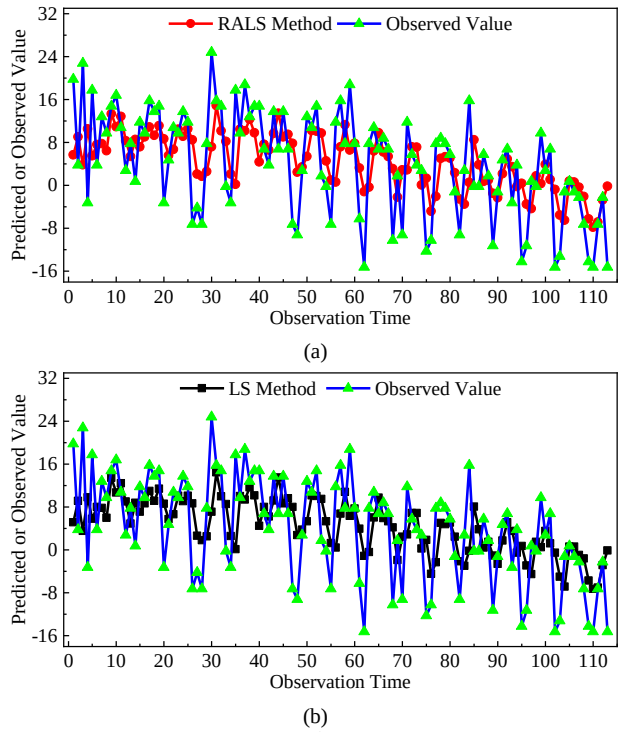

Fig. 5. Fitting Trip Historical Data of Station 7

The data created by the $G(1; 14)$ can well track the trend of the observed data, but the prediction errors are relatively large. When RALS method is used, the *i*AIC and the MSE are 4.0854 and 58.8160, respectively. When LS method is used, the *i*AIC and the MSE are 4.1252 and 61.1877, respectively. Therefore, the fitting accuracy of the GNAR model with RALS method is slightly better than that of GNAR model with LS method.

The good fitting accuracy of the GNAR model has been proved from the above analysis. However the GNAR model also has shortcomings. The fitting accuracy of the GNAR model on the data with obvious trend items and relatively violent changes is significantly reduced. Therefore, in the future, on the one hand, the adaptability of the GNAR model to various types of data should be further improved. On the other hand, effective measures should be taken to separate the trend items of the data and the data compression methods must be studied to reduce the rate of change thus to make the data more consistent with the requirements of time series modeling for data stability.

Although the data shown in Fig. 1, Fig. 3 and Fig. 4 have relatively large differences in their change rules, these data can be well fitted by using the GNAR models. The data shown in Fig. 4 is quite different from those data shown in the previous three images. It has obvious trend items and the data changes rapidly and substantially. The GNAR model can fit the trend items very well, but the prediction errors are relatively large. Further research needs to be carried out from different perspectives, such as modeling strategy or preprocessing.

2) Characteristic Analysis of RALS Estimator

The parameters of GNAR model are estimated by LS method and RALS method, respectively. Then the prediction experiments are carried out on the aforementioned dataset with 23 records. The MSEs of 23 prediction series which are obtained by using the two parameter estimation methods are calculated respectively.

Here the relative difference between MSE of LS method and MSE of RALS method is defined as the percentage of the former over the latter. According to Table II, among the 23 relative differences, only 1 is less than zero. Among the other 22 relative differences, the maximum value is as high as 325.30%, and the second largest value is 241.52%. It indicates that the overall fitting errors of the GNAR models with RALS estimator are smaller than those of the GNAR models with LS estimator.

TABLE II
RELATIVE DIFFERENCES OF MSE FOR RECORDS IN TABLE I

No.	RESULTS	No.	RESULTS	No.	RESULTS
1	0.27%	9	325.30%	17	-0.46%
2	20.06%	10	4.55%	18	4.03%
3	46.78%	11	14.37%	19	1.71%
4	7.00%	12	28.38%	20	2.93%
5	1.77%	13	241.52%	21	0.31%
6	3.14%	14	5.71%	22	0.31%
7	6.32%	15	0.09%	23	0.48%
8	7.13%	16	1.53%	/	/

When the *i*AIC takes the minimum value, the evaluation index r defined in (21) is calculated according to the GNAR model's prediction errors which are equal to the differences between observations and predictions. The evaluation index r for each record in Table I is calculated accordingly and the 23 values of the evaluation indexes are plotted as a histogram, as shown in Fig. 6.

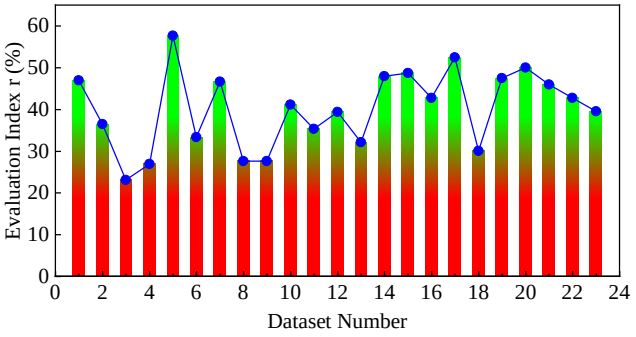

Fig. 6. Evaluation Index r for Each Record in Table I

According to Fig. 6, among the above 23 results of the evaluation indexes, there are 4 results in the interval [20%, 30%), 7 results in the interval [30%, 40%), 9 results in the interval [40%, 50%) and 3 results in the interval [50%, 60%). The maximum and minimum results are 57.69% and 23.08%, respectively. Thus, the proportion of r greater than 50% is only about 13%. According to (21), the smaller the r, the better the RALS estimator. Consequently, it is cautiously concluded that the performance of the RALS method is significantly better than that of the LS method.

3) Characteristic Analysis of iAIC

All records in the dataset are modelled by using 770 GNAR models and the parameters of the GNAR models are estimated by the LS method and the RALS method. Subsequently, the *i*AICs of each model and the MSEs of each prediction series are calculated. Two groups of results for each record are selected according to its *i*AICs, as shown in Table III.

There are two groups of *i*AICs for each record, each group containing 770 results. The two groups of *i*AICs for 11^{th} record are shown in the Fig. 7. The x-axis of Fig. 7 is the number of GNAR models. The y-axis of Fig. 7 (a) is *i*AIC of GNAR models of which the parameters are estimated by LS method and the y-axis of Fig. 7 (b) is *i*AIC of GNAR models of which the parameters are estimated by RALS method.

As shown in Fig. 7 (a), the minimum *i*AIC is obtained at the red point, where Number = 149, *i*AIC is 3.5768, and the MSE of the prediction series is 34.0597. If the parameters of the GNAR model are estimated by the RALS method, the *i*AIC result is 3.6337 and the MSE of the prediction series is 35.9406. As shown in Fig. 7 (b), the minimum *i*AIC is obtained at the blue point, where Number = 86, *i*AIC is 3.5508, and the MSE of the prediction series is 33.9517. If the parameters of the GNAR model are estimated by the LS method, the *i*AIC result is 3.6434 and the MSE of the prediction series is 38.8320.

According to Table III, the results can be divided into four groups. In the first group, when the GNAR model adopts the RALS estimator, if its *i*AIC takes the minimum value, its MSE of the prediction series using the model also takes the minimum value, which occurred 13 times. In the second group, when the GNAR model adopts LS estimator, if its *i*AIC takes the minimum value, its MSE of the prediction series using the model also takes the minimum value, which

occurred 3 times. In the third group, when the GNAR model adopts the RALS method, its *i*AIC takes the minimum value, but its MSE of the prediction series using the model is not the minimum value, which occurred only once. In the fourth group, when the GNAR model adopts the RALS method, although its *i*AIC is not the minimum value, its MSE of the prediction series using the model is the minimum value, which occurred 6 times. Furthermore, when the MSE is the minimum value, the *i*AIC of the model using the RALS method must be smaller than that of the model using the LS method. Therefore, the parameter defined, i.e., *i*AIC, can be effectively used to determine the order of the GNAR model.

TABLE III
*i*AICs OF GNAR MODELS AND MSEs OF PREDICTION SERIES

No.	GNAR Model	*i*AIC		MSE	
		LS	RALS	LS	RALS
1	765	1.8564	**1.8534**	6.3017	**6.2848**
	767	1.8572	1.8613	6.2950	6.3226
2	16	4.1687	**4.1509**	75.9160	**63.2321**
	17	4.1883	4.1661	76.7056	64.0166
3	704	3.2482	2.8917	25.2387	**17.1949**
	754	**2.7498**	2.9313	18.9489	19.0450
4	661	3.7480	**3.6883**	41.7453	**39.0159**
	756	3.7083	3.7035	40.2138	40.1198
5	597	4.4528	4.4044	89.3864	**87.8292**
	757	**4.3191**	4.4323	93.0499	93.1106
6	37	3.1185	3.0835	25.3443	**24.5728**
	138	**2.9284**	3.0919	29.1436	29.7318
7	39	1.7584	**1.7066**	5.8009	5.4559
	470	1.7096	2.0905	**4.8379**	7.0401
8	765	5.5836	**5.5201**	278.9744	**260.4055**
	766	5.5857	5.5470	287.0549	273.1585
9	398	6.9005	**5.5360**	1108.4019	**260.6138**
	581	6.3302	5.5768	645.3006	274.2127
10	*758*	*5.1252*	*5.0856*	*165.6604*	*158.4550*
	728	***4.9870***	*5.3222*	***137.7130***	*211.2578*
11	86	3.6434	**3.5508**	38.8320	**33.9517**
	149	3.5768	3.6337	34.0597	35.9406
12	152	4.3037	**4.0572**	69.2739	**53.9594**
	188	4.1604	4.2343	62.2769	66.9795
13	685	6.2727	**5.0227**	506.9140	**148.4276**
	67	5.3148	6.6199	194.2124	761.7608
14	547	4.1113	**4.0859**	65.1733	**61.6534**
	442	4.1080	4.0864	65.2370	61.7297
15	759	4.5411	4.5410	92.9383	**92.8528**
	42	**4.5144**	4.5482	93.0820	95.2946
16	658	3.7971	3.7814	43.9967	**43.3322**
	336	**3.7590**	3.7863	43.6818	44.0482
17	*761*	***5.3111***	*5.3154*	***201.0017***	*201.9218*
	762	*5.3168*	*5.3205*	*202.4222*	*203.1690*
18	754	4.1252	**4.0854**	61.1877	**58.8160**
	756	4.1253	4.0899	61.1417	59.0295
19	*472*	*4.3102*	*4.3069*	*75.4830*	*74.2147*
	579	***4.2853***	*4.3838*	***72.2323***	*79.5114*
20	332	4.7148	**4.7077**	115.0621	**111.7893**
	231	4.7138	4.7090	115.7256	113.6226
21	759	4.1219	**4.1190**	60.8581	**60.6689**
	758	4.1334	4.1326	61.5682	61.5119
22	755	3.8781	**3.8768**	48.7600	**48.6079**
	754	3.8787	3.8782	48.8409	48.7415
23	436	3.9307	3.9278	52.2209	**51.9715**
	544	**3.9190**	3.9296	52.8133	53.4795

Note: Bold indicates that the indice performance of the coresponding method is the best.

