Energy Demand Forecast of Iran's Industrial Sector Using Markov Chain Grey Model

Aliyeh Kazemi[1], Mohammad Modarres[2], M.Reza Mehregan[3]

Abstract

The aim of this paper is to develop a prediction model of energy demand of Iran's industrial sector. For that matter a Markov Chain Grey Model (MCGM) has been proposed to forecast such energy demand. To find the effectiveness of the proposed model, it is then compared with Grey Model (GM) and regression model. The comparison reveals that the MCGM model has higher precision than those of the GM and the regression. The MCGM is then used to forecast the annual energy demand of industrial sector in Iran up to the year 2020. The results provide scientific basis for the planned development of the energy supply of industrial sector in Iran.

Keywords: Energy Demand, Industrial Sector, Forecasting, GM, MCGM, Regression.

1. Department of Industrial Management, Faculty of Management, University of Tehran, Tehran, Iran.
   aliyehkazemi@ut.ac.ir
2. Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran.
3. Department of Industrial Management, Faculty of Management, University of Tehran, Tehran, Iran.

1. Introduction

In today's competitive world, the better an organization could predict and analyze the developing future trend based on past facts, the better chance it has to survive. During the past decades, a variety of forecasting methods and computerized systems have been developed either to improve accuracy or to increase computational efficiency. Selecting an appropriate and efficient forecasting method is a challenging and sometimes a complicated decision.

Various technical and statistical methods for energy demand forecast have been proposed in the last few decades with varying results. In several studies, future predictions have been achieved under different scenarios of those indicators without predicting future values of indicators. But, for the developing countries, especially for Iran, aforementioned indicators have neither stable nor predictable trends, even in a long term, owing to political and economic uncertainties or fluctuations taking place in international markets. In other words, it is rather difficult to incorporate quantitatively and handle the uncertain factors such as socio-economic indicators for the energy

demand projections in Iran. Yet, using socio-economic indicators for future energy demand forecasting, estimation of these indicators seems to be another problem that should additionally be studied on them. However, GM only needs recent year's data for reliable and acceptable accuracy for future prediction. This is one of the considerable advantages of GM over previous studies. Besides, it is more practicable and user friendly when compared to Box–Jenkins models and artificial intelligence techniques which require more effort and time for parameter identification and model building phases (Akay and Atak, 2007).

The applications of grey model for energy forecasting problems have resulted in several research papers. Zhang and He (2001), developed a Grey–Markov model for forecasting the total power requirement of agricultural machinery in Shangxi Province. Akay and Atak (2007), formulated a Grey prediction model with rolling mechanism for electricity demand forecasting of Turkey. A Grey–Markov forecasting model was developed by Huang *et al.* (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.

Deng initially presented the Grey system theory in 1982. The Grey forecasting model adopts the essential part of the Grey system theory. The GM forecasting model can be used in circumstances with relatively little data and it can use a first-order differential equation to characterize an unknown system. So the GM forecasting model 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 (Huang *et al.*, 2007).

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 (Huang *et al.*, 2007).

To improve the accuracy of the

forecasting, it is necessary to combine the two models of prediction. The proposed GM forecasting model which combines with Markov Chain is defined as MCGM forecasting model.

Because of many factors, including economical development, industrial technologies, the national policy, etc., fluctuation in the energy consumption of industrial sector appeared obvious. Thus, according to the data of the total energy consumption of the sector, this paper presents a MCGM forecasting model to forecast and analyze the energy demand of industrial sector in Iran.

In this paper, the energy demand of industrial sector in Iran has been forecasted using MCGM for the time span 2011 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 section, grey forecasting model is introduced. MCGM is presented in third section. Details of applying GM, MCGM and regression model for energy demand forecast of industrial 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.

2. Grey Forecasting Model

In recent years, the grey system theory has become a very effective method of solving uncertainty problems under discrete data and incomplete information. The theory includes five major parts, which include grey prediction, grey relation, grey decision, grey programming and grey control. Prediction is to analyze the developing trend in the future according to past facts. Most of the prediction methods need a large number of history data, and will make use of the statistical method to analyze the characteristics of the system. Furthermore, because of additional noise from the outside and the complex interrelations among the system or between the system and its environment, it is more difficult to analyze the system. As a prediction model, the grey dynamic model has the advantages of establishing a model with few and uncertain data and has become the core of grey system theory (Li *et al.*, 2007).

Grey forecasting model has three basic operations: accumulated generating operator (AGO), inverse accumulating operator (IAGO) and grey model (GM) (Li *et al.*, 2007). The steps of GP are shown below.

2.1. AGO

One of the important data operations in grey system theory is the accumulative generating operation, abbreviated as AGO, which is a linear transformation of the original data vector. The role of AGO is to partially eliminate the fluctuation in the original discrete data sequence as long as the original data is strictly positive. Assume that $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2),..., x^{(0)}(n)\}$ is an original series of real numbers with irregular distribution. Then $x^{(1)}$ is viewed as 1-AGO generation series for $x^{(0)}$, if $\forall x^{(1)}(j) \in x^{(1)}$ can satisfy

$$x^{(1)}(j) = \sum_{i=1}^{j} x^{(0)}(i) \qquad (1)$$

Then

$$x^{(1)} = \{\sum_{i=1}^{1} x^{(0)}(i), \sum_{i=1}^{2} x^{(0)}(i),..., \sum_{i=1}^{n} x^{(0)}(i)\} \qquad (2)$$

which is the first-order AGO series obtained from $x^{(0)}$ (Li *et al.*, 2007).

2.2. IAGO

From Eq. (1), it is obvious that the original data $x^{(0)}(i)$ can be easily recovered from $x^{(1)}(i)$ as

$$x^{(0)}(i) = x^{(1)}(i) - x^{(1)}(i-1) \tag{3}$$

where $x^{(0)}(1) = x^{(1)}(1), x^{(1)}(i) \in x^{(1)}$. This operation is called first-order IAGO (Li *et al*., 2006).

2.3. GM

If we have $n \geq 4, x^{(0)}, x^{(1)} \in R^+$, the grey dynamic prediction model GM can be expressed by a one-variable, first-order differential equation.

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{4}$$

The whitening version of Eq. (4) is shown as

$$x^{(0)}(i) + az^{(1)}(i) = b \tag{5}$$

where $z^{(1)}(i)$ is called the background value of $\frac{dx^{(1)}}{dt}$ and is calculated by

$$z^{(1)}(i) = \frac{1}{2}(x^{(1)}(i) + x^{(1)}(i+1)) \tag{6}$$

This kind of method does not have any mathematical proof, we just call it as "whiteness processing" and Eq. (4) is called the "shadow" equation of Eq. (5). Parameter a is called the development coefficient of GM and parameter b is called the grey controlled variable. Both parameters are unknown variables (Hsu nad Wang, 2009). By least-squares method, the coefficients a and b can be obtained as

$$A = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad X_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad \beta = \begin{bmatrix} a \\ b \end{bmatrix} \tag{7}$$

by solving Eq. (5)

$$X_n = \begin{bmatrix} -az^{(1)}(2) + b \\ -az^{(1)}(3) + b \\ \vdots \\ -az^{(1)}(n) + b \end{bmatrix} \tag{8}$$

therefore

$$X_n = A\beta \tag{9}$$

By least-squares method

$$\sum_{i=1}^{n} e_i^2 = e'e = (X_n - A\beta)'(X_n - A\beta) = X_n'X' - 2\beta'A'X_n + \beta'A'A\beta \tag{10}$$

$$\frac{\partial(e'e)}{\partial \beta} = -2A'X_n + 2A'A\beta = 0 \tag{11}$$

$$\beta = \begin{bmatrix} a \\ b \end{bmatrix} = (A'A)^{-1}A'X_n' \tag{12}$$

By using Laplace transform, Eq. (4) can be expressed in frequency domain as

$$sx^{(1)}(s) - u(0) + ax^{(1)}(s) = \frac{b}{s} \tag{13}$$

where u(0) is the initial value of the system. $u(0) = x^{(0)}(1) = x^{(1)}(1)$.

$$x^{(1)}(s) = \frac{x^{(0)}(1) - \frac{b}{a}}{s+a} + \frac{\frac{b}{a}}{s} \tag{14}$$

By Laplace inversion transform, the solution of continuous system form and discrete system form are obtain by

$$\hat{x}^{(1)}(i+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ai} + \frac{b}{a} \tag{15}$$

By IAGO, the predicted equation is,

$$\hat{x}^{(0)}(i+1) = \hat{x}^{(1)}(i+1) - \hat{x}^{(1)}(i) =$$

$$(x^{(0)}(1) - \frac{b}{a})(1-e^a)e^{-ai} \qquad (16)$$

By Eq. (16), the data series $\{x^{(0)}(1), x^{(0)}(2),....,x^{(0)}(n)\}$ are called fitted series, while series $\{\hat{x}^{(0)}(n+1), \hat{x}^{(0)}(n+2),....,\hat{x}^{(0)}(n+k)\}$ are called predicted series.

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, and 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, and then the possible correction for the predicted value can be made from those Markov matrices. The detailed procedure is shown as follows (Li *et al*., 2007).

3.1. The Division of State
3.1.1. Establishment of GM Forecasting Model

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.

3.1.2. Division state by Markov chain

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

$$S_{ij} \in [L_{ij}, U_{ij}], \qquad j=1,2,...,r \qquad (17)$$

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.

$$L_{ij} = \min e(i) + \frac{j-1}{r}(\max e(i) - \min(i)) \qquad (18)$$

$$U_{ij} = \min e(i) + \frac{j}{r}(\max e(i) - \min(i)) \qquad (19)$$

$e(i)$ is residual error of GM.

3.2. Establishment of Transition Probability Matrix of State

If the transition probability of state is written as

$$P_{ij}^{(m)} = \frac{M_{ij}^{(m)}}{M_i}, \qquad j=1,2,...,r \qquad (20)$$

6

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 ith 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

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

3.3. Obtaining the Predicted Value

The residual error series $e(i)$ is divided in to 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

$$\tilde{x}^{(0)}(T+1) = \hat{x}^{(0)}(T+1) + \sum a_i(T)v_i \quad (22)$$

where

$$a^{(T)} = [a_1(T), a_2(T),...,a_r(T)] = a^{(T-1)}R^{(m)} \quad (23)$$
and

$$\begin{cases} a^{(T+1)} = a^{(T)}R^{(m)} \\ a^{(T+2)} = a^{(T+1)}R^{(m)} \\ \quad \vdots \\ a^{(T+k)} = a^{(T+k-1)}R^{(m)} \end{cases} \quad (24)$$

where $m=1$.

4. Energy Demand Forecast of Industrial Sector in Iran

In 2008, total energy consumption of Iran's industrial sector was 236.32 million barrel of oil equivalent (MBOE). This figure equals of the 19.9% total final energy consumption in Iran. A high percentage of the total energy consumption in industry relates to the large industrial factories. Thus, the industrial sector has been in highest priority from conservation policy perspective. There are many factors which could influence the energy demand of industrial sector like population, economic growth, price of energy, the industrial technologies, etc., so the time series of energy demand of industrial sector shows large random fluctuations. As Table 1 shows, the historical data series of the energy consumption of industrial 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 the sector. This model will compare with GM and

7

regression model. In this section GM, MCGM and regression model will establish.

Table 1 Statistics of Energy Consumption of Industrial Sector in Iran from 1990 to 2008 (Institute for International Energy Studies, 2010)

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

4.1. Establishment of GM

Based on the historical data of the energy consumption of industrial sector in Iran from 1990 to 2008, a trend curve equation is built by GM. The model is established by Eq. (16). As the results, the fitted and predicted generated data series $\left\{\hat{x}^{(0)}(i), i = 1,2,...,n\right\}$ and original data are plotted in Fig. 1.
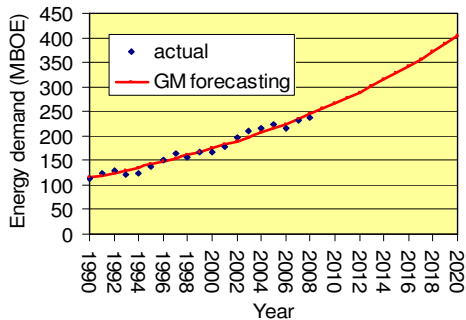

Fig 1 Fitted and predicted Values by GM for the Increase in Energy Demand of Industrial Sector

4.2. Establishment of MCGM
4.2.1. The Division of State

According to the predicted data series $\left\{\hat{x}^{(0)}(i), i = 1,2,...,n + k\right\}$ by GM, 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 states based on their residual errors are defined and the results are also listed in Table 2.

Table 2 State Table of MCGM (state number=4)

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

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

4.2.2. Establishment of Transition Probability Matrix of State

By the state of each entry as shown in Table 2, the transition probability matrices of state $R^{(m)}$, $m=1$, can be evaluated as

$$R^{(1)} = \begin{bmatrix} 0.25 & 0.75 & 0 & 0 \\ 0.2 & 0 & 0.60 & 0.2 \\ 0.33 & 0 & 0 & 0.67 \\ 0.33 & 0.17 & 0 & 0.50 \end{bmatrix} \quad (25)$$
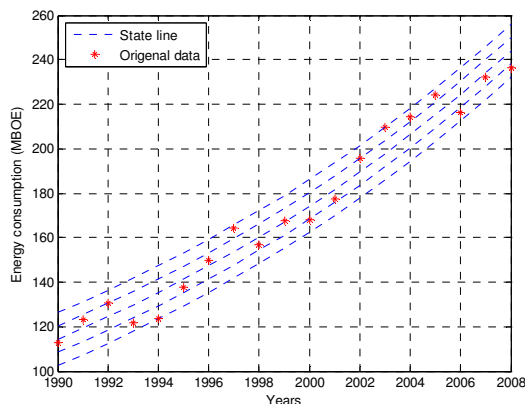

Fig 2 The State Division by Residual Errors for Energy Consumption of Iran's Industrial Sector 1990-2008.

4.2.3. Obtaining Predicted Values

According to the four states, we can calculate their center values. Then $v_1 = -9.21$, $v_2 = -3.26$, $v_3 = 2.68$ and $v_4 = 8.63$ are obtained. The model fitted and predicted values by MCGM, and the experimental original data are plotted in Fig. 3.
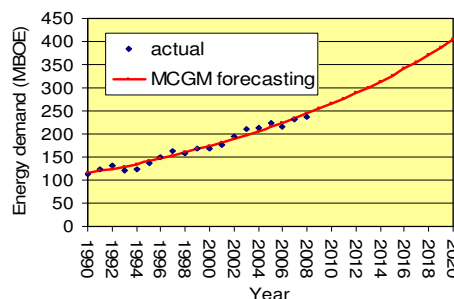

Fig.3: The fitted and predicted values by MCGM for the increase in energy demand of industrial sector.

4.3. Establishment of Regression Model

Conventional regression analyzes is one of the most used statistical tools to explain the variation of a dependent variable Y in terms of the variation of explanatory variables X as: $Y = f(X)$ where $f(X)$ is a linear function. It refers to a set of methods by which estimates are made for the model parameters from the knowledge of the values of a given input–output data set.

Based on the historical data of the energy consumption in industrial sector of Iran from 1990 to 2008, a trend curve equation is built by regression forecasting model. Regression forecasting model was established by Eq. (26).

$$\hat{x}(i) = 7.2185x(i) - 14258 \quad (26)$$

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

9

The actual extraction.

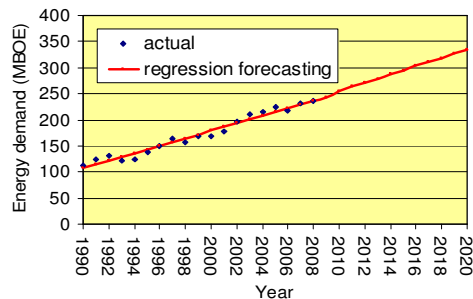regression model, and the experimental original data are plotted in Fig. 4.

Fig.4: The fitted and predicted values by regression model for the increase in energy demand of industrial sector

5. Comparison of Forecast Precision between GM, MCGM and Regression Model

As the above, the forecast values from 1979 to 2008 calculated by GM, MCGM and regression model. The forecast values of test data 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

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e^2(i) \tag{27}$$

$$AME = \frac{1}{n}\sum_{i=1}^{n} |e(i)| \tag{28}$$

$$AAEP = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{e(i)}{x(i)}\right| \times 100 \tag{29}$$

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

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

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 industrial sector by MCGM from 2011 to 2020 is given in Table 4.

Table 4 Predicted Value of Energy Demand of Iran's Industrial Sector by MCGM

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

In 2020, the energy demand of industrial sector will reach to a level of 335 MBOE. The accurate prediction will reasonably help in establishing energy supply and demand management policies and technological system plans of Iran.

10

Industrial sector is the most energy intensive sector of Iran, implying that for a given level of value added, the industry consumes the highest amount of energy. It should be mentioned there are high energy saving potentials and extensive managerial drivers for improving energy efficiency.

Conclusion

The major purpose of this paper was to develop the prediction model of energy demand of industrial sector in Iran. Through using the statistics data of the energy consumption of industrial sector from 1990 to 2008, three forecasting models presented and compared. The results showed that the accuracy of MCGM in forecast energy demand of industrial sector is higher than those of GM and regression model. Moreover, energy demand of Iran's industrial sector from 2011 to 2020 was forecasted.

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

References

[1] Akay, D. and M. Atak (2007). Grey Prediction with Rolling Mechanism for Electricity Demand Forecasting of Turkey. *Energy* 32(9): 1670-1675.

[2] Hsu, L.C. and C.H. Wang (2009). Forecasting the Output of Integrated Circuit Industry Using a Grey Model Improved by the Bayesian Analysis, *Technological Forecasting & Social Change* 36(4): 7898-7903.

[3] Huang, M., H. Yong and H. Cen (2007). Predictive Analysis on Electric Power Supply and Demand in China. *Renewable Energy* 32(7): 1165–1174.

[4] Institute for International Energy Studies (2010). Hydrocarbon Balