Control and Optimization in Applied Mathematics (COAM)
Vol. 2, No. 2, Autumn-Winter 2017(15-31), ©2016 Payame Noor University, Iran

Application of Grey System Theory in Rainfall Estimation

D. Darvishi Salookolaei[1,*], S. Liu[2], P. Babaei[3]

[1,3] Department of Mathematics, Payame Noor University, P.O. Box 19395-3697, Tehran, Iran,
[2] College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, China.

Received: Oct 15, 2018; **Accepted:** Mar 23, 2019.

Abstract. Considering the fact that Iran is situated in an arid and semi-arid region, rainfall prediction for the management of water resources is very important and necessary. Researchers have proposed various prediction methods that have been utilized in such areas as water and meteorology, especially water resources management. The present study aimed at predicting rainfall amounts using Grey Prediction Methods. It is a novel approach in confrontation with uncertainties in the aquiferous region of Babolrud to serve for the water resources management purposes. Therefore, expressing the concepts of Grey Prediction Method using the collected data, at a 12-year timeframe of 2006 to 2017, rainfall prediction in 2018 to 2022 were also implemented with three methods GM(1,1), DGM(2,1) and Verhulest models. According to the calculated error and the predictive power, GM(1,1) method is better than other models and was placed within the set of good predictions. Also, we predict that in 2027, there might be a drought. According to the small samples and calculations required in this approach, the method is suggested for rainfall prediction in inexact environments. The authors can use fuzzy grey systems to predict the amount of rainfall in uncertain environments.

Keywords. Prediction, Grey system, Water resources management, Rainfall amount, Absolute prediction error.

MSC. 90C34; 90C40.

* Corresponding author
d_darvishi@pnu.ac.ir, p.babaeivalohi@gmail.com, sfliu@nuaa.edu.cn
http://mathco.journals.pnu.ac.ir

1 Introduction

Rainfall is one of the most complicated and most accidental natural phenomena which serves important and essential functions in hydrology cycle and hydrological models‘ inputs. Estimation and prediction of rainfall amount play an essential and effective role in the management of dams‘ reservoir management, flood control and drought. It also exerts a large deal of effect on the water resource management. Thus, the precision and reliability of the data pertaining to that in temporal and spatial scales are abundantly important for climate change simulations and flood prediction. Due to the same reason, this issue has been considered by hydrologists [18]-[22]. The characteristics of rainfall amount are important for the precise modeling of the other hydrological streams like surface run-off, evaporation and transpiration.

The reason is that rainfall controls their statistical behavior as well as their intermediate and intensive distribution. The precise rainfall amount prediction during time is an important part of hydrological streams‘ evaluation even more substantially in arid and semi-arid basins wherein the runoff and discharge measurements might be impractical [6]. The prediction of any incident constitutes the basis for the crisis management, which is made feasible when one can have an available appropriate prediction model. It is necessary to perform studies in this area. Doing so, water resources management and accurate prediction can be accurately carried out for both the agriculturalists, to be able to make proper sowing plans, (especially in regard of rain-fed cultivation) and the aqueous resources and reservoirs planners to perform more successfully and eventually bar any incurrence of irreparable losses. Decision makers of water resources need confident predictions for their management decision-making. Meanwhile, various prediction methods have attempted the determination of the relationship between independent and dependent variables and many concepts and statistical methods have been utilized for the prediction and premonition of the climatic variations. In their adoption of various methods, the authors have sought to obtain acceptable results in this field for various spots around the globe. During the recent decades, extensive researches have been conducted regarding the prediction of rainfall amount (monthly and seasonal) using macro scale climatic signals and local variables (temperature, humidity, pressure and evaporation and so forth) [3].

Time series prediction refers to the process through which the future values of a system is forecasted based on the information obtained from the past and current data points [11]. In literature, statistical and artificial intelligence based approaches are the two main techniques used for time series predictions [8]. There are several prediction methods. These forecasting methods could be generally divided into three categories of statistical analysis models, computational intelligence models and grey prediction models. Up to now, there have been various methods for predicting the atmospheric parameters, including rainfall amounts. Artificial neural network, autoregressive model, moving average calculation, support vector machine, time-series models, wavelets theory and tree regression are amongst the methods that have been applied up to now in the area of rainfall prediction [7]-[20]. Time-based prediction methods rely primarily on past and objective information as well as simple procedures for prediction calculations. This method is very suitable for short-term forecasts. Many researches have been considered the use of time series in hydrological parameters‘ modeling like rainfall, temperature and river streams. In the investigation of the prior research, the scientists have endeavored to model and then simulate the climatic parameters to enable an analysis of their changes. Modeling in Auto Regressive Moving Average (ARMA), Auto Regressive Integrated Moving (ARIMA) and Seasonal Auto Regressive Integrated Moving (SARIMA) are important and credible methods for the stimulation of the climatic parameters [2]. The objective of time-series is to determine rule-based nature of the data and identification of their behaviors in line with making future predictions. There is a scarcity of instruments applicable to the rainfall prediction in the short run [19]-[25]. In the majority of the cases, the prediction of rainfall amount considers the current climatic conditions using numerical weather prediction models and/or extraction of data from remote sensing observations (radar data and satellite images). Although numerous studies have been investigated in the area of climatic parameters‘ predictions, these models have certain advantages as well as many shortcomings and disadvantages.

The larger the volume of the available information, the more precise the predictions would be. However, the collection of information necessitates the spending of time and costs. Though it causes an increase in the precision of the predictions, spending of more time is accompanied by reductions in the relevance and timeliness of the information. On the other hand, not all the required information is always available. Thus, it is important to use techniques that can offer precise predictions with the least available data. However, not only the statistical models are not as accurate as the neural network-based approaches for nonlinear problems, but also they are too complex to be used in predicting future values of a time series [8].The results of these methods are always ensured just under a basic assumption of known distribution or large sample. However, it is sometimes impossible to collect time series with large sample sizes [24]. In this paper, the use of grey prediction theory was proposed to alleviate the problem. This method faced difficulties of having incomplete or insufficient information [8] in many situations. When encountering uncertain data, the use of fuzzy gray systems was needed.

Grey Prediction Method is amongst the predictions that is simple and understandable with the least information available. This method can be administered in a short period and offer an appropriate precision level. Grey prediction model is one of the most important components of grey system theory and it is useful in solving indefinite problems featuring small sample volumes and many weaknesses. During the past three decades, grey model has been extensively utilized in such fields as agriculture, industry, society, economy, transportation, geology, hydrology and meteorology, environment, education, energy and healthcare [10]-[13]. Grey model has been largely considered during the recent years.

Grey prediction model is an authentic method which is widely applied within the societies [1]-[21]-[25]-[27]-[28]. GM(1,1) type of grey model is the most widely used model in the literature, pronounced as ‘‘Grey Model First Order One Variable”. This model is a time series forecasting model [8]. As a superiority to conventional statistical models, grey models require only a limited amount of data to estimate the behavior of unknown systems [5].

The main objective of this study was around rainfall amount prediction using grey prediction model of GM(1,1) in the aquiferous region of Babolrud and the rainfall prediction in the upcoming years parallel to water resources management. The remainder of this paper is organized as follows:
The grey systems theory is briefly presented in Section 2. The basic knowledge of the GM(1,1),

DGM(2,1) and Verhulest models, grey disaster prediction and some statistical measurements, including relative percentage error, root mean squared error and mean of average percentage error are demonstrated in Section 3. Rainfall prediction based on collected data using grey prediction methods GM (1,1), DGM(2,1) and Verhulest models and forecast accuracy of these model have been carried out in section 4. Finally, conclusions are presented in section 5.

2 Grey System Theory

Grey model theory is a mathematical description of indetermination. This theory can be used alone or in connection with other mathematical theories dealing with uncertainty, such as fuzzy theory. The theory of grey systems was proposed in 1982 by Deng and it was later employed by Hang for solving uncertainty problems [4]. During late 1960s, Hang performed many studies on the prediction and control of economic and fuzzy systems when facing systems with many uncertainties. The indices of these systems could be roughly described by fuzzy mathematics and/or statistics and probabilities. To solve these systems optimally, Deng published an article under the title of "The Control Problems of Grey Systems" in 1982 to introduce grey systems theory. The major advantage of the grey systems theory is its need for small samples and poor information. In fact, grey systems theory has been posited as an effective method for solving the problems with discrete data and imperfect information [26]. The term "grey" in the grey system theory is a combination of black and white wherein black indicates indefinite information and white indicates perfect information. Grey points to incomplete information; in other words, the information that is somewhat clear and somewhat uncertain. These systems are called grey systems. To get more familiar with the grey systems theory, the interested reader can refer to the studies in [12]-[16]. The grey system theory include the following fields: grey generating, grey relational analysis, grey forecasting, grey decision making, and grey control. The majority of the prediction methods need a large number of data and statistical methods that are utilized to investigate the system properties. Moreover, systematic investigation is very difficult for the exogenous confusion in the system and the complex mutual interrelationships between the system and the peripheral environment. Grey prediction model, as the core of the grey system theory, features the advantage of creating a model using few and uncertain data and it is an appropriate instrument for the prediction of systems with complex, uncertain and disorganized structure. The grey prediction model is a lot more applicable and simpler than the other prediction methods. In these studies and the others, it is seen that grey system theory-based approaches can achieve good performance characteristics when applied to real-time systems. The reason behind is that grey predictors adapt their parameters to new conditions as new outputs become available. Therefore, grey predictors are more robust with respect to noise and lack of modeling information when compared to conventional methods [8].

3 Grey Prediction Model

Grey prediction model investigates the preliminary data assisted by grey differential equation to extract the rules governing the system. The model creates a dynamic and continuous differential equation from the series of discrete data to actualize time-series prediction. Each grey model is expressed in the form of GM(n,m) wherein n is the order of the differential equation and m determines the number of variables.

3.1 GM(1,1) Model

Grey model analysis permits the prediction of system behavior, such as extrapolation of data.The grey forecasting model GM(1,1) is a time series prediction model encompassing a group of differential equations adapted for parameter variance as well as a first-order differential equation. In this section we focus on the grey prediction model, GM(1,1), which has been applied in many aspects of social and natural science, including decision-making, finance, economics, engineering and meteorology. GM(1,1) is the most applied models of time-series prediction model and it is basically an exponential model [17]. Liu and Deng studied the range suitable for GM(1,1) model based on a simulated test. The area of validity, the area to be used carefully, the area not suitable for use and the prohibited area of GM(1,1) have been divided clearly according to the threshold of the developing coefficients [16]. In order to smooth the randomness, the primitive data obtained from the system to form the GM(1,1) is subjected to an operator, named Accumulating Generation Operator (AGO) [5]. The operator reveals the internal order pattern of the data or the trends of the data series. Then, the differential equation operationalizes system prediction in n stages. Finally, the prediction values and Inverse Accumulated Generating Operator (IAGO) are applied to figure out the main data estimates [5]-[23]-[26]. The procedure of GM(1,1) grey prediction model can be summarized as follows:

Step 1: Let $x^{(0)} = \big(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\big)$ denote a non-negative sequence of original data, where n is the length of the raw data sequence and $n \geq 4$.

Step 2: The new cumulative data sequence $x^{(1)} = \big(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\big)$, which is accumulated generating operator [4] of $x^{(0)}$, is obtained as $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, $k = 1, 2, 3, \cdots, n$. The data sequence $X^{(1)}$ could weaken the randomness of $x^{(0)}$. It is obvious that it is monotonically increasing.

Step 3: The generated mean sequence of $x^{(1)}$ is defined as:

$$Z^{(1)} = \Big(z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)\Big) \tag{1}$$

where $Z^{(1)}(k)$ is the mean value of adjacent data, i.e.

$$Z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), \quad k = 2, 3, \cdots, n. \tag{2}$$

The least square estimate sequence of the grey difference equation of GM(1,1) is defined as follows:

$$Z^{(1)}(k) = \alpha x^{(1)}(k) + (1-\alpha)x^{(1)}(k-1), \quad k = 2, 3, \cdots, n \tag{3}$$

$$x^{(0)}(k) + aZ^{(1)}(k) = b. \tag{4}$$

Where a is the development coefficient and b is the input grey coefficient or grey parameter.

Step 4: Define the first-order differential equation of sequence $x^{(1)}$ as:

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b. \tag{5}$$

In above mentioned equitation, t denotes the independent variables, a represents the grey developed coefficient of GM(1,1) model, and b is the grey controlled variable of the GM(1,1) model.

Step 5: Utilize the least squares estimation, we can derive the estimated first-order AGO sequence $x_p^{(1)}(k+1)$ and the estimated inversed AGO sequence $x_p^{(0)}(k+1)$ as follows,

$$x_p^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-ak} + \frac{b}{a} \tag{6}$$

$$x_p^{(0)}(k+1) = x_p^{(1)}(k+1) - x_p^{(1)}(k) \tag{7}$$

where $k = 1, 2, 3, \cdots, n$. parameters a and b can be conducted by the least square estimation methods as follows:

$$[a, b]^T = \left[B^T, B\right]^{-1} B^T Y \tag{8}$$

where $Y = \left[x^{(0)}(2), \cdots, x^{(0)}(n)\right]^T$

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(n) & 1 \end{bmatrix}. \tag{9}$$

To obtain the predicted value of the primitive data at time $(k+1)$, IAGO is used to establish the following grey model:

$$\begin{aligned} x_p^{(0)}(k+1) &= x_p^{(1)}(k+1) - x_p^{(1)}(k) \\ &= (x^{(0)}(1) - \frac{b}{a})(1 - e^a)e^{-ak}, \quad k = 1, 2, \cdots, n. \end{aligned} \tag{10}$$

The predicted value of the primitive data at time $(k+h)$ can be obtained as follows:

$$x_p^{(0)}(k+h) = (x^{(0)}(1) - \frac{b}{a})(1 - e^a)e^{-a(k+h-1)}, \quad k = 1, 2, \cdots, n. \tag{11}$$

In large data areas, the grey system prediction method based on small data mining as a new force suddenly rises, which becomes an effective tool for valuable information extraction from a mass of data. It is a very meaningful job to build more normal model testing standards based on the grey system prediction model testing method and statistical testing theory [16].

3.2 Grey Verhulst Model

As for non-monotonic wavelike development sequences, or saturated sigmoid sequences, one can consider establishing a grey Verhulst model. The main purpose of Velhulst model is to limit the whole development for a real system and it is effective in describing some increasing processes. The grey Verhulst model can be defined as [26]:

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b(x^{(1)})^2. \tag{12}$$

Grey difference equation of Eq.(12) is

$$x^{(0)}(k) + aZ^{(1)}(k) = b(Z^{(1)}(k))^2. \tag{13}$$

Similar to the GM(1,1) model

$$[a,b]^T = \left[B^T, B\right]^{-1} B^T Y \tag{14}$$

where $Y = \left[x^{(0)}(2), \cdots, x^{(0)}(n)\right]^T$

$$B = \begin{bmatrix} -Z^{(1)}(2) & (Z^{(1)}(2))^2 \\ -Z^{(1)}(3) & (Z^{(1)}(3))^2 \\ \vdots & \vdots \\ -Z^{(1)}(n) & (Z^{(1)}(n))^2 \end{bmatrix}. \tag{15}$$

The solution of $x^{(1)}(t)$ at time k:

$$x_p^{(1)}(k+1) = \frac{ax^{(0)}(1)}{bx^{(0)}(1) + (a - bx^{(0)}(1))e^{ak}}. \tag{16}$$

Applying the IAGO, the solution of $x^{(0)}(t)$ at time k:

$$x_p^{(0)}(k) = \frac{ax^{(0)}(1)(a - bx^{(0)}(1))}{(bx^{(0)}(1) + (a - bx^{(0)}(1))e^{ak-1}} \times \frac{(1-e^a)e^{a(k-2)}}{(bx^{(0)}(1) + (a - bx^{(0)}(1))e^{ak-2}}. \tag{17}$$

3.3 DGM(2,1) Model

The DGM(2,1) model [10] is a single sequence second-order linear dynamic model and is fitted by differential equations. GM(2,1) is proposed to change the linear structure of GM(1,1) model and expands application scope of grey prediction theory. It is an important model among the group of grey prediction models. The structure of GM(2,1) model and solving process have limited its application. Assume an original series to be $x^{(0)}$

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right) \tag{18}$$

a new sequence $x^{(1)}$ is generated by the accumulated generating operation (AGO).

$$x^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right) \tag{19}$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \quad k = 1, 2, 3, \cdots, n. \tag{20}$$

Setting up a second-order differential equation:

$$\frac{d^2x^{(1)}(1)}{dt^2} + a\frac{dx^{(1)}(t)}{dt} = b, \tag{21}$$

where a is the developing coefficient and b is the grey input coefficient. The coefficients a, b can be calculated using the least-squares method as shown below:

$$[a,b]^T = \left[B^T, B\right]^{-1} B^T Y \tag{22}$$

$$Y = \begin{bmatrix} (x^{(0)}(2) - x^{(0)}(1)) \\ (x^{(0)}(3) - x^{(0)}(2)) \\ \vdots \\ (x^{(0)}(n) - x^{(0)}(n-1)) \end{bmatrix} \tag{23}$$

$$B = \begin{bmatrix} -x^{(0)}(2) & 1 \\ -x^{(0)}(3) & 1 \\ \vdots & \vdots \\ -x^{(0)}(n) & 1 \end{bmatrix}. \tag{24}$$

According to Eq.(21) , we have:

$$x_p^{(1)}(k+1) = \left[\frac{b}{a^2} - \frac{x^{(0)}(1)}{a}\right] e^{-ak} + \frac{b}{a}(k+1) + (x^{(0)}(1) - \frac{b}{a})(\frac{1+a}{a}). \tag{25}$$

The prediction values of original sequence can be obtained by applying inverse AGO to $x^{(1)}$ namely,

$$\begin{aligned} x_p^{(0)}(k+1) &= x_p^{(1)}(k+1) - x_p^{(1)}(k) \\ &= \left[\frac{b}{a^2} - \frac{x^{(0)}(1)}{a}\right](1-e^a)e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \cdots, n-1. \end{aligned} \tag{26}$$

3.4 Grey Disaster Prediction

The basic idea of grey disaster predictions is essentially the prediction of abnormal values. For Those values that are considered abnormal, it is commonly determined based on individuals‘ experiences and qualitative analysis [14]. The task for disaster predictions is to pinpoint the time moment(s) for one or several abnormal values to occur so that relevant parties can have enough time to make preparations for disasters to happen [12].

Definition 1. [14] Assume that $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}$ is a sequence of raw data. Then

(1) for a given upper abnormal (or catastrophe) value ζ, the sub-sequence of X

$$X_\zeta = \{x[q(1)], x[q(2)], \cdots, x[q(m)]\} = \{x[q(i)] | x[q(i)] \geq \zeta;\ i = 1, 2, \cdots, m\} \tag{27}$$

is known as the upper catastrophe sequence.

(2) For a given lower abnormal (or catastrophe) value ζ, the sub-sequence

$$X_\zeta = \{x[q(1)], x[q(2)], \cdots, x[q(l)]\} = \{x[q(i)] | x[q(i)] \leq \zeta; i = 1, 2, \cdots, l\}. \tag{28}$$

Definition 2. [14] Assume that $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}$ is a sequence of raw data. The following sub-sequence of X

$$X_\zeta = \{x[q(1)], x[q(2)], \cdots, x[q(m)]\} \subset X \tag{29}$$

is a catastrophe sequence. Then,

$$Q^{(0)} = \{q(1), q(2), \cdots, q(m)\} \tag{30}$$

will be referred to as the catastrophe date sequence [14].
Disaster prediction is about finding any patterns, through the study of catastrophe date sequences to predict future dates of occurrences of catastrophes. In grey system theory, each disaster prediction was realized through establishing GM(1,1) models for relevant catastrophe date sequences.

Definition 3. [14] If $Q^{(0)} = \{q(1), q(1), \cdots, q(m)\}$ is a catastrophe date sequence, the following

$$Q^{(1)} = \{q(1)^{(1)}, q(1)^{(1)}, \cdots, q(m)^{(1)}\} \tag{31}$$

is the 1-AGO sequence of the catastrophe date sequence$Q^{(0)}$, $Z^{(1)}$ is the adjacent neighbor mean generated sequence of $Q^{(1)}$, and

$$q(k) + az^{(1)}(k) = b \tag{32}$$

is referred to as a catastrophe model of GM(1,1). For the available sequence, $X = \{x(1), x(2), \cdots, x(n)\}$ of raw data, if n stands for the present and the last entry $q(m)(\leq n)$ in the corresponding catastrophe, and date sequence $Q^{(0)}$ represents when the last catastrophe occurred, then the predicted value $\hat{q}(m+1)$ represents the next forthcoming catastrophe and for any $k > 0$, $\hat{q}(m+k)$ stands for the predicted date for the k^{th} catastrophe to occur in the future.