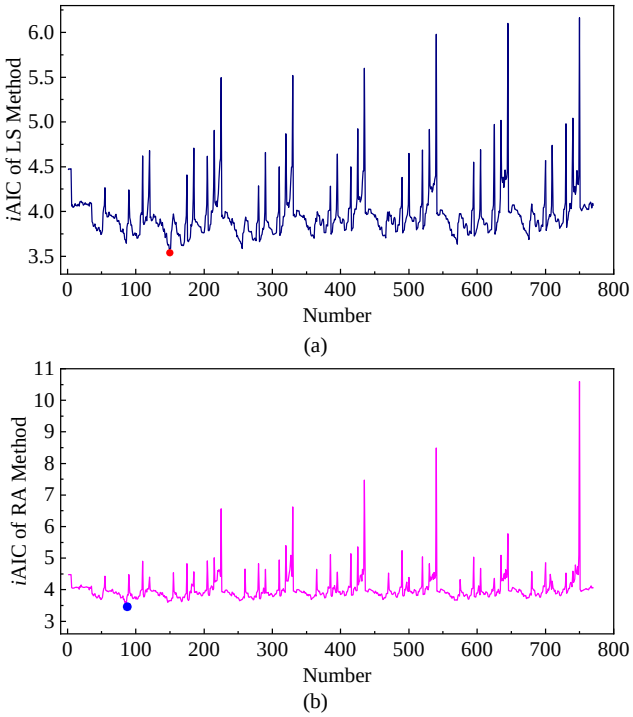

Fig. 7. Results of *i*AIC for GNAR Models

V. CONCLUSION

For the prediction of complex nonlinear time series, the GNAR model is proposed. The RALS estimator and its process are proposed to deal with the situation that time series will inevitably be contaminated by outliers. Since more nonlinear terms are considered in the GNAR models, the computational complexity of the GNAR model is modelled and introduced into the AIC to propose an improved AIC, named *i*AIC. Finally, the performance of GNAR model and its parameter estimation method and order determination method have been evaluated experimentally and quantitatively by using the established dataset. The results show that the MES of the prediction series of the GNAR model with RALS estimator are relatively small. The MES of the LS estimator is 325.3% higher than that of the RALS estimator at most. The effectiveness of the *i*AIC reaches 82.61%. Consequently, the proposed order determination method for GNAR model is effective and the high fitting accuracy of the GNAR model with RALS estimator has been achieved.

Although the GNAR model can fit various time series well, the performance of the GNAR model needs to be further improved, especially when the time series has obvious trend items and changes significantly. Therefore, on the one hand, the adaptability of GNAR model can be improved by optimizing the structure of GNAR model. On the other hand, the stability of time series can be improved by trend item separation and data compression which is used to reduce the magnitude of data changes. Finally, the fitting accuracy of GNAR models has to be further improved.

REFERENCES

[1] P. s. D. Online. (2020, 2020-10-24). Ministry of Transport: Support and Encourage the Development of New Forms of Transport. Available: http://hi.people.com.cn/auto/n2/2020/1024/c336672-34371434.html

[2] Q. TAN, M. DAI, and S. MEI, "Research on electric vehicle carbon quota and demand response in electric power system dispatching," (in Chinese), Power System and Clean Energy, vol. 37, no. 7, pp. 79-86, July 2021.

[3] H. Xu, F. Duan, and P. Pu, "Dynamic bicycle scheduling problem based on short-term demand prediction," Applied Intelligence, vol. 49, no. 5, pp. 1968–1981, May 2019.

[4] W. Jiang, "Bike sharing usage prediction with deep learning: a survey," Neural Computing and Applications, vol. 34, no. 18, pp. 15369-15385, September 2022.

[5] B. Zhang, "Research on location model of bike sharing launch points," Chang'an University, Xi'an, China, 2019.

[6] L. Wang, "Traffic flow prediction and rental location of bike-sharing system," Beijing Jiaotong University, Beijing, China, 2020.

[7] C. Zheng, C. Yang, Y. Xiao, L. Luo, and W. Wang, "Demand prediction analysis of electric vehicle charging facilities based on data fusion," Jiangxi Electric Power, vol. 40, no. 2, pp. 9-12, February 2016.

[8] Z. YANG and Z. GAO, "Location method of electric vehicle charging station based on data driven," Journal of Transportation Systems Engineering and Information Technology, vol. 18, no. 5, pp. 143-150, October 2018.

[9] F. Tian and H. Chen, "Research on planning of electric vehicle charging station considering user choice preference," Computer Engineering and Applications, vol. 58, no. 15, pp. 294-301, April 2022.

[10] M. Suleyman, C. Z. Abidin, and Ö. Eren, "Location and coverage analysis of bike- sharing stations in university campus," Business Systems Research, vol. 9, no. 2, pp. 80-95, July 2018.

[11] I. Frade and A. Ribeiro, "Bike-sharing stations: A maximal covering location approach," Transportation Research Part A: Policy and Practice, vol. 82, pp. 216-227, December 2015.

[12] H. Liu, Y. He, T. Song, and P. Xu, "Demand forecast and scheduling optimization of shared bicycles," (in Chinese), Science Technology and Engineering, vol. 21, no. 35, pp. 15247-15254, December 2021.

[13] X. Gao and G. M. Lee, "Moment-based rental prediction for bicycle-sharing transportation systems using a hybrid genetic algorithm and machine learning," Computers & Industrial Engineering, vol. 128, pp. 60-69, February 2019.

[14] M. F. Akbar, E. Octaviany, T. Toharudin, and B. N. Ruchjana, "The application of vector autoregressive integrated with exogenous variable to model the relationship between covid-19 positive numbers with population growth rate," presented at the 10th International Conference and Workshop on High Dimensional Data Analysis, Sanur-Bali, Indonesia, Oct. 12 - 15, 2021. Available: http://dx.doi.org/10.1088/1742-6596/1722/1/012081

[15] M. Alawiyah, D. A. Kusuma, and B. N. Ruchjana, "Application of generalized space time autoregressive integrated (GSTARI) model in the phenomenon of covid-19," presented at the 10th International Conference and Workshop on High Dimensional Data Analysis, Sanur-Bali, Indonesia, Oct. 12 - 15, 2021. Available: http://dx.doi.org/10.1088/1742-6596/1722/1/012035

[16] J. Lei, X. Wang, Y. Zhang, L. Zhu, and L. Zhang, "Policy and law assessment of COVID-19 based on smooth transition autoregressive model," Complexity, vol. 2021, pp. 1-13, January 2021.

[17] J. A. L. Marques, F. N. B. Gois, J. Xavier Neto, and S. J. Fong, "Forecasting COVID-19 time series based on an autoregressive model," in predictive models for decision support in the COVID-19 crisis (Springer Briefs in Applied Sciences and Technology: Springer Science and Business Media Deutschland GmbH, 2021, pp. 41-54.

[18] A. Meimela, S. S. S. Lestari, I. F. Mahdy, T. Toharudin, and B. N. Ruchjana, "Modeling of covid-19 in Indonesia using vector autoregressive integrated moving average," presented at the 10th International Conference and Workshop on High Dimensional Data Analysis, Sanur-Bali, Indonesia, Oct. 12 - 15, 2021. Available: http://dx.doi.org/10.1088/1742-6596/1722/1/012079

[19] Sunayana, S. Kumar, and R. Kumar, "Forecasting of municipal solid waste generation using non-linear autoregressive (NAR) neural models," Waste Management, vol. 121, pp. 206-214, February 2021.

[20] A. Razzaq, A. Sharif, A. Najmi, M. Tseng, and M. K. Lim, "Dynamic and causality interrelationships from municipal solid waste recycling to economic growth, carbon emissions and energy efficiency using a novel bootstrapping autoregressive distributed lag," Resources, Conservation and Recycling, vol. 166, pp. 1-14, March 2021.

[21] M. B. Rahmoune, A. Hafaifa, A. Kouzou, X. Chen, and A. Chaibet, "Gas turbine monitoring using neural network dynamic nonlinear autoregressive with external exogenous input modelling," Mathematics and Computers in Simulation, vol. 179, pp. 23-47, January 2021.

[22] N. Khraief, M. Shahbaz, M. K. Mahalik, and M. Bhattacharya, "Movements of oil prices and exchange rates in China and India: New evidence from wavelet-based, non-linear, autoregressive distributed lag estimations," Physica A: Statistical Mechanics and its Applications, vol. 563, pp. 1-15, February 2021.

[23] C. Berninger, A. Stocker, and D. Rugamer, "A Bayesian time-varying autoregressive model for improved short-term and long-term prediction," Journal of Forecasting, vol. 41, no. 1, pp. 181-200, January 2022.

[24] X. Zhang, H. Wen, T. Yamamoto, and Q. Zeng, "Investigating hazardous factors affecting freeway crash injury severity incorporating real-time weather data: Using a Bayesian multinomial logit model with conditional autoregressive priors," Journal of Safety Research, vol. 76, pp. 248-255, February 2021.

[25] F. Zhang, P. Li, L. Gao, Y. Liu, and X. Ren, "Application of autoregressive dynamic adaptive (ARDA) model in real-time wind power forecasting," Renewable Energy, vol. 169, pp. 129-143, May 2021.

[26] M. O. Yano, L. G. G. Villani, S. da Silva, and E. Figueiredo, "Autoregressive model extrapolation using cubic splines for damage progression analysis," Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 43, no. 1, pp. 1-19, January 2021.

[27] Y. Zhang, D. Zhang, and E. J. Miller, "Spatial autoregressive analysis and modeling of housing prices in city of toronto," Journal of Urban Planning and Development, vol. 147, no. 1, pp. 1943-5444, March 2021.

[28] S. A. Zulkeflee, F. S. Rohman, S. Abd Sata, and N. Aziz, "Autoregressive exogenous input modelling for lipase catalysed esterification process," Mathematics and Computers in Simulation, vol. 182, pp. 325-339, April 2021.

[29] W. Li et al., "Output Optimization of nuclear power steam turbine based on nonlinear autoregressive neural network and random forest algorithm," (in Chinese), Proceeding of the CSEE, vol. 41, no. 02, pp. 409-416, October 2021.

[30] W. Xiong, W. Sun, and J. Ma, "Autocorrelation constrained NARX dynamic soft sensing model," (in Chinese), Kongzhi yu Juece/Control and Decision, vol. 35, no. 4, pp. 816-822, April 2020.

[31] L. Yao, R. Ma, and H. Wang, "Baidu index-based forecast of daily tourist arrivals through rescaled range analysis, support vector regression, and autoregressive integrated moving average," Alexandria Engineering Journal, vol. 60, no. 1, pp. 365-372, February 2021.

[32] H. Allende, J. Galbiati, and R. Vallejos, "Robust image modeling on image processing," Pattern Recognition Letters, vol. 22, no. 11, pp. 1219-1231, September 2001.

[33] X. Kong, "M-estimation for moderate deviations from a unit root," Communications in Statistics - Theory and Methods, vol. 44, no. 3, pp. 476-485, February 2015.

[34] W. Xu and H. Zhao, "Robust constrained recursive least M-estimate adaptive filtering algorithm," Signal Processing, vol. 194, pp. 1-12, May 2022.

[35] O. H. Bustos and V. J. Yohai, "Robust estimates for ARMA models," Journal of the American Statistical Association, vol. 81, no. 393, pp. 155-168, March 1986.

[36] O. H. Bustos, M. Ruiz, S. Ojeda, R. Vallejos, and A. C. Frery, "Asymptotic behavior of RA-estimates in autoregressive 2D processes," Journal of Statistical Planning and Inference, vol. 139, no. 10, pp. 3649-3664, October 2009.

[37] R. Vallejos, S. Ojeda, and O. Bustos, "Robust RA and GM Estimators for Spatial AR Models," 2007.

[38] S. M. Ojeda, R. O. Vallejos, and M. M. Lucini, "Performance of robust RA estimator for bidimensional autoregressive models," Journal of Statistical Computation and Simulation, vol. 72, no. 1, pp. 47-62, January 2002.

[39] S. M. Ojeda, H. Torre, and M. Allende, "Robust RA estimators in AR-2D models for images," in Proceedings of the 2nd Latino-American Seminar on Radar Remote Sensing, Santos, Sao Paulo, Brazil, 1998, pp. 1153-1159.

[40] J. Shi, S. Zhu, and R. Chen, "order selection for general expression of nonlinear autoregressive model based on multivariate stepwise regression," presented at the 2017 4th International Conference on Mechanical, Materials and Manufacturing, Atlanta, GA, United states, Oct. 25 - 27, 2017. Available: http://dx.doi.org/10.1088/1757-899X/272/1/012010

[41] M.-B. Hossain, J. Moon, and K. H. Chon, "Estimation of ARMA model order via artificial neural network for modeling physiological systems," IEEE Access, vol. 8, pp. 186813-186820, October 2020.

[42] J. Moon, M. B. Hossain, and K. H. Chon, "AR and ARMA model order selection for time-series modeling with ImageNet classification," Signal Processing, vol. 183, June 2021.