مجلة

كلية التـــراث الجامعة

مجلة علمية محكمة

متعددة التخصصات نصف سنوية

العدد الثامن والثلاثون

15 شباط 2024

ISSN 2074-5621

رقم الايداع في دار الكتب والوثائق 719 لسنة 2011

مجلة كلية التراث الجامعة معترف بها من قبل وزارة التعليم العالي والبحث العلمي بكتابها المرقم (ب 3059/4) والمؤرخ في (4/7 /2014)

A Comparison of Traditional and Optimized Multiple Grey Regression Models with Water Data Application

Mohammad Ismael Rasheed **Omar Abdulmohsin Ali**

University of Baghdad

Abstract

Grey system theory is regarded as a multidisciplinary scientific approach, which deals with systems that have partial information (such as: small samples). Grey modeling as a vital portion such theory that yields a sufficient results despite of limited amount of data. Grey Models can be divided into two types; univariate and multivariate grey models. Univariate grey model GM (1,1) is the cornerstone of this theory, but it doesn't take the relative factors in account. While multivariate grey models GM(1,M) takes those factor in account. However, they have a complex structure and some defects in " modeling mechanism", "parameter estimation "and "model structure". Therefore, GM(1,M) in the traditional version was submitted to many trials of optimizations to getting rid this defects.

This paper aims to:

- introduce the characteristics of the traditional GM(1,M) as well as its problem.
- impose a method to overcome such problem, through two optimized multivariable grey model of one order derivative equation.
- present an Optimized Grey Model abbreviated as OGM(1, M) by adding the linear correction term h_1(M-1)and greyaction quantity term (h_2) to the traditional model GM(1,M).
- use an Optimized Background value Grey Model OBGM(1,M) by optimizing the Background value of the last model OGM(1,M).
- apply real data-set represents water consumption in Baghdad at the period (2016-2022) to compare both optimized models with the traditional model.
- use the mean absolute percentage error (MAPE) and the determination coefficient R^2.

The results of comparison showed that both OGM(1, M) and OBGM(1,M) have superiority compared with the traditional model GM(I,M).

Paper type: Research paper.

Keywords: Ordinary least squares, Grey Models, Multiple Regression Model, Background value, mean absolute percentage error criterion.

1. Introduction

Time series data are taken a sequence and equal intervals of time.The prediction in time series refers to the process by which we can forecast the future values of a system from what is available data in the past and present. It is a familiar matter to use the ordinary statistical methods which are defined as"Box Jenkins"methods. In accurate forecasting of time series,but when the ability of this methods is being weak,because of its fuzzy and short period (poor information),the researchers turn to forecasting models with different background and mathematical rules, then grey systems models were found to be the suitable solution for such problems(Geng,J,Ye,D.and Luo,P,2015).Grey system theory mainly aims to provide techniques and thoughts to analyse the solution of the complex eigen(latent)systems.The pioneer Gulong Deng enabled to treat the lack of information by an important mathematical method based on transform the irregular raw data to a new more organized series through greygenerating technique and construct a derivative model named greyModel(GM) by using at least(4)entries data to replace difference modeling which require large quantities of data(Julong,,1989).

Grey system theory has a new important methodology which is widely used in the last three decades especially in the fields of prediction and control. The major achievement was the eighties of the last century by "Julong Deng" to treat the uncertain systems with incomplete information and Inaccuracies in data (Kayacan, et al, 2010).

1.1 Literature Review

Julong Deng has defined Grey system theory according to the concept of black box as the system which contain a partially known information and partially unknown information (Cheng, M. et al. 2022). The colour which describe the system refers to the clearness (amount of accuracy) of information and its abundance, so the system which has clear and enough amount of information described as white system, while the system which has unclear and few amount of information (less than 4) is described as black system, but the system with noisy and poor information (but greater than 4 information) is described as grey systems. In real life every system can be regarded as grey system because amount of uncertainty is usually exist, this belong to the noisy in both the external and the internal environment of the system in addition to limitation of our cognitive, so the information that we can reach about the system is always uncertain and limited in its scope. (Kayacan, et al, 2010).

The official appearance of grey system theory is represented by publishing the first paper on greysystems theory "The problem of grey systems" in the international journal of North Holland "systems and control letters**"** by the Chinees professor Gulong Deng in 1982, after that the journal of Huazhong university of science and technology published the first paper on greysystem theory by J. Deng also in chinese language, and as soon as these two works appears many scholars and scientific researchers around the world draw a great interest on this theory and many well-known scientists has strongly supported the validity and livelihood of this theory and many young man and woman scholars from china, united state, England, Germany, Japan, Australia, Canada, Rusia, Netherland, Turkey, Iran, and others has actively applied the theoretical aspect of the theory in various fields of applications such as in agriculture, Industrial, Economics, social, financial, transportation, scientific and technology, mechanical, meteorogical, ecology, geological, medical, military affairs, etc., the successful applications of greysystem theory in those great many fields have won the attention of the international world of learning.(Deng, 1982).

In 1989 the British journal "The journal of grey systems theory" was launched and published numerous researches and applications on grey systems theory in England. In 1997, a Chinese publication, named Journal of Grey System, is launched in Taiwan. It is later in 2004 that this publication becomes all English. Additionally, a new journal, named Grey Systems: Theory and application, edited by the faculty of Institute for Grey Systems Studies at University of Aeronautics and Astronautics **NUAA**, will be launched by Emerald in 2011. There are currently over one thousand different professional journals in the world that have accepted and published papers in grey systems theory. Many finest universities around the world offer courses in grey systems theory. For instance, Nanjing University of Aeronautics and Astronautics (NUAA) not only offers such courses to PhD level and master level students, but also provides a service course on grey systems to all undergraduate students in different majors. Huazhong University of Science and Technology, NUAA, Wuhan University of Technology, Fuzhou University, and several universities in Taiwan recruit and produce PhD students focusing on the research in grey systems. It is estimated that well-over several thousands of graduate students from around the world employ the thinking logic and methodology of grey systems in their research and the writing of their dissertations. Many publishers from around the world, such as Science Press (mainland China), Press of National Defense Industry (mainland China), Huazhong University of Science and Technology Press (mainland China), Taiwan Quanhua Science and Technology Books, Science and Engineering Press of Japan, the IIGSS Academic Press and Taylor and Francis Group (USA), Springer-Verlag (Germany),etc., have published over 100 different kinds of monographs in grey systems.

In 2006, the conference of grey systems theory and applications was held successfully in Beijing with financial sponsorship provided by Chinese Center for Advanced Science and Technology, headed by Li Zhengdao (a Nobel Laureate), Zhou Guangzao and Lu Yongxian (academicians of Chinese Academy of Sciences). In 2008 and 2010, the 16th and 19th National Conference of Grey Systems once again was financially supported by Chinese Center for Advanced Science and Technology. Many important international conferences, such as the Conference on Uncertain Systems Modeling, Systems Prediction and Control, Congress of World Organization of Cybernetics and Systems, IEEE conferences on System, Man, and Control, International Conferences of Computers and Industrial Engineering, etc., have listed grey systems theory as a special topic of interest. In August 2003, the 32nd International Conference on Computers and Industrial Engineering, held in Ireland, opened four different sessions for grey systems theory. in March 2005, the 13th Congress of the WOSC, held in Slovak, in October 2005, the IEEE Conference on Systems, Man, Control, held in Taiwan, China, in October 2006, the IEEE Conference on Systems, Man and Control, held in Montreal, Canada, in October 2007, the 14th Congress of the WOSC, held in Poland in September 2008, the IEEE Conference on Systems, Man and Control, held in Singapore in October 2008, the IEEE Conference on Systems, Man and Control, held in San Antonio, USA, in October 2009, etc. all arranged special topic sessions for grey systems research. Grey systems theory has caught the attention of many important international conferences and become a center of discussion at many international events.
Since 2010 numerous of researchers in china, USA, England, Roman, South Africa, Germany, Japan, Australia, Canada, Poland, Aspan, Cupa, Russia, Turkey, Irland, Iran and other countaries submitted thousands of researches and application on greysystems theory and in different fields which refers to its effective and successful application ability.But the most important works were what introduced by Tien (2012) who studied accuracy of GM(1,1) and GM(1,N) in different case studies and concluded that the GM(1,N) is incorrect.. Wu et al. (2015) developed a novel GM(1,N) model considering opposite direction AGO and predicted CO2 emission in BRICS countries. Zeng et al. (2016) stated that the GM(1,N) is associated with three defects such as parameter estimation, model structure and modeling mechanism. He proposed an optimization model for the GM(1,N) model called as OGM(1,N) by introducing a grey action quantity and a linear correction term and improved performance of GM(1,N) . Ye et al. (2020) developed a novel accumulative time-delay GM(1,N) model and analyzed the CO2 emissions from the transport vehicles in china. Based on the above studies, it is noticed that the multiple grey prediction models are developed by different authors and reduced errors in traditional GM (1, N) model.

2. Material and Methods

2.1 Characteristics of Uncertain systems

The essential characteristic of the uncertain systems are represented by **incompleteness** and **inadequacy** in their information which are belongs to the dynamics of system evaluations, the biological limitations of the human sensing organs, and the constraints of relevant economic conditions and technological availabilities.

2.1.1 Incomplete Information

The situation involving incomplete system information can have the following four cases (Liu, Forrest, and Yang, 2012)

i. The information about the elements (parameters) is incomplete;
ii. The information about the structure of the system is incomplete;
iii. The information about the boundary of the system is incomplete.
iv. The information on the system's behaviors is incomplete.

2.1.2 Inaccuracies in Data

The concepts of Uncertainty and Inaccuracy are roughly the same, both are refer to errors and deviation from the actual data values. The essential causes of inaccuracy can be categorized into three categories: the conceptual, level, and prediction types. (Liu and Forrest, 2010)

2.2 Accumulating and (Inverse Accumulating) generation operation (AGO) and (IAGO)

It is one of the most important components of grey systems theory, and the first procedure done on the positive irregular raw data $X^{(0)}$ to lessen its randomness to be more regular by transforming it to a new sequence $X^{(1)}$ with exponential increasing form useful in modeling process through sequential accumulating additions on the raw data sequence. In addition, through this accumulation we can discover the development tendency and laws of integration hidden in the chaotic original data (Wang, Zhang, Moon, and Sutherland, 1998) (Geng, Ye, and Luo, 2015).

To explain the this procedure assume the following origin sequence

$X^{(0)}$= [$x^{(0)}(1)$, $x^{(0)}(2)$, ….. , $x^{(0)}(n)$] , $n \geq 4$ ……..(1)

Then we can express this procedure as:

$X^{(1)}$ AGO :

$X^{(0)}$ ⟹

] (2) $x^{(1)}(n)$, .. , ... $x^{(1)}(2)$,) [

$x^{(1)}$ (1 = $X^{(1)}$

And we can write it mathematically as

$$X^{(1)}(k)= \sum_{m=1}^{k} x^{(0)}(m) \quad , \quad k= 1,2,3,\ldots\ldots\ldots n \qquad (3)$$

The inverse of the accumulative generation Operator (IAGO) is the reciprocal operation of AGO ,
so that:

$$x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \quad , \quad k= 2,3,\ldots.n \qquad (4)$$

where $x^{(0)}(1) = x^{(1)}(1)$

Grey

Modeling: 2.3

On the subject of modeling, for biological phenomena, biological medicine and economy, it had better build a differential model, this has been impossible in the past. In grey system theory a dynamic model with a group of differential equations is built based on the generating series rather than on the raw data. This model is called grey differential model (GM). In grey modeling only a few data (as few as 4) are needed to distinguish a GM model. (Deng, 1989).

2.3.1 Grey Model GM (1,1):

The first order grey model in one variable GM(1,1) stand for the base stone of grey systems theory and grey prediction model which is more used in numerous fields for its simple calculation procedure and high prediction accuracy, its mainly thoughts represented by the full benefit of the minimum bound of the available data and the predicting the system motion with the poor and uncertain information.

The dynamic of this model is represented by one independent variable $X^{(0)}$ stands for a time series of size (n), all its value must be positive. We can express the sequence as:

$$X^{(0)} = [\, x^{(0)}(1),\ x^{(0)}(2),\ \ldots,\ x^{(0)}(n) \,],\quad n \geq 4$$

and $X^{(1)} = [\, x^{(1)}(1),\ x^{(1)}(2),\ \ldots,\ x^{(1)}(n) \,]$

represent the Accumulation generating sequence (AGO) where;

$$X^{(1)}(\mathbf{k}) = \sum_{m=1}^{k} x^{(0)}(m)\ ,\quad k = 1,2,3,\ldots\ldots n \qquad ;\ \text{then}$$

$$\frac{\mathbf{dx^{(1)}}}{\mathbf{dt}} + \mathbf{a\, x^{(1)} = b} \qquad (5)$$

Is called the ' grey differential equation of one order" (whitening equation) and:-

$$X^{(0)}(K) + a\, Z^{(1)}(k) = b \qquad \ldots\ldots\ (K = 1,2,\ldots..n) \qquad (6)$$

It is referred to as the basic form of the GM (1, 1) model, (a) is the development coefficient and (b) is greyaction quantity which reflects the uncertain external grey factors effects on system development trend, and it is considered a grey number also (Liu, S., Forrest, J. and Yang, Y., 2012). $\mathbf{Z^{(1)}}$ is a new exponential more smooth sequence, which is generated from $\mathbf{x^{(1)}}$ by taking the adjacent neighbor means

$$\mathbf{Z^{(1)} = [\, z^{(1)}(1),\ z^{(1)}(2),\ \ldots,\ z^{(1)}(n) \,]} \qquad (7)$$

$$\mathbf{Z^{(1)}(K) = 0.5\, X^{(1)}(K) + 0.5\, X^{(1)}(K-1),\quad K = 1,2,\ldots\ldots,n} \qquad (8)$$

Now; by using the sequences $\mathbf{X^{(0)}}$, $\mathbf{X^{(1)}}$ and $\mathbf{Z^{(1)}}$, we can estimate the least square parameter sequence $\hat{p}$ through finding the parameters which make the sum square errors as small as possible.

the model (6) can be set up as follows:

$$\mathbf{X^{(0)}(K) = a\, Z^{(1)}(k) - b} \qquad (9)$$

and errors is be given as (Mohammed Ali, N. S. and Mohammed, F. A., 2018)**:**

$$\mathbf{e = X^{(0)}(K) - a\, Z^{(1)}(k) + b} \qquad (10)$$

and the sum square errors is given as: $S = \sum_{k=2}^{n}(x^{(0)}(k) + a\,Z^{(1)}(k) - b)^2$ then by applying the partial derivative with respect to **a** and **b** and equaling the result to zero we get the OLS parameter estimations. (Mao-lin, Yong, and Yan-qiu, 2007).

By Using matrices, The set of natural equations for this model (9) can be expressed in the form $Y = B\hat{p}$

Where

$$\hat{p} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad B = \begin{bmatrix} -Z^{1}(2) & 1 \\ -Z^{1}(3) & 1 \\ \vdots & \vdots \\ -Z^{1}(n) & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} X^{0}(2) \\ X^{0}(2) \\ \vdots \\ X^{0}(2) \end{bmatrix} \qquad (11)$$

The error vector is given in the form: $U = Y - B\hat{p}$ (12)

And the sum of squares of error (S)

$S = U^{T}U = [Y - B\hat{p}]^{T}[Y - B\hat{p}]$

$S = \mathbf{Y^{T}Y} - 2P^{T}B^{T}Y + P^{T}B^{T}BP$

and equaling the result to zero we get the OLS $\hat{p}$

then by applying the partial derivative with respect to parameter estimation vector as:

$\hat{p} = (a, b)^{T} = (B^{T}B)^{-1}B^{T}Y$ .... (13)

We mention here that OLS method is best than other parameter estimation method such as the weighted least square method (WLS), total least square method (TLS) and the Gradient descent method (GD) in grey modelling, and its parameter gives less MAPE amount. (Mohammed Ali and Mohammed, F., 2018)

According to the whitening equation (5) we can calculate the prediciton values $\hat{x}^{(1)}(k)$ which is called the time response function as (Liu and Forrest, 2010)

$$\hat{\mathbf{x}}^{(1)}(\mathbf{K+1}) = \left[x^{(0)}(1) - \frac{\mathbf{b}}{\mathbf{a}}\right]\mathbf{e}^{-\mathbf{ak}} + \frac{\mathbf{b}}{\mathbf{a}} \qquad (14)$$

then we will apply the Inverse Accumulating Generating Operator **IAGO** to find the predicted values of original data.

$$\hat{\mathbf{x}}_{\mathbf{1}}^{(\mathbf{0})}(\mathbf{k+1}) = \hat{\mathbf{x}}_{\mathbf{1}}^{(\mathbf{1})}(\mathbf{k+1}) - \hat{\mathbf{x}}_{\mathbf{1}}^{(\mathbf{1})}(k) \qquad (15)$$

2.3.2 The traditional GM (1, M)

The multivariate grey model GM(1,M) is an important causal relationship forecasting model. This model is consist of a system characteristic sequence (dependent variable) and (M-1) related factor sequences(independent variable). The modeling process takes the effect of the relevant factors in the system change in account, and makes full use of the limited information contained in the relevant factore sequencs so that the multivariate grey forecasting model makes up the deficiency of the one fold structure and limited simulation capacity of uni-variate models. (Liu and Forrest, 2010) (Xie, Yan, Wu, Liu, Bai, Liu, and Tong, 2021).

To describe this model, consider $X_1^{(0)}$ the system characteristic sequence,

$$X_1^{(0)} = (X_1^{(0)}(1), X_1^{(0)}(2), X_1^{(0)}(3), \ldots\ldots\ldots\ldots, X_1^{(0)}(n)), \quad n \geq 4 \qquad (16)$$

and $X_j^{(0)}, (j = 2,3, \dots, M)$ be the independent variable sequences which have a highly correlation with the sequence $X_1^{(0)}$

$$X_j^{(0)} = (X_j^{(0)}(1) , X_j^{(0)}(2) , X_j^{(0)}(3), \dots\dots\dots, X_j^{(0)}(n)) \quad , \quad n \geq 4 , \; j = (1,2, \dots\dots M) \tag{17}$$

Let $X_j^{(1)}$ represent the cumulative generation operator 1-AGO where

$$X_j^{(1)} = (X_j^{(1)}(1) , X_j^{(1)}(2) , X_j^{(1)}(3), \dots\dots\dots, X_j^{(1)}(n)) \quad , \quad n \geq 4 , j = (1,2, \dots\dots M) \tag{18}$$

Where; $X_j^{(1)}(k) = \sum_{s=1}^{k} X_j^{(0)}(s)$, k= 1,2,………n
(19)

Let $Z_1^{(1)}$ represent the average series generated by the successive neighbors of $X_1^{(1)}$
where:

$$Z_1^{(1)} = (Z_1^{(1)}(2), Z_1^{(1)}(3) , \dots\dots\dots, Z_1^{(1)}(n)) \tag{20}$$

21) (n ,........2 ,3 k = **)** , $Z_1^{(1)}(k) = 0.5 \,\mathbf{(}\, X_1^{(1)}(k) + X_1^{(1)} (k-1$

where;

Then :

$$X_1^{(0)}(k) + a\, Z_1^{(1)}(k) = \sum_{i=2}^{M} b_i\, x_1^{(1)}(k) \tag{22}$$

Represent the conventional multivariate gray model GM (1, N).

Where (a) represents the system development coefficient, ($\sum_{\mathbf{i=2}}^{\mathbf{M}} \mathbf{b_i\, x_1^{(1)}(k)}$) represents the driving term, and (bi) represents the driving coefficient.

And p ̂ = [a, b2, …, bn]^T represents the series of parameters that must be estimated by the formula (13)

$$\hat{p} = [\, B^T B \,]^{-1} B^T Y$$

$$B = \begin{pmatrix} -z_1^1(2) & x_2^1(2) & x_3^1(2) & \dots & x_N^1(2) \\ -z_1^1(3) & x_2^1(3) & x_3^1(3) & \dots & x_n^1(3) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -z_1^1(n) & x_2^1(n \dots) & x_3^1(n) & \dots & x_n^1(n) \end{pmatrix} \qquad Y = \begin{pmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{pmatrix} \tag{23}$$

For the parameter series $\widehat{P}$ = [a, b2,…,bn]^T, the whitening equation, sometimes called the shadow equation, for the first-order grey multivariate model GM(1,N) takes the following form:

$$\frac{\mathbf{dx^{(1)}}}{\mathbf{dt}} + a\, x_1^{(1)}(k) = \sum_{i=2}^{M} b_i\, x_1^{(1)}(k) \;\dots\dots\dots\dots\dots (24)$$

The series of approximate predictive values $\hat{x}_1^{(1)}$(k+1) called the time response function of the model GM(1,N) is calculated with the formula

$$= [\ X_1^{(0)}(k) - \frac{1}{a}\sum_{i=2}^{M} b_i x_i^{(1)}(k+1)]\, e^{-ak} + \frac{1}{a}\sum_{i=2}^{M} b_i x_i^{(1)}(k+1) \quad (25) \quad k+1)\ (\hat{x}_1^{(1)}$$

Where $x_1^{(1)}(0) = x_1^{(0)}(1)$ which is the initial value of the model GM(1,N) (He, and Tao, 2014)

The recovery value of the inverse of accumulation is given by the model GM (1,N) as in formula (15)

2.3.3 The optimized multivariate Grey model OGM (1,M):

Although the multivariate grey model GM(1,M) takes into account the influence of other factors on the system characteristic sequence (dependent variable sequence) and addresses the shortcomings exist in the univariate grey mode GM(1,1) because it doesn't reflect the effect of external environment changes on the direction of the system's evolution. However, GM(1,M) has some important defects, because it has a more complex structure than GM(1,1), which may cause significant errors in simulation and prediction work. as a result of numerous studies and applications for the GM(1,N) model. Three obvious important defects concerning with the modeling mechanism, parameter estimation and model structure were observed, which led to decrease the model accuracy, so a new model was proposed by introducing the linear correction term h_1(k-1) and the " Grey action quantity term " (h_2) to the traditional GM(1,N) model. The new model has a consistent and reasonable modeling process and a more stable structure, which solves the three problems of the traditional model, in addition to being a completely compatible model with the univariate discrete grey model DGM(1,1) As well as with the multiple dynamic prediction model for GM (0, N), the new developed model has proven its effectiveness through its use in simulating tensile strength of steel material. Therefore, the relative change rate and percentage of prediction errors for the new model are 0.0707% and 5.7369%. Compared to the traditional GM (1,M) model and the traditional GM (1,1) model, which were 6.0011%, 18.4228%, 1.1020%, and 12.05190%, respectively. The results showed that the new model has better performance . On one hand, it acknowledges the validity of the analysis of defects, and on the other hand validates effectiveness and efficiency of restructuring the traditional GM(1,M), so its form will be as (Li, C., Li, Y. and Xing, J., 2023) & (Zeng, B., Ma, X. and Shi, J., 2020).

$$X_1^{(0)}(k) + a\, Z_1^{(1)}(k) = \sum_{i=2}^{M} b_i x_1^{(1)}(k) + h_1(k-1) + h_2 \ \ldots\ldots\ldots\ldots (26)$$

This model is called the developed model of traditional model GM (1,M) and is symbolized as OGM(1, M).

Where;

The linear correction term h_1(k-1) reflects the linear relationships between the dependent variable and the independent variables and the grey quantity term (h_2) describes the relationships of changing data in the series of the dependent variable. The strengths and weaknesses of these relationships can be modified through their transactions.

According to grey model OGM(1,N), the vector of parameters $\hat{p} = [a, b2, \ldots, bn, \mathbf{h_1}, \mathbf{h_2}]^T$ is given by:

$$[\ \mathbf{B^T B}\ ]^{-1}\ \mathbf{B^T Y} = \hat{\mathbf{p}}$$

Then:

$$B = \begin{pmatrix} x_2^{(1)}(2) & x_3^{(1)}(2) & \dots & x_N^{(1)}(2) & -Z_1^{(1)}(2) & 1 \\ x_2^{(1)}(3) & x_3^{(1)}(3) & \dots & x_n^{(1)}(3) & -Z_1^{(1)}(3) & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_2^{(1)}(n) & x_3^{(1)}(n) & \dots & x_n^{(1)}(n) & -Z_1^{(1)}(n) & n-1 \end{pmatrix} \begin{matrix} 1 \\ 1 \\ \vdots \\ 1 \end{matrix} \quad , \quad Y = \begin{pmatrix} X_1^{(0)}(2) \\ X_1^{(0)}(3) \\ \vdots \\ X_1^{(0)}(n) \end{pmatrix} \qquad (27)$$

$$\widehat{X}_1^{(0)}(K) = \sum_{i=2}^{M} b_i \widehat{X}_1^{(1)}(k) - a\, Z_1^{(1)}(k) + h_1(k-1) + h_2 \qquad \dots\dots\dots\dots(28)$$

The difference model of the model is called OGM (1,N).

For the same developed model, the time response expression is given as follows:

$$\widehat{\mathbf{X}}_1^{(1)}(\mathbf{k}) = \sum_{t=1}^{k-1}[\mu_1 \sum_{i=2}^{M} \mu_2^{t-1} b_i x_i^{(1)}(k-t+1)] + \mu_2^{k-1}\widehat{\mathbf{X}}_1^{(1)}(\mathbf{1}) + \sum_{j=0}^{k-2} \mu_2^{j}[(k-j)\mu_3 + \mu_4],$$
$$(\mathbf{k}= 2,3,\dots\dots\dots\mathbf{n}) \qquad \dots\dots\dots\dots(29)$$

Where:

$$\boldsymbol{\mu_1} = \frac{\mathbf{1}}{\mathbf{1+0.5a}},\ \boldsymbol{\mu_2} = \frac{\mathbf{1-0.5a}}{\mathbf{1+0.5a}},\ \boldsymbol{\mu_3} = \frac{\mathbf{h_1}}{\mathbf{1+0.5a}},\ \boldsymbol{\mu_4} = \frac{\mathbf{h2-h1}}{\mathbf{1+0.5a}},\ \text{and } \widehat{\mathbf{X}}_1^{(1)}(\mathbf{k}) = \mathbf{x}_1^{(0)}(\mathbf{k}) \qquad (30)$$

And that μ_1, μ_2 , μ_3 , μ_4 are known constants, and that $\widehat{X}_1^{(1)}(k)$ It is the initial value of the model OGM(1,N), which is also a known value. Therefore, for a given value of k, it is possible to calculate $\widehat{X}_1^{(1)}(k)$ according to equation (2-27). It is noted from the same equation that the term "μ_2" which is calculated from the parameter a is an exponential formula, so any slight change in (a) leads to a large change in the simulation and prediction values of the developed model OGM(1,N), so the term μ_2 is considered The most sensitive parameter on the results among the four parameters. (Zeng, Luo, Liu, Bai, and Li, 2016)

For the developed model OGM (1, N), the respond expression for the inverse of accumulation is given as in the eq. (15):

$$\widehat{X}_1^{(1)}(k-1) - \widehat{X}_1^{(0)}(k) = \widehat{X}_1^{(1)}(k)$$

In the developed model, the sum $\sum_{i=2}^{M} b_i x_1^{(1)}(k)$ can be treated as a grey constant with a very small range of variation, and at the same time adding the linear correction term $h_1(k-1)$ and greyaction quantity h_2 in the developed model OGM(1,N) led to solving the defects of the "modeling mechanism, "parameter estimation," and "model structure" of the traditional model GM(1,N). By using the above three measurements, the possibility of converting between the model was proven OGM(1,N) and the traditional grey prediction model with one variable corresponding to GM(1,1) by making N=1, which is not possible with the traditional gray model GM(1,N) due to a defect in its structural structure. This confirms the soundness of the structure of the new developed model and that the analysis of defects was correct.

2.3.4 The Optimized Grey Model with the optimal background value OBGM (1,M):

The generation coefficient (α) of the background value is one of the factors affecting the prediction accuracy of the multi-grey model, it is usually equal to (0.5) so that the background value represents the average of the previous and subsequent values in the $X^{(1)}$ series resulting from the cumulative generation process of the first order(1-AGO). In another attempt to improve the prediction accuracy of the GM(1,M) model, we try to choose the optimal value for the coefficient (α) which gives us a minimum Mean Absolute percentage error (MAPE), we can determine this value of (α) using the programe MATLAB to select the optimal value in the open interval (0 ,1). The new optimized Grey model is named OBGM(1,M). (Wang, and Hao, 2016) (Li, and Xing, 2023).

The OBGM (1,M) model was defined through the following equation:

$$X_1^{(0)}(k) + a\,\alpha\, x_1^{(1)}(k) + a\,(1-\alpha)\, x_1^{(1)}(k-1) = \sum_{i=2}^{M} b_i\, x_i^{(1)}(k) + (kc) + d \qquad (31)$$

$$(0 < \alpha < 1), \qquad k = 2,3,\ldots n$$

The vector of parameters $\hat{p} = [b_2, b_3, \ldots, b_n, a, c, d]^T$ is calculated using the least squares method. If a certain value of α reduces the average absolute relative error (MAPE) of the series $X_1^{(0)}$, then the coefficient will represent the optimal posterior basis value for the model OGM (1, M). This model is called OBGM (1, M). The estimation of the vector of least squares parameters $\hat{p} = [b2, b3, \ldots, bn, a, c, d]^T$ is verified.

$$\hat{p} = (E^T E)^{-1} E^T H \qquad (32)$$

The matrix E and vector S can be represented as follows ((Li, and Xing, 2023)

$$E = \begin{pmatrix} x_2^{(1)}(2) & x_3^{(1)}(2) & \ldots & x_M^{(1)}(2) & -\alpha x_1^{(1)}(2) - (1-\alpha)x_1^{(1)}(1) & 1 & 1 \\ x_2^{(1)}(3) & x_3^{(1)}(3) & \ldots & x_M^{(1)}(3) & -\alpha x_1^{(1)}(3) - (1-\alpha)x_1^{(1)}(2) & 2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_2^{(1)}(n) & x_3^{(1)}(n) & \ldots & x_M^{(1)}(n) & -\alpha x_1^{(1)}(n-1) - (1-\alpha)x_1^{(1)}(n) & n-1 & 1 \end{pmatrix} \quad H = \begin{pmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{pmatrix} \qquad (33)$$

Then, the differential equation can be rewritten as the following difference equation:

$$\hat{x}_1^{(0)}(k) + a_1\,\alpha\, \hat{X}_1^{(1)}(k) + (1-\alpha)\, \hat{X}_1^{(1)}(k-1) = \sum_{i=2}^{M} b_i\, x_i^{(1)}(k) + kc + d, \quad k=1,2,\ldots,n, \qquad (34)$$

Since,

$$\hat{X}_1^{(1)}(k-1) - \hat{X}_1^{(0)}(k) = \hat{X}_1^{(1)}(k)$$

Hence,

$$\hat{X}_1^{(1)}(k) - \hat{X}_1^{(1)}(k-1) + a\,\alpha\, \hat{X}_1^{(1)}(k) + (1-\alpha)\, \hat{X}_1^{(1)}(k-1) = \sum_{i=2}^{M} b_i\, x_i^{(1)}(k) + kc + d \qquad (35)$$

The response function for $\hat{X}_1^{(1)}(k)$ can be obtained through:

$$\hat{X}_1^{(1)}(k) = v_1 \sum_{i=2}^{M} b_i\, x_i^{(1)}(k) + v_2\, \hat{X}_1^{(1)}(k-1) + v_3\, k + v_4, \quad k=2,3,\ldots,n \qquad (36)$$

Whereas:

$$v_1 = \frac{1}{1+a\,\alpha}, \quad v_2 = 1 - \frac{1}{1+a\,\alpha}, \quad v_3 = \frac{c}{1+a\,\alpha}, \quad v_4 = \frac{d}{1+a\,\alpha} \qquad (37)$$

Then the the time response is given as equation (15) also as :

$$\hat{X}_1^{(0)}(k) = \hat{X}_1^{(1)}(k) - \hat{X}_1^{(1)}(k-1)$$

3. Discussion of Results

3.1 Data Description of Water Consumption

Water is one of the most important natural resources to which the issue of food security relates, and the survival of humanity depends on it. Given the importance of water and the importance of regulating its use, in the early nineties the Human Development Report issued by the United Nations for the year 1990 put forward the idea of water security, which means that a person should enjoy enough safe water at affordable prices. Reasonable to live a clean, healthy and productive life. Moreover, water issue is related to sustainable development.

Many visits have been done to the Ministry of Water Resources and the departments concerned with drinking water and domestic consumption, represented by the Baghdad Water Department of the Baghdad Municipality, for request of providing data-set related to drinking water consumption in the center of Baghdad, and the Baghdad Water Directorate in Karrada for the purpose of providing us with data related to drinking water consumption in the outskirts of Baghdad. So, real data concerns drinking water consumption were got for (2016-2022).

The data studied in this research are seven-year time series (n = 7), representing the quantities of drinking water consumption and domestic use in Baghdad Governorate as a dependent variable representing the characteristic series of the system measured in the unit of measurement (1000 m3 / year). The factors affecting water consumption are: the number of population (X_1) in Baghdad and the annual average of temperature (X_2). The population number in Baghdad Governorate was obtained from the official website of the Central Bureau of Statistics. While the average annual temperatures (maximum points) were obtained from the Iraqi Meteorological Department.

3.2 Numerical results of real application

The numerical results were obtained by using the programing language Matlab.2018a, as below.

Table (1): Data on water consumption quantities in Baghdad, its population, and the annual average of maximum temperatures for the period (2016-2022)

years	Water consumption quantity (1000 m^3) (X_0)	Population (X_1)	Annual average temperature (X2)
2016	1258528.08	8280013	31.9
2017	1563269.36	8361497	32.296
2018	1369313.14	8199317	31.55
2019	1399495.44	8387229	31.7
2020	1459894.04	8472762	32.316
2021	1568069.40	8565671	33.16
2022	1559386.89	8669783	32.33

Table (2): The estimated values of water consumption in Baghdad in (2016-2022) for GM ,OGM and OBGM models

Year	$X^{(0)}$	$\widehat{X^{(0)}}$GM	$\widehat{X^{(0)}}$OGM	$\widehat{X^{(0)}}$OBGM ($\alpha = 0.500127$)
2016	1258528.08	1258528.08	1258528.08	1258528.08
2017	1563269.36	1542863.70 8	1563115.858	1563272.819
2018	1369313.14	1472319.22	1371021.38	1371035.486

2019	1399495.44	1305827.574	1394693.328	1394698.65
2020	1459894.04	1394193.174	1465503.767	1465511.976
2021	1568069.40	1557353.375	1565256.572	1565268.348
2022	1559386.89	1627248.812	1559786.513	1559788.502

It can be noticed that the results of estimated water consumption in table (2) are very close to those real data in table (1).

Table (3): The estimated parameters of (GM)

Parameters (GM)	**a**	**b_1**	**b_2**
Estimation	-387759	1.017828	-2.68839

Table (4): The estimated parameters of (OGM)

Parameters (OGM)	**a**	**h_1**	**h_2**	**b_1**	**b_2**
Estimation	109750.2	0.478342	1.299324	5121412	5670591

Table (5): The estimated parameters of (OBGM)

Parameters (OBGM)	**a**	**h_1**	**h_2**	**b_1**	**b_2**
Estimation	109768.3	0.478421	1.299538	5122258	5671524

The tables (3), (4), and (5) show the estimated parameters for (GM), (OGM), and (OBGM), respectively. The last two models have used the same mechanism with closely values in their parameters.

Table (6): MAPE values of (GM , OGM, OBGM) models

Model	**GM**	**OGM**	**OBGM ($\alpha = 0.500127$)**
MAPE	0.035795	0.001524	0.001511

Table (6) denoted the values of comparison criterion (MAPE) for GM, OGM, and OBGM models, respectively.

Below, Figure (1) represents a visual comparison for GM and OGM models.

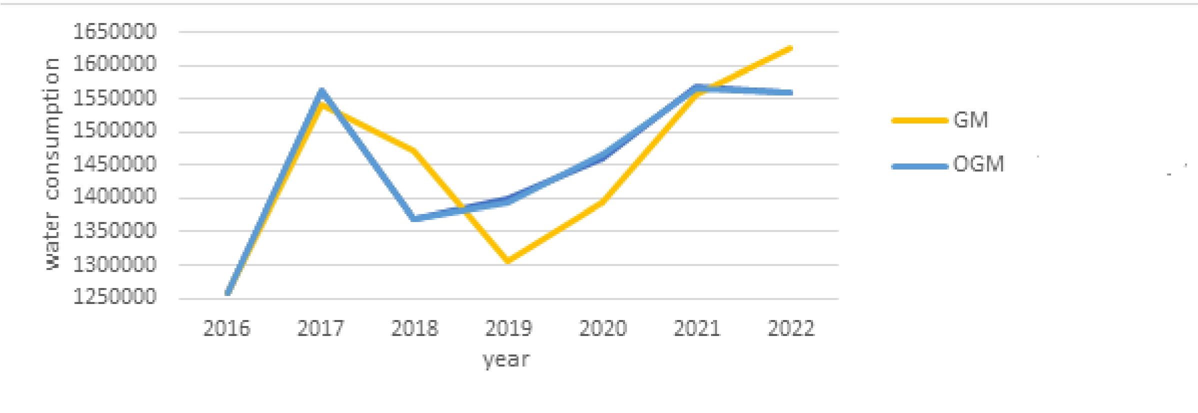

Fig (1): The real and estimated values of GM and OGM(1,3) for water consumption in Iraq

Furthermore, Figure (2) represents a visual comparison for GM and OBGM models.

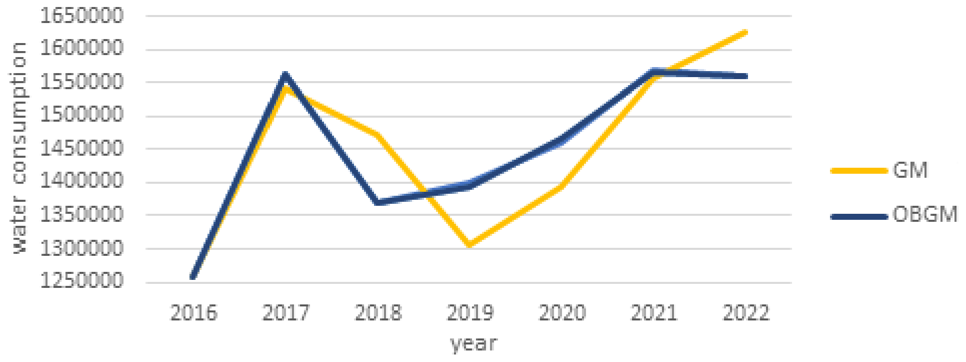

Fig (2): The real and estimated values of GM and OBGM(1,3) for water consumption in Iraq

4. Conclusion

This paper offered the characteristics of the traditional multivariate grey model GM(1,M) and showed the defects which it suffer from and affected on the simulation and prediction accuracy. To compensate this lack in GM (1,M),the linear correction term (h_1) and grey action term (h_2) were added beside the traditional model to improve its efficiency and the optimized grey model OGM(1,M) was introduced. Another optimization has been done on the optimized model OGM(1,M) through select the optimal generation coefficient (α) of the background value ($Z^{(1)}$) and OBGM(1,M) was introduced also. The two optimization on GM(1,M) was studied by a practical application using a realist data represent the water consumption in IRAQ on the period (2016-2022). The results showed that the prediction errors of the multivariate grey model decreases after each of the two optimizations, the MAPE of the OGM(1,3) model decreased from 0.035795 in GM(1,3) to (0.001524) in OGM(1,3) and to (0.001511) in OBGM(1,3).That reflects

the effectiveness of model optimization, proves the practicability of the OGM(1,3) and OBGM(1,3) models, the defect analysis is correct and the validity of the structure reform for the trariditional multivariate grey model GM (1,M). On other hand the idea of improving the GM(1,M) model by optimizing the background value through using the optimal generation coefficient (α) has less contribution in rising the accuracy of the model due to nearness of the generation coefficient's value ($\alpha = 0.500127$) from (0,5) while this type of optimization achieves more contribution in other studies and applications (Rao, K.V. et al., 2021) & (Wang, Y. et al. , 1998).

5. Future Works:

The current work can be expanded as hybridization with artificial intelligence algorithms such

as swarms, or artificial neural networks to grantee the precision of optimization.

References

1. Cheng, M., Li, J., Liu, Y. and Xiang, M., 2022. Parameter estimation of modified gray model GM (1, N) and model application. Communications in Statistics-Simulation and Computation, 51(9), pp.5447-5464.

2. Deng, J.L., 1982. The Control problem of grey systems. System & Control Letter.
3. Geng, J., Ye, D. and Luo, P., 2015. Predicting Severity of Software Vulnerability Based on Grey System Theory. In Algorithms and Architectures for Parallel Processing: ICA3PP International Workshops and Symposiums, Zhangjiajie, China, November 18-20, 2015, Proceedings 15 (pp. 143-152). Springer International Publishing.
4. He, F. and Tao, T., 2014. An Improved Coupling Model of Grey-System and Multivariate Linear Regression for Water Consumption Forecasting. Polish Journal of Environmental Studies, 23(4).
5. Julong, D., 1989. Introduction to grey system theory. The Journal of grey system, 1(1), pp.1-24.
6. Kayacan, E., Ulutas, B. and Kaynak, O., 2010. Grey system theory-based models in time series prediction. Expert systems with applications, 37(2), pp.1784-1789.
7. Li, C., Li, Y. and Xing, J., 2023. Multivariate Grey Prediction Model Application in Civil Aviation Carbon Emission Based on Fractional Order Accumulation and Background Value Optimization. Sustainability, 15(11), p.9127.
8. Liu, S. and Forrest, J.Y.L., 2010. Grey systems: theory and applications. Springer Science & Business Media.
9. Liu, S., Forrest, J. and Yang, Y., 2012. A brief introduction to grey systems theory. Grey Systems: Theory and Application, 2(2), pp.89-104.
10. Mao-lin, L., Yong, W. and Yan-qiu, D., 2007. A new method to estimate parameter of grey model. In 2007 IEEE International Conference on Grey Systems and Intelligent Services.
11. Mohammed Ali, N. S. and Mohammed, F. A., 2018. Comparison Some Estimation Methods Of GM(1,1) Model With Missing data and Practical Application. Journal of Economic and administration science, Vo (24), No (103).
12. Rao, K.V., Kumar, Y.P., Singh, V.K., Raju, L.S. and Ranganayakulu, J., 2021. Vibration-based tool condition monitoring in milling of Ti-6Al-4V using an optimization model of GM (1, N) and SVM. The International Journal of Advanced Manufacturing Technology, 115(5-6), pp.1931-1941.
13. Wang, Y., Zhang, G., Moon, K.S. and Sutherland, J.W., 1998. Compensation for the thermal error of a multi-axis machining center. Journal of materials processing technology, 75(1-3), pp.45-53.
14. Wang, Z.X. and Hao, P., 2016. An improved grey multivariable model for predicting industrial energy consumption in China. Applied Mathematical Modelling, 40(11-12), pp.5745-5758.
15. Wei, B. and Xie, N., 2020. On unified framework for discrete-time grey models: Extensions and applications. ISA transactions, 107, pp.1-11.
16. Xie, M., Yan, S., Wu, L., Liu, L., Bai, Y., Liu, L. and Tong, Y., 2021. A novel robust reweighted multivariate grey model for forecasting the greenhouse gas emissions. Journal of Cleaner Production, 292, p.126001.
17. Zeng, B., Luo, C., Liu, S., Bai, Y. and Li, C., 2016. Development of an optimization method for the GM (1, N) model. Engineering Applications of Artificial Intelligence, 55, pp.353-362.
18. Zeng, B., Ma, X. and Shi, J., 2020. Modeling method of greyGM (1, 1) model with interval grey action quantity and its application. Complexity, 2020, pp.1-10.
19. Zhang, Y. and Zhang, J., 2010. A hybrid model of neural network and grey theory for air traffic passenger volume forecasting. Key Engineering Materials, 439, pp.818-822.

مقارنة نماذج الانحدار الرمادي المتعدد التقليدية والمحسنّة مع تطبيق على المياه

محمد اسماعيل رشيد
جامعة بغداد، كلية الادارة والاقتصاد،
قسم الاحصاء.
Mohammed.Ismail1201a@coadec.uobaghdad.edu.iq

أ.د. عمر عبدالمحسن علي
جامعة بغداد، كلية الادارة والاقتصاد،
قسم الاحصاء.
dromar72@coadec.uobaghdad.edu.iq

المستخلص

تعتبر نظرية النظام الرمادي بمثابة منهج علمي متعدد التخصصات، يتناول الأنظمة التي تحتوي على معلومات جزئية (مثل: العينات الصغيرة). وتعتبر النمذجة الرمادية جزءًا حيويًا من هذه النظرية التي تعطي نتائج كافية على الرغم من محدودية كمية البيانات. يمكن تقسيم النماذج الرمادية إلى نوعين؛ نماذج رمادية أحادية المتغير ومتعددة المتغيرات. النموذج الرمادي أحادي المتغير GM(1,1) هو حجر الزاوية في هذه النظرية، لكنه لا يأخذ العوامل النسبية في الاعتبار. بينما تأخذ النماذج الرمادية متعددة المتغيرات GM(1,M) هذه العوامل في الاعتبار. إلا أنها ذات بنية معقدة وبعض العيوب في "آلية النمذجة" و"تقدير المعلمة" و"بنية النموذج". ولذلك خضع GM(1,M) في النسخة التقليدية للعديد من تجارب التحسينات للتخلص من هذه العيوب.

ويهدف هذا البحث الى :

- التعريف بخصائص GM(1,M) التقليدي بالإضافة إلى مشكلته.
- فرض طريقة للتغلب على هذه المشكلة من خلال نموذجين رماديين متعددي المتغيرات لمعادلة مشتقة من رتبة واحدة.
- تقديم نموذج رمادي محسن ومختصر بـ OGM(1, M) عن طريق إضافة مصطلح التصحيح الخطي h1(M-1) ومصطلح كمية الفعل الرمادي (h2) إلى النموذج التقليدي GM(1,M).
- استخدم قيمة الخلفية المحسنة للنموذج الرمادي OBGM(1,M) عن طريق تحسين قيمة الخلفية للنموذج الأخير OGM(1,M).
- تطبيق مجموعة بيانات حقيقية تمثل استهلاك المياه في مدينة بغداد للمدة (2016-2022) لمقارنة كلا النموذجين المحسنين مع النموذج التقليدي.
- استخدام متوسط النسبة المئوية للخطأ المطلق (MAPE) ومعامل التحديد R^2.

أظهرت نتائج المقارنة تفوق كلا من OGM(1, M) و OBGM(1,M) مقارنة بالنموذج التقليدي GM(I,M).

نوع البحث: ورقة بحثية .

الكلمات الرئيسة: المربعات الصغرى، الانموذج الرمادي، انموذج الانحدار المتعدد، القيمة الخلفية، معيار متوسط النسبة المئوية للخطأ المطلق.