3.5 Model Evaluation Scales

It should be noted that most predictions do not exactly match reality, and should try to minimize the forecast error. There are various techniques for forecasting, each of which has its own application. To compare the model precision, there are three common tools such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). As a judgment method, prediction error determination indicates the success of a prediction model. This study adopted three criteria to evaluate the performance of the grey rainfall-forecasting model. RMSE is a part of a standard for evaluating the prediction precision. Standard deviation designates an example of the differences between the real and estimated values. The relative root mean square error is defined as:

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}(x^{(0)}(k) - \hat{x}^{(0)}(k))^2}{n}} \tag{33}$$

where $x^{(0)}(k)$ denotes the observed cumulative rainfall at time t, $\hat{x}^{(0)}(k)$ is the forecast cumulative rainfall at time t. The RMSE represent a quantitative judgment of model performance.

MAE measures the difference between the real and the estimated values and it is expressed as below:

$$MAE = \frac{1}{n}\sum_{k=1}^{n} |x^{(0)}(k) - \hat{x}^{(0)}(k)|. \tag{34}$$

MAPE indicates the mean value of the prediction error ratio and it is explained as below:

$$MAPE = \frac{1}{n}\sum_{k=1}^{n} \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100. \tag{35}$$

To compare the prediction power of MAPE, there are four considered regions. If these values are below 10%, the model prediction power could be envisaged excellent. Values between 10 and 20 percent are suggestive of good prediction and values in a range from 20% to 50% are indicative of an acceptable prediction. Values above 50% are expressive of imprecise prediction [9].

4 Materials and Methods

Babolsar station is a rainfall measurement station in Mazandaran province and it is located at the side of Babolrud River on the southbound of Caspian Sea in Iran. The annual mean temperature was $18.4C$ and annual precipitation was 791 mm. It is said that forecasting rainfall for various purposes such as flood estimation, drought, catchment management, agriculture, and others is very important. There are several methods for forecasting such as ARMA, ARIMA, SARIMA, Decomposition, Holt-Winters and so on. In the current research paper, the annual rainfall time series data obtained from Babonlsar station were utilized to evaluate the prediction accuracy of GM(1,1) grey prediction model. In evaluating the prediction accuracy of GM(1,1), DGM(2,1) and Verhulest grey prediction models, the annual rainfall statistics for the periods of 2006 to 2017 were applied. The evaluation of these models authenticity and the predication results and error values can be found in the following table:

$$\begin{aligned} X^{(0)} &= \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(12)\} \\ &= \{956.9, 812.7, 1032.9, 1234.8, 619.2, 1081.2, 1047.1, 713.3, 725.6, \\ &\quad 944.3, 908.5, 626.1\} \end{aligned}$$

and

$$\begin{aligned} X^{(1)} &= \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(12)\} \\ &= \{956.9, 1769.6, 2802.5, 4037.3, 4656.5, 5737.7, 6784.8, 7498.1, 8223.7, \\ &\quad 9168, 10076.5, 10702.6\}. \end{aligned}$$

The mean sequence based on consecutive neighbors of $X^{(1)}$ is given by

$$\begin{aligned} Z^{(1)} &= \{z^{(1)}(2), z^{(1)}(3), \cdots, z^{(1)}(12)\} \\ &= \{1363.25, 2286.05, 3419.9, 3419.9, 4346.9, 5197.1, 6261.25, 7141.45, 7860.9, \\ &\quad 8695.85, 9622.1, 1038.25\}. \end{aligned}$$

Let $x^{(0)}(k) + az^{(1)}(k) = b$. From

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(12) & 1 \end{bmatrix}$$

$$Y = \left[x^{(0)}(2), \cdots, x^{(0)}(12)\right]^T$$
$$= [812.7, 1032.9, 1234.8, 619.2, 1081.2, 1047.1, 713.3, 725.6, 944.3, 908.5, 626.1]^T .$$

It follows that

$$[a, b]^T = \left[B^T, B\right]^{-1} B^T Y = [0.02424, 1032.7399]$$
$$X_p^{(0)}(k+1) = (x^{(0)}(1) - \frac{b}{a})(1 - e^a)e^{-ak}, \quad k = 1, 2, \cdots, n-1$$
$$X_p^{(0)}(k+1) = 1021.8738 \times e^{-ak}.$$

Droughts and water shortage crisis in West Asia and the Middle East (including a large part of Iran) have an increasing trend over the next three decades. The phenomenon fluctuating near the Middle East is one of the factors affecting annual rainfall reduction. One of the reasons for falling of the rainfall is the global warming. In this situation, when faced with a serious water crisis, what matters is the use of rainfall to both reducing the groundwater levels and preventing flood damage.

Table 1: Actual and estimated rainfall prediction values

Year	Actual value	GM(1,1)	DGM(2,1)	Grey Verhulst
2006	956.9	956.9	956.9	956.9
2007	812.7	997.4	1001.5	935.5
2008	1032.9	973.5	1220.3	906.1
2009	1234.8	950.2	1965.1	866.4
2010	619.2	927.4	4500.3	814.5
2011	1081.2	905.2	13129.9	749.1
2012	1047.1	883.5	42504.6	670.8
2013	713.3	862.4	142495	582.2
2014	725.6	841.7	482857.6	488.3
2015	944.3	821.6	1641436.4	395.5
2016	908.5	801.9	5585185.3	309.8
2017	626.1	782.7	19009525.5	235.5
2018	-	763.9	64705363.8	126.9
2019	-	745.6	2202511900.8	90.7
2020	-	727.7	749725175.8	64.2
2021	-	710.3	2552027910.1	45
2022	-	693.3	8686983723.1	31.3

According to Table 1, based on the evaluations, the rainfall predictions for the years between 2018 to 2022 by GM(1,1) model were 763.9, 745.6, 727.7, 710.3 and 693.3 mm, respectively. Also, the predicted values of the DGM(2,1) and Grey Verhulst models are in columns 4 and 5, respectively. More accurate precipitation estimation as an input to many analytical models is needed for multiple purposes in climate .In Figure 1, the comparison of the predicted values of the GM(1,1), and Verhulest models with actual values and linear forecast trendline is shown.

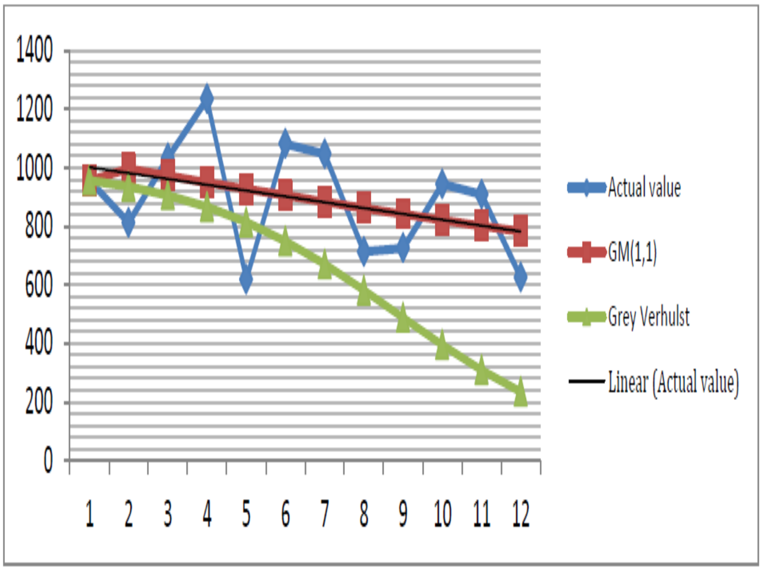

Figure 1: Actual value and estimated rainfall prediction values by GM(1,1) and Verhulest models

According to Table 2, for the years from 2006 to 2017, the error estimations for MAE, RMSE and MAPE in GM(1,1) model were 166.14, 180.41 and 19.98%, respectively. Hence, it can be stated that the model prediction was in a good level based on the MAPE values being in a range from10% to 20%. Also, the error estimations of DGM(2,1) and Grey Verhulst models are in columns 3 and 4, respectively.

Table 2: Compare errors in GM(1,1), DGM(2,1) and Grey Verhulst models

Model	GM(1,1)	DGM(2,1)	Grey Verhulst
MAE	166.14	2446916	311.65
RMSE	180.41	5995968.73	349.36
MAPE	19.98	356054.54	35.72

In order to carry out a drought prediction for this specific region, the following sequence gives the annual average precipitations (in mm) of a certain region for 21 years, where $x(1), x(2), \cdots, x(21)$ are respectively, the data for the years of 1997, 1998, ..., 2017:

$$\begin{aligned} X &= \{x(1), x(2), \cdots, x(21)\} \\ &= \{1157.3, 640.6, 861.8, 772.1, 828.7, 1350.9, 1026.3, 1137.6, 666.9, 956.9, \\ &\quad 812.7, 1032.9, 1234.9, 619.2, 1081.2, 1047.1, 713.3, 725.6, 944.3, 908.5, 626.1\}. \end{aligned}$$

According to the experience of scientists and weather experts, we take $\zeta = 666.9$ mm as a lower abnormal (drought) value. Therefore, the following lower catastrophe sequence is obtained

$$X_\zeta = (x[2], x[9], x[14], x[21]) = (640.6, 666.9, 619.2, 626.1)$$

with the corresponding catastrophe date sequence

$$Q^{(0)} = (q[1], q[2], q[3], q[4]) = (640.6, 666.9, 619.2, 626.1)$$

the GM(1,1) ordinality response sequence of the catastrophe date sequence is

$$a = -0.41266$$
$$b = 6.41344$$
$$\hat{q}(k+1) = 5.9310 \times e^{-ak}.$$

Thus, we can obtain a simulated sequence for $Q^{(0)}$ as follows

$$\hat{Q}^{(0)} = (\hat{q}(1), \hat{q}(2), \hat{q}(3), \hat{q}(4)) = (5.9310, 8.9607, 13.5381, 20.4539)$$

From $\varepsilon(k) = q(k) - \hat{q}(k), k = 1, 2, 3, 4$, we obtain the error sequence as follows:

$$\varepsilon^{(0)} = (\varepsilon(1), \varepsilon(2), \varepsilon(3), \varepsilon(4)) = (3.931, 0.0393, 0.4619, 0.5461)\,.$$

And from $\Delta_k = |\dfrac{\varepsilon(k)}{q(k)}|; k = 1, 2, 3, 4$, it follows that the sequence of relative errors is

$$\Delta = (\Delta_2, \Delta_3, \Delta_4) = (0.4\%, 3.3\%, 2.6\%)\,.$$

From this sequence, we calculate the average relative error $\bar{\Delta} = \frac{1}{3}\sum_{k=2}^{4}\Delta_k = 2.1\%$. With $1 - \bar{\Delta} = 97.9\%$ as the average relative accuracy, and $1 - \Delta_4 = 97.4\%$ Therefore, we can use

$$\hat{q}(k+1) = 5.9310 \times e^{-ak}$$

to carry out our predictions. Because $\hat{q}(5) \approx 30.9024, \hat{q}(5) - \hat{q}(4) \approx 10$ We predict that in 10 years, counting from the time of the last drought in 2027, there might be a drought. Therefore, managers should plan future policies to maximize the cost and efficiency of their facilities.

5 Conclusion

Rainfall has an important role in the global water and energy cycle. The prediction and estimation of atmospheric precipitation for each region and watershed is considered as one of the most important climatic parameters in the optimal use of water resources. There are various methods for predicting the natural accidents and incidents. The majority of the prediction methods need a large number of data. Hence, statistical methods are employed to investigate the systems properties. Grey prediction model features the advantage of creating a model using a low volume of the uncertain data for an appropriate precision and it is an appropriate instrument for predicting systems in uncertain environments.By comparing the predicted values of rainfall amount by these models of GM(1,1), DGM(2,1) and Verhulest models, it is concluded that the forecasting accuracy of the GM(1,1) method is better than other models. The GM(1,1) model used in this paper was one of the most commonly used models in the grey system theory. According to the prediction of rainfall amount and investigation of its rate using Actual values, it can be stated that there was an acceptable error level in this model. Also, We predict that in 2027, there might be a drought. The other methods may have other parameters influencing the rainfall amount prediction. On the contrary, grey prediction method used only rainfall values to perform rainfall prediction for the forthcoming years. The more advanced models with more variables and/or even nonlinear models can be applied based on this method to perform the predictions of various kinds of hydrological streams and determine the relationships between independent and dependent variables. When data be uncertain, the authors may seek to use fuzzy gray systems to predict rainfall in their next studies.

References

[1] Alvisi S., Franchini M. (2012). “Grey neural networks for river stage forecasting with uncertainty”, Physics and Chemistry of the Earth, 42-44, 108–118.

[2] Box G.E., Jenkins G.M., Reinsel G.C., Ljung G.M. (2015). “Time Series Analysis: Forecasting and Control”, John Wiley & Sons: Hoboken, NJ, USA.

[3] Chuan C. (1997). “Weather prediction using artificial neural network”, Hydrology, 230, 110–119.

[4] Deng J.L. (1982). “The control problems of grey systems”, System Control Letter, 5, 288–294.

[5] Deng J.L. (1989). “Introduction to grey system theory”, The Journal of Grey System, 1(1), 1–24.

[6] Diez J., Deljesus, M. (2017). “A rainfall analysis and forecasting tool”, Environmental Modeling Software, 97, 243–258.

[7] EL-Shafie A., Mukhlisin M., Najah A.A., Taha M.R. (2011). “Performance of artificial neural network and regression techniques for rainfall-runoff prediction”, Physical Sciences, 6(8), 1997–2003.

[8] Kayacan E., Ulutas B., Kaynak O. (2010). “Grey system theory-based models in time series prediction”, Expert Systems with Applications 37, 1784–1789.

[9] Lewis C.D. (1982). “Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting”, Butterworth Scientific London, UK.

[10] Li J. (2012). “The optimization DGM(2,1) model and its application”, Journal of Grey System, 24(2), 181–186.

[11] Li, C., Cabrera D., Oliveira J.D., etc, (2016). “Extracting repetitive transients for rotating machinery diagnosis using multiscale clustered grey infogram”, Mechanical System on Signal Processing. 76-77, 157–173.

[12] Li K., Zhang, T. (2018). “Forecasting Electricity Consumption Using an Improved Grey Prediction Model ”, Information, 9, 1–18.

[13] Liu S., Lin L. (2006). “Grey Information Theory and Practical Applications”, Springer, London.

[14] Liu S.F., Forrest J., Yang Y.J. (2012). “A brief introduction to grey systems theory”, Grey System: Theory Application, 2(2), 89–104.

[15] Liu S.F., Lin Y. (2010). “Grey Systems Theory and Applications ”, Springer-Verlag, Berlin Heidelberg, 107–147.

[16] Liu S., Yang Y. and Forrest J. (2017). “Grey Data Analysis”, Springer, Singapore.

[17] Liu S., Yang Y., Xie N., Forrest J. (2016). “New progress of grey system theory in the new millennium”, Grey Systems: Theory and Application, 6(1), 2–31.

[18] Liu S., Zeng B., Liu J., Xie N., Yang Y. (2015). “Four basic models of GM(1,1) and their suitable sequences”, Grey Systems: Theory and Application, 5(2), 141–156.

[19] Miao C., Ashouri H., Hsu K., Sorooshian S., Duan Q. (2015). “Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over china”, Journal of Hydrometeorology, 16(3), 1387–1396.

[20] Neiman P., White A., Ralph F., Gottas D., Gutman S. (2009). “A water vapour flux tool for precipitation forecasting”, Proceedings of the Institution of Civil Engineers Water Management, 162, 83–94.

[21] Obukhov A.M. (1971). “Turbulence in an atmosphere with a non-uniform temperature”, Boundary-layer Meteorology, 2(1), 7–29.

[22] Ren J. (2018). “GM(1,N) method for the prediction of anaerobic digestion system and sensitivity analysis of influential factors”, Bioresource Technology, 247, 1258–1261.

[23] Sorooshian S., AghaKouchak A., Arkin P., Eylander J., Foufoula-Georgiou E., Harmon R., et al., (2011). “Advancing the remote sensing of precipitation”, Bulletin of the American Meteorological Society. 92(10), 1271–1272.

[24] Tsai S.B., Xue Y. Zhang J., Chen Q., Liu Y., Zhou J., Dong W. (2017). “Models for forecasting growth trends in renewable energy”, Renewable and Sustainable Energy Reviews, 77, 1169–1178.

[25] Weia B., Xie A., Hu A. (2018). “Optimal solution for novel grey polynomial prediction model”, Applied Mathematical Modelling, 62, 717–727.

[26] Wen K. L., Huang Y. F. (2004). “The development of grey Verhulst toolbox and the analysis of population saturation state in taiwan–fukien”, International Conference on Systems, Man and Cybernetics, Netherlands, 6, 5007–5012.

[27] Wright D., Mantilla R., Peters-Lidard C. (2017). “A remote sensing-based tool for assessing rainfall-driven hazards”, Environmental Modelling and Software. 90, 34–54.

[28] Wu L., Liu S., Fang Z., Xu H. (2015). “Properties of the GM(1,1) with fractional order accumulation”, Applied Mathematics and Computation, 252, 287–293.

[29] Ye R., Lu X., Liu H. (2013). “Local tomography based on grey model”, Neurocomputing, 101(4), 10–17.

[30] Zeng B., Tan Y., Xu H., Quan J., Wang L., Zhou X. (2018). “Forecasting the electricity consumption of commercial sector in Hong Kong using a novel grey dynamic prediction model”, Journal of Grey System, 30(1), 159–174.

کاربرد نظریه سیستم های خاکستری در پیش بینی بارش باران

درویشی سلوکلایی، د.
استادیار گروه ریاضی- نویسنده مسئول
ایران، تهران، دانشگاه پیام نور، گروه ریاضی، صندوق پستی ۳۶۹۷-۱۹۳۹۵
d_darvishi@pnu.ac.ir

لیو، س.
استاد گروه اقتصاد و مدیریت،
چین، نانجینگ، دانشگاه هوانوردی و فضانوردی، گروه اقتصاد و مدیریت، مرکز پژوهشی مطالعات سیستم های خاکستری
sfliu@nuaa.edu.cn

بابایی ولوی، پ.
کارشناسی ارشد، گروه ریاضی،
ایران، تهران، دانشگاه پیام نور، گروه ریاضی، صندوق پستی ۳۶۹۷-۱۹۳۹۵
p.babaeivalohi@gmail.com

تاریخ دریافت: ۲۳ مهر ۱۳۹۷ تاریخ پذیرش: ۱۰ اردیبهشت ۱۳۹۸

چکیده

با وجود ایران در اقلیم خشک و نیمه‌خشک، پیش‌بینی میزان بارش به منظور مدیریت منابع آب بسیار مهم و ضروری است. روش‌های پیش‌بینی مختلفی در زمینه پیش‌بینی توسط پژوهشگران ارائه گردیده که در زمینه آب و هواشناسی خصوصا مدیریت منابع آب مورد استفاده قرارگرفته است. هدف از این پژوهش، پیش‌بینی میزان بارش باران با استفاده از مدل پیش‌بینی خاکستری که رویکردی نوین در مواجهه با عدم قطعیت است، در منطقه آبخیز بابل رود به منظور مدیریت منابع آبی می‌باشد. ازاین‌رو، ضمن بیان مفاهیم روش پیش‌بینی خاکستری با استفاده از داده‌های جمع‌آوری‌شده در یک دوره ۱۲ ساله ۲۰۰۶ تا ۲۰۱۷، پیش‌بینی بارش سه روش $GM(1,1)$،$DGM(2,1)$ و Verhulest در سال‌های ۲۰۱۸ تا ۲۰۲۲ محاسبه گردید. با توجه به خطای محاسبه شده و قدرت پیش‌بینی، روش $GM(1,1)$ از روش‌های دیگر بهتر بوده و در دسته پیش‌بینی‌های خوب قرار گرفته است. همچنین پیش بینی شده است در سال ۲۰۲۷ در این منطقه خشکسالی داشته باشیم. با توجه به حجم کم داده‌های مورد نیاز و محاسبات در این رویکرد پیشنهاد می‌شود در محیط‌های عدم قطعیت، این رویکرد برای پیش‌بینی‌ها استفاده گردد. پژوهشگران می‌توانند از سیستم‌های فازی خاکستری برای پیش‌بینی میزان بارش باران در محیط‌های عدم قطعیت استفاده کنند.

کلمات کلیدی

پیش‌بینی، سیستم خاکستری، مدیریت منابع آب، میزان بارش باران، خطای پیش‌بینی مطلق.