

## Water Demand Prediction of Grey Markov Model Based on GM(1, 1)

Hang Li<sup>1</sup>, Chi Zhang<sup>4</sup>, Dong Miao<sup>4</sup>, Tong Wang<sup>1,3</sup>, Yang Feng<sup>1</sup>,  
Hao Fu<sup>1</sup>, Chunsheng Zhang<sup>1</sup>, Ming Zhao<sup>2,\*</sup>

<sup>1</sup>School of Environmental Science and Engineering, Chang'an University, Xi'an, 710054, China

<sup>2</sup>School of Municipal and Environmental Engineering, Harbin Institute of Technology, Harbin, 150001, China

<sup>3</sup>Tianjin Sambo Water Technology Co., Ltd. Tianjin, 300070, China

<sup>4</sup>Tianjin Jinbin Veolia Water Co., Ltd. Tianjin, 300011, China

\*email: zhming1188@126.com

**Keywords:** Grey Model; Markov Model; Water Demand Prediction; Software MATLAB

**Abstract.** When the raw data has greater volatility, alone using GM (1, 1) model can lead to big error of water demand prediction results. For this problem, based on the grey GM (1, 1) model and Markov model theory and combine the two algorithm to constitute a grey Markov model which is reasonable and applied to the domestic water demand forecasting project of water supply area in the JV Company of Tianjin. Results show that in the case of original data volatile grey Markov model overcomes the disadvantages of grey GM (1, 1) model and the insufficiency, improves the prediction precision, and enhances the scientific nature and reliability of results.

### Introduction

Water demand prediction [1,2] of urban water supply system, has a prior estimate on the change and the characteristics of urban water demand, is to guarantee the safe operation of the water supply system[3], and realizes the scientific management and scheduling as well as the important content of necessary to make reasonable planning, also is an effective means of urban water resources planning and management. Therefore, urban water demand forecasting has very important significance. And there are a lot of methods [4-6] for urban water demand prediction, which is often used with grey model [7,8], the exponential smoothing model [9], regression analysis model [10], BP neural network model [11], etc. How to effectively use these methods, and providing the prediction model of urban water demand reasonably is an important problem to need study and solve at present urgently. A case study of domestic water demand prediction of the JV Company's area of water supply in Tianjin, by using the grey Markov model based on GM (1, 1), the author used the software MATLAB to program and forecast, which played the advantages of less data and high prediction accuracy for grey system, at the same time, utilized Markov chains' characteristics of forecasting precision for volatile data. Compared with GM (1, 1) model, the grey Markov model is proven to be high accuracy and reliability.

### Grey Markov Model Based on GM(1, 1)

Establishment of GM(1, 1)

A progression of historical water consumption as:

$$q^{(0)} = \{q^{(0)}(1), q^{(0)}(2), \dots, q^{(0)}(n)\} \quad (1)$$

Accumulative calculation at a time is:

---

Fund project: Research and Demonstration on Key Technologies of Intelligent Management System for Urban Water Supply Network during the Twelfth Five-year Plan Period(No.2014ZX07406003).

$$q^{(1)} = \{q^{(1)}(1), q^{(1)}(2), \dots, q^{(1)}(n)\} \quad (2)$$

Differential equation made up of grey module (2) is:

$$\frac{dq^{(1)}}{dt} + \alpha q^{(1)} = u \quad (3)$$

In the differential equations (3),  $Q$  value occurs in the average time in  $k$  and  $k + 1$ . Namely the formula (3) can be written as:

$$\alpha^{(1)} [q^{(1)}(k+1)] + \frac{1}{2} \alpha [q^{(1)}(k+1) + q^{(1)}(k)] = u \quad (4)$$

Written in matrix form as:

$$\begin{bmatrix} q^{(0)}(2) \\ q^{(0)}(3) \\ \dots \\ q^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} [q^{(1)}(1) + q^{(1)}(2)] & 1 \\ -\frac{1}{2} [q^{(1)}(2) + q^{(1)}(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2} [q^{(1)}(n-1) + q^{(1)}(n)] & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ u \end{bmatrix} \quad (5)$$

Use the following formula corresponding to the matrix which is:

$$Y_n = BA \quad (6)$$

In the above equation,  $A$  and  $B$  as known quantity,  $A$  is undetermined parameters. Because  $a$  and  $u$  are the only two variables, and the number of equations is  $n - 1$ , and  $n - 1 > 2$ , so the equations are without solution. Least-square approximation solution can be used by the least square method. so the equation (6) can be written as:

$$Y_n = B\hat{A} + E \quad (7)$$

In the above equation,  $E$  is the error term and  $\hat{A}$  is the least squares estimate.

Least squares calculation formula is:

$$\min \|Y_n - B\hat{A}\|^2 = \min (Y_n - B\hat{A})^T (Y_n - B\hat{A}) \quad (8)$$

The least squares solution of equation (7) is obtained as:

$$\hat{A} = (B^T B)^{-1} B^T Y_n = \begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix} \quad (9)$$

Obtain a time response equation of the model by replacing (6) as:

$$\hat{q}^{(1)}(t) = [q^{(1)}(1) - \frac{\hat{u}}{\hat{a}}] e^{-\hat{a}t} + \frac{\hat{u}}{\hat{a}} \quad (10)$$

Written in a discrete form as:

$$\hat{q}^{(1)}(k+1) = [q^{(0)}(1) - \frac{\hat{u}}{\hat{a}}] e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}} \quad (k=0,1,2,\dots) \quad (11)$$

So prediction model equation of GM (1,1) is:

$$\hat{q}^{(0)}(k+1) = \hat{q}^{(1)}(k+1) - \hat{q}^{(1)}(k) \quad (12)$$

Precision of model test

Backward error detection is a test method according to the statistical situation between model prediction and the actual value, which are transplanted from probability prediction method. Its content is: based on the residual error (absolute error), according to the size of each phase of the absolute value of the residuals, examine the concept of residual smaller point appears, and the size of the prediction error variance related indexes.

A progression of historical water consumption as:

$$q = \{q(1), q(2), \dots, q(n)\} \quad (13)$$

Predictive value sequence as:

$$\hat{q} = \{\hat{q}(1), \hat{q}(2), \dots, \hat{q}(n)\} \quad (14)$$

The residual in  $K$  time is:

$$e(k) = q(k) - \hat{q}(k) \quad (k=1, 2, \dots, n) \quad (15)$$

The mean  $\bar{q}$  of actual value  $Q(k)$  is:

$$\bar{q} = \frac{1}{n} \sum_{k=1}^n q(k) \quad (k=1, 2, \dots, n) \quad (16)$$

The mean  $\bar{e}$  of residual  $e(k)$  is:

$$\bar{e} = \frac{1}{n} \sum_{k=1}^n e(k) \quad (k=1, 2, \dots, n) \quad (17)$$

The variance  $S_1^2$  of historical data is:

$$S_1^2 = \frac{1}{n} \sum_{k=1}^n [q(k) - \bar{q}]^2 \quad (k=1, 2, \dots, n) \quad (18)$$

The variance  $S_2^2$  of residual is:

$$S_2^2 = \frac{1}{n} \sum_{k=1}^n [e(k) - \bar{e}]^2 \quad (k=1, 2, \dots, n) \quad (19)$$

$$C = S_2 / S_1 \quad (20)$$

$$p = p(|e(i) - \bar{e}| < 0.6745S_1) \quad (21)$$

Depending on the expression of  $S_1$  and  $S_2$ , small error probability  $P$  and posterior error ratio of  $C$  can be obtained by the following equation. And model accuracy level standard is seen in Table 1:

**Table 1.** Model precision grade

Model precision grade	C	P
Level 1 (Good)	$\leq 0.35$	$\geq 0.95$
Level 2 (Qualified)	$0.35 < C \leq 0.50$	$0.80 \leq p < 0.95$
Level 3 (Reluctant)	$0.50 < C \leq 0.65$	$0.70 \leq p < 0.80$
Level 4 (Unqualified)	$> 0.65$	$< 0.70$

Markov prediction steps

Compare the predicted value and the actual value of GM (1, 1) model and get errors to form a residual series, then the series is divided into “s” status, and finally obtain the resulting state  $E$  and then do the Markov prediction in order to increase the prediction precision of the model.  $P_{ij}^{(k)}$  is the probability of predict object transferring by  $K$  times from  $E_i$  state to  $E_j$  state as follow:

$$P_{ij}^{(k)} = \frac{M_{ij}^{(k)}}{M_i} \quad (i, j=1, 2, 3, \dots, n) \quad (22)$$

In the above equation,  $M_{ij}^{(k)}$  is the times transferring by  $K$  times from  $E_i$  state to  $E_j$  state. And  $M_i$  is the total times of  $E_i$  state occurring.

$K$ th-transition probability matrix is:

$$P^{(k)} = \begin{bmatrix} P_{11}^{(k)} & P_{12}^{(k)} & \dots & P_{1s}^{(k)} \\ P_{21}^{(k)} & P_{22}^{(k)} & \dots & P_{2s}^{(k)} \\ \dots & \dots & \dots & \dots \\ P_{s1}^{(k)} & P_{s2}^{(k)} & \dots & P_{ss}^{(k)} \end{bmatrix} \quad (i=1, 2, \dots, s) \quad (23)$$

Select “s” objects from the forecast object recently and set their corresponding state as initial state. In the transfer matrix of the transfer step  $K$ ,  $P_i^{(k)}$  is the corresponding row vector of the initial state. The corresponding state of  $\max\{P_i = \sum_{k=1}^s P_{si}^k, i \in E\}$  is the state of forecasting object, and the prediction value of this object is:

$$\hat{q}(k) = \hat{q}^{(0)}(k) + 0.5(a_{1i} + a_{2i}) \quad (24)$$

### case analysis

Tianjin JV Company is a water company, the main scope of the present water supply is north of the Haihe River, south of outer area, referred to as Shibeil Mature District. Some of the areas in the future there will be the JV Company's scope of water supply, which is called Development Area and Extended Area respectively. Based on the future of the company's development plan, it needs to supply regional water demand prediction. In the project water demand is divided into four parts such as Life water demand, Production water demand, Ecological water demand, and Other water demand, among which life water demand occupy a large proportion, so accurately predict life water demand is very necessary. Now from Tianjin JV Company get data of life water demand in the area of water supply in 2007-2014, and the results of prediction using GM(1,1) model is seen in Table 2 as follow:

**Table 2.** Life water demand of Tianjin JV Company water supply and forecast results

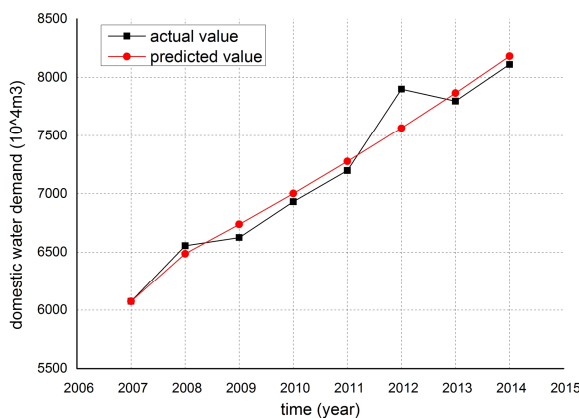
Time/(year)	Actual value/ ( $10^4\text{m}^3$ )	Predicted value/ ( $10^4\text{m}^3$ )	Absolute error	relative error (%)
2007	6075	6075	0	0
2008	6556	6487	69	1.0525
2009	6626	6739	-113	1.705
2010	6932	7002	-70	1.01
2011	7198	7277	-79	1.098
2012	7900	7565	335	4.2405
2013	7797	7866	-69	0.885
2014	8111	8182	-71	0.875

From Table 2 it shows that the error between the actual value and the predicted value has certain volatility, if separately use GM (1, 1) model to forecast, it will cause large error, therefore based on GM (1, 1) Grey Markov model to modify predicted results can be relatively accurate forecasts. According to Table 1, the absolute error data sequence can be divided into five state, and the upper limit is included in the range, otherwise the lower limit is not included in the range, specific seen in Table 3.

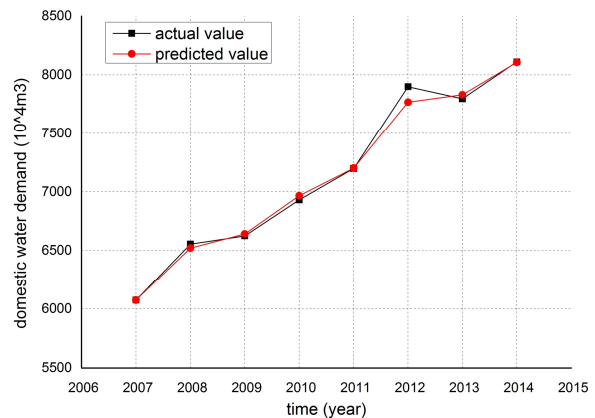
**Table 3.** State partition

State	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$
State limit	(-115,-80)	(-80,-70)	(-70,0)	(0,70)	(70,335)

According to the state partition of the table 3, it can be easily obtained that the state from 2007 to 2014 respectively is  $E_4, E_4, E_1, E_3, E_2, E_5, E_3, E_2$ . with the prediction steps of Markov model to calculate, forecast results after correction can be obtained as shown in Fig 2.



**Fig 1.** Forecast of life water demand based on GM (1,1)



**Fig 2.** Forecast of life water demand based on Grey Markov model

By comparison of Fig 1 and Fig 2, it is known that the results of Grey Markov model based on GM (1, 1) predict more in line with the actual income. And the state of prediction in 2015 is  $E_5$  according to above model. Therefore, the volume of life water demand predicting in 2015 is 87.15 million m<sup>3</sup>.

## Conclusion

When using GM (1,1) model fits relatively large fluctuations in data, the error will be relatively large, and the accuracy is not high enough. Through the establishment of Grey Markov model based on GM(1,1), it can to some extent be overcome by a Markov chain by transition probability, the prediction accuracy of gray Markov model is more higher than the GM (1, 1) model, and the effect is more better. Thus, in the field of water demand prediction we can get a lot of use of this model.

## Acknowledgments

This project is supported by The Key Projects in the National Science & Technology Pillar Program for China during the Twelfth Five-year Plan Period .And special thanks to Hongbin Zhao professor of Harbin Institute of Technology.

## References

- [1] Hall, J. W., Watts, G., Keil, M., de Vial, L., Street, R., Conlan, K., O'Connell, P. E., Beven, K. J., Kilsby, C. G.:Water and Environment Journal, 26 (1):118-129, 2012.

- [2] Emmanuel A. Donkor, S.M.ASCE; Thomas A. Mazzuchi; Refik Soyer; and J. Alan Roberson, P.E. :Water Resource Plan Manage, 140(2): 146-159,2014.
- [3] Bakker, M., Vreeburg, J.H.G., Palmén, L.J., Sperber, V., Bakker, G., Rietveld, L.C.: J. Water Supply: Res. Technol. AQUA 62(1):1-13,2013
- [4] Christopher J. Hutton, Zoran Kapelan. :Environmental Modelling & Software 66 (2015) :87-97.
- [5] Alvisi, S., Franchini, M., Marinelli, A.:J. Hydroinform. 9 (1): 39-50,2007.
- [6] Herrera, M., Torgo, L., Izquierdo, J., Perez-García, R.:J. Hydrol. 387 (1): 141-150,2010.
- [7] Wang Ming-dong, Kong Bin, Li Xiang-yang, Wang Xue-ke.: Electrical Power and Energy Systems 79 (2016) :188–195.
- [8] Zhaoqing Hu, Chengxiong Mao, Jiming Lu.:Proc EPSA 2002;14(1):30–3.
- [9] Federico Poloni, Giacomo Sbrana.: J. Production Economics 162 (2015) :143–150.
- [10]Junhong Li , Wenyi Zeng , Jianjun Xie , Qian Yin.:Engineering Applications of Artificial Intelligence 52 (2016) :54–64.
- [11]Cheng Cheng, Xiaosheng Cheng, Ning Dai, Xiaotong Jiang, Yuchun Sun, Weiwei Li. :Computers in Biology and Medicine 66 (2015): 103–112.