2011 3rd International Conference on Information and Financial Engineering
IPEDR vol.12 (2011) © (2011) IACSIT Press, Singapore

A Markov Chain Grey Forecasting Model: A Case Study of Energy Demand of Industry Sector in Iran

A. Kazemi [1+], M. Modarres [2], M.R. Mehregan [1], N. Neshat [2] and A. Foroughi.A [3]

[1] Faculty of Management, University of Tehran, Tehran, Iran

[2] Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran

[3] National Iranian Oil Company, Tehran, Iran

Abstract. The aim of this paper is to develop a prediction model of energy demand of industry sector in Iran. A Markov chain grey model (MCGM) is proposed to forecast the energy demand of industry sector. It is then compared to grey model (GM) and regression model. The comparison reveals that the MCGM has higher precision than GM and regression model. The MCGM is then used to forecast the annual energy demand of industry sector in Iran up to the year 2020. The results provide scientific basis for the planned development of the energy supply of industry sector in Iran.

Keywords: Energy demand, Industry sector, Forecasting, GM, MCGM, Regression

1. Introduction

In today's competitive world, the better an organization can predict and analyze the developing trend in future based on past facts, the better chance it has to survive. Various technical and statistical methods for energy demand forecasting have been proposed in the last few decades with varying results. GM is suitable for forecasting the competitive environment where decision makers can refer only to a limited historical data. But the forecasting precision for data sequences with large random fluctuation is low. The Markov-chain forecasting model can be used to forecast a system with randomly varying time series. It is a dynamic system which forecasts the development of the system according to transition probabilities between states which reflect the influence of all random factors. So the Markov-chain forecasting model is applicable to problems with random variation, which could improve the GM forecasting model [1].

The applications of grey model for energy forecasting problems have resulted in several research papers [1-3]. In 2001, Zhang and He have developed a Grey–Markov forecasting model for forecasting the total power requirement of agricultural machinery in Shangxi Province [2]. In 2007, Akay and Atak have formulated a Grey prediction model with rolling mechanism for electricity demand forecasting of Turkey [3]. A Grey–Markov forecasting model has been developed by Huang, He and Cen in 2007. This paper was based on historical data of the electric-power requirement from 1985 to 2001 in China, and forecasted and analyzed the electric- power supply and demand in China [1].

In this paper energy demand of industry sector in Iran has been forecasted using the MCGM for the time span 2009 to 2020. For the estimation, time series data covering the period 1990 to 2008 have been used. This model is compared with GM and regression model. The remaining parts of the paper are organized as follows. In the second and third sections, GM and MCGM are presented. Details of applying GM, MCGM and regression model for energy demand of industry sector in Iran and obtained numerical results are described in the section 4. Section 5 analyzes and compares the empirical results obtained from the three forecasting models. A brief review of the paper and the future research are in Section 6.

[+] Corresponding author.

E-mail address: aliyehkazemi@ut.ac.ir

2. GM forecasting model

Grey prediction (GP) has three basic operations: accumulated generating operator (AGO), inverse accumulating operator (IAGO) and grey model (GM). The steps of GP are shown below.

Step 1: Original time sequence with n samples (time point) is expressed as

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\} \tag{1}$$

AGO operator is used to convert chaotic series $x^{(0)}$ into monotonically increasing series $x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\}$, where $x^{(1)}(j) = \sum_{i=1}^{j} x^{(0)}(i)$

It is obvious that the original data $x^{(0)}(i)$ can be easily recovered from $x^{(1)}(i)$ as

$$x^{(0)}(i) = x^{(1)}(i) - x^{(1)}(i-1) \tag{2}$$

where $x^{(0)}(1) = x^{(1)}(1), x^{(1)}(i) \in x^{(1)}$. This operation is called IAGO [4].

Step 2: Form the GM model by establishing a first order grey differential equation

$$x^{(0)}(i) + az^{(1)}(i) = b \tag{3}$$

where $z^{(1)}(i) = \frac{1}{2}(x^{(1)}(i) + x^{(1)}(i+1)$

In Eq. (3), $i(i = 2,...,n)$ is a time point. a is called the development coefficient and b is called driving coefficients. Using least mean square estimation technique coefficients, $[a,b]^T$ can be estimated as

$$\begin{bmatrix} a \\ b \end{bmatrix} = (A'A)^{-1} A'X_n' \quad , \quad A = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad , \quad X_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \tag{4}$$

Step 3: According to the estimated coefficients a and b, GP equation can be obtained by solving differential equation in Eq. (3). By IAGO, the predicted equation is

$$\hat{x}^{(0)}(i+1) = \hat{x}^{(1)}(i+1) - \hat{x}^{(1)}(i) = (x^{(0)}(1) - \frac{b}{a})(1 - e^{a})e^{-ai} \tag{5}$$

3. MCGM forecasting model

In this section, Markov chain is presented to enhance the predicted accuracy of GM. The new model is defined as MCGM. The original data are first modelled by the GM, then the residual errors between the predicted values and the actual values for all previous time steps are obtained. The idea of the MCGM is to establish the transition behavior of those residual errors by Markov transition matrices, then the possible correction for the predicted value can be made from those Markov matrices. The detailed procedure is shown as follows [4].

3.1. The division of state

For original data series, use GM forecasting model to obtain the predicted value $\hat{x}^{(0)}(i)$. Then, the residual error $e(i) = x^{(0)}(i) - \hat{x}^{(0)}(i)$ can also be obtained. Assume that there exists some regular information in the residual error series of GM. We can establish Markov state transition matrices; r states are defined for each time step. Thus the dimension of the transition matrix is $r \times r$. The residual errors are partitioned into r equal portions called states. Each state is an interval whose width is equal to a fixed portion of the range between the maximum and the minimum of the whole residual error. Then, the actual error can be classified into those states. Let s_{ij} be the jth state of the ith time step $S_{ij} \in [L_{ij}, U_{ij}], j = 1,2,...,r$ where L_{ij} and U_{ij} are the lower boundary and upper boundary of the jth state for the ith time step of the residual error series. $e(i)$ is residual error of GM.

$$L_{ij} = \min e(i) + \frac{j-1}{r}(\max e(i) - \min(i)), \qquad U_{ij} = \min e(i) + \frac{j}{r}(\max e(i) - \min(i)) \tag{6}$$

3.2. Establishment of transition probability matrix of state

If the transition probability of state is $P_{ij}^{(m)} = \frac{M_{ij}^{(m)}}{M_i}$, $j = 1,2,...,r$, where $P_{ij}^{(m)}$ is the probability of transition from state i to j by m steps. $M_{ij}^{(m)}$ is the transition times from state i to j by m steps and M_i is the number of data belonging to the i th state. Because the transition for the last m entries of the series is indefinable, M_i should be counted by the first as $n-m$ entries; n is the quantity of entries of the original series. Then, the transition probability matrix of state can be written as

$$R^{(m)} = \begin{bmatrix} P_{11}^{(m)} & P_{12}^{(m)} & \cdots & P_{1r}^{(m)} \\ P_{21}^{(m)} & P_{21}^{(m)} & \cdots & P_{2r}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{r1}^{(m)} & P_{r2}^{(m)} & \cdots & P_{rr}^{(m)} \end{bmatrix} \quad (7)$$

3.3. Obtaining the predicted value

The residual error series $e(i)$ is divided into r states, then there is r transition probability row vectors. The possibilities of a certain error state for the next step are obtained by the probabilities in r row vectors, denoted as $\{a_i(T), i = 1,2,...,r\}$ at time step T. Define the centres of r states as $v_i (i = 1,2,...,r)$. Then, the predicted value for the next step is

$$\tilde{x}^{(0)}(T+1) = \hat{x}^{(0)}(T+1) + \sum a_i(T) v_i \quad (8)$$

where $a^{(T)} = [a_1(T), a_2(T),..., a_r(T)] = a^{(T-1)} R^{(m)}$ and

$$\begin{cases} a^{(T+1)} = a^{(T)} R^{(m)} \\ a^{(T+2)} = a^{(T+1)} R^{(m)} \\ \vdots \\ a^{(T+k)} = a^{(T+k-1)} R^{(m)} \end{cases} \quad (9)$$

where $m = 1$.

4. Energy demand forecasting of industry sector in Iran

In 2008, total energy consumption of industry sector was 236.32 million barrel of oil equivalent (MBOE). This figure equals of the 19.9% total final energy consumption in Iran. Thus, the industry sector has been in highest priority from conservation policy perspective. Because of many undermining factors, including economical development, industries technology, the national policy, etc., fluctuation in the energy consumption of industry sector appeared obvious. As Table 1 shows, the historical data series of the energy consumption of industry sector in Iran from 1990 to 2008 is rising, but fluctuating randomly. Therefore, this paper proposed a MCGM forecasting model to forecast energy demand of industry sector in Iran. This model will compare with GM and regression forecasting model.

TABLE I. STATISTICS OF THE ENERGY CONSUMPTION OF INDUSTRY SECTOR IN IRAN FROM 1990 TO 2008 [5]

Years	Energy consumption (MBOE)	Years	Energy consumption (MBOE)	Years	Energy consumption (MBOE)	Years	Energy consumption (MBOE)
1990	112.9	1995	137.85	2000	168	2005	224.27
1991	123.3	1996	149.68	2001	177.55	2006	216.4
1992	130.6	1997	164.4	2002	195.73	2007	232.09
1993	121.6	1998	157.02	2003	209.63	2008	236.32
1994	123.51	1999	167.41	2004	214.36	2005	224.27

3.4. Establishment of prediction model GM and MCGM

Based on the historical data of the energy consumption of industry sector in Iran from 1990 to 2008, a trend curve equation is built by GM forecasting model. GM forecasting model was established by Eq. (5). As the results, the fitted and predicted generated data series $\{\hat{x}^{(0)}(i), i = 1,2,...,n\}$ and original data are plotted in

Fig. 1. According to the predicted data series $\{\hat{x}^{(0)}(i), i=1,2,...,n+k\}$ by GM forecasting model, its residual error series $e(i)$ can obtained as listed in Table 2. From the obtained residual errors, the corresponding intervals are divided into four states for this study. The four states are $v_1 = [-12.18, -6.24]$, $v_2 = [-6.24, -0.29]$, $v_3 = [-0.29, 5.65]$, $v_4 = [5.65, 11.60]$. The four states based on their residual errors are defined and the results are also listed in Table 2.

TABLE II. TABLE. 2: STATE TABLE OF MCGM (STATE NUMBER=4)

Years	$e(i)$	State	Years	$e(i)$	State	Years	$e(i)$	State	Years	$e(i)$	State
1990	-1.8035	2	1995	-3.6611	2	2000	-6.584	1	2005	8.8836	4
1991	3.6757	3	1996	2.098	3	2001	-4.5237	2	2006	-8.2266	1
1992	5.8438	4	1997	10.4867	4	2002	5.8453	4	2007	-2.1731	2
1993	-8.5083	1	1998	-3.4962	2	2003	11.5992	4	2008	-7.9931	1
1994	-12.1799	1	1999	0.0076	3	2004	7.8336	4			

In addition, $\{\hat{x}^{(0)}(i) + v_j, for \quad i=1,2,...,n; j=1,2,3,4\}$ is used to divide original data $\{\hat{x}^{(0)}(i), i=1,2,...,n\}$, and the results are plotted in Fig. 2.

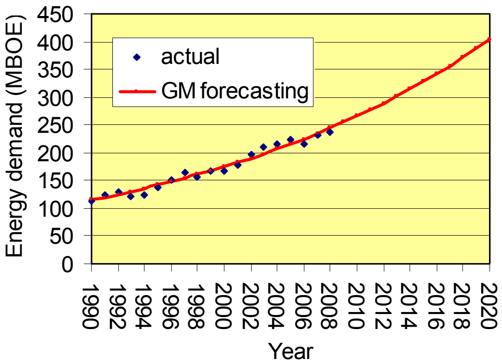

Fig. 1: The fitted and predicted values by GM.

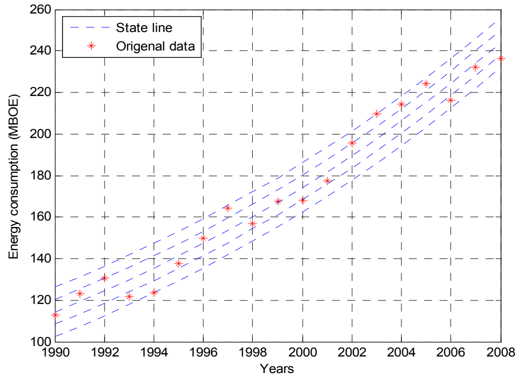

Fig. 2. The state division by residual errors.

By the state of each entry as shown in Table 2, the transition probability matrices of state $R^{(m)}$, $m=1$, can be evaluated. According to the four states, we can calculate their centre values. The model fitted and predicted values by MCGM forecasting model, and the experimental original data are plotted in Fig. 3.

3.5. Establishment of prediction model regression

Based on the historical data of the energy consumption in industry sector of Iran from 1990 to 2008, a trend curve equation is built by regression forecasting model. Regression forecasting model was established by $\hat{x}(i) = 7.2185\text{x(i)} - 14258$ where $x(i)$ is the actual value and $\hat{x}(i)$ is the predicted value. As the results, the model fitted and predicted values by regression forecasting model, and the experimental original data are plotted in Fig. 4.

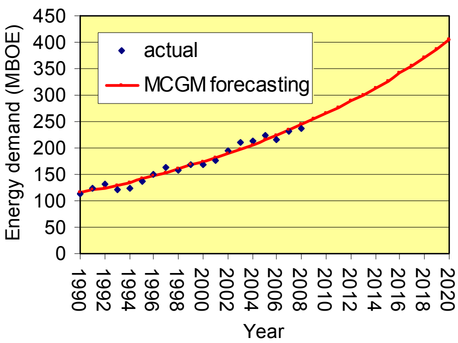

Fig.3: The fitted and predicted values by MCGM.

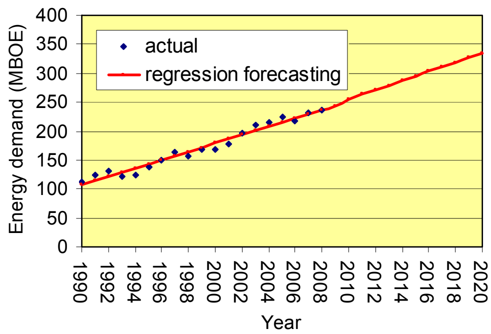

Fig.4: The fitted and predicted values by regression model.

5. Comparison of forecast precision between the GM, MCGM and regression forecasting model

As the above, the forecast values from 1979 to 2006 calculated by GM, MCGM and regression model. The forecast values between the three models are compared and the results are presented in Table 3. The three criteria are used for comparing three models. They are the mean square error (MSE), absolute mean error (AME) and average absolute error percentage (AAEP) which are calculated as

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e^2(i), \qquad AME = \frac{1}{n}\sum_{i=1}^{n}\left|e(i)\right|, \qquad AAEP = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{e(i)}{x(i)}\right| \times 100 \tag{10}$$

where $e(i) = x(i) - \hat{x}(i)$

Table 3 shows that the MCGM forecasting model is better for forecasting the energy demand of industry sector in Iran. The forecast values of MCGM forecasting model are more precise than GM forecasting model and regression forecasting model.

TABLE III. TABLE.3: COMPARISON OF FORECAST RESULTS WITH THREE DIFFERENT METHODS.

Models	Mean Square Error (MSE)	Absolute Mean Error (AME)	Average Absolute Error Percentage (AAEP)
Regression	53.02	6.51	4.17%
GM	48.46	6.07	3.65%
MCGM	48.29	6.04	3.61%

The estimated energy demand of industry sector by MCGM forecasting model from 2009 to 2020 is given in Table 4. In 2020, the energy demand of industry sector will reach to a level of 335 MBOE.

TABLE IV. THE PREDICTED VALUE OF THE ENERGY DEMAND OF INDUSTRY SECTOR IN IRAN BY MCGM FORECASTING MODEL.

Years	Energy demand (MBOE)	Years	Energy demand (MBOE)	Years	Energy demand (MBOE)
2009	244	2013	280	2017	311
2010	256	2014	287	2018	319
2011	264	2015	295	2019	327
2012	272	2016	303	2020	335

6. Conclusions

The major purpose of this paper was to develop the prediction model of energy demand of industry sector in Iran. Through using the statistics data of the energy consumption of industry sector from 1990 to 2008, three forecasting models presented and compared. The results showed that the accuracy of MCGM in forecast energy demand of industry sector is higher than those of GM and regression forecasting model. Also energy demand of industry sector in Iran from 2009 to 2020 was forecasted.

The MCGM forecasting model could be applied to forecast other time series problems with large random fluctuation.

7. Acknowledgements

This research was supported by the Iranian Fuel Conservation Organization (IFCO) and Invited Collaborative Research Program (ICRP).

8. References

[1] Min Huang, Yong He, Haiyan Cen; "Predictive analysis on electric-power supply and demand in China", *Renewable Energy* 2007; 32(7): 1165–1174.

[2] Zhang SJ, He Y. A Grey–Markov forecasting model for forecasting the total power requirement of agricultural machinery in Shangxi Province. *J Shanxi Agric Univ (Nat Sci Edi)* 2001; 21(3): 299–302.

[3] Diyar Akay, Mehmet Atak; "Grey prediction with rolling mechanism for electricity demand forecasting of Turkey", *Energy* 2007; 32(9): 1670-1675.

[4] Guo-Dong Li, Daisuke Yamaguchi , Masatake Nagai; "A GM(1,1)–Markov chain combined model with an application to predict the number of Chinese international airlines", *Technological Forecasting & Social Change*

2007; 74(8): 1465-1481.

[5] Institute for International Energy Studies (IIES), Islamic Republic of Iran (2008) Hydrocarbon Balance, 2010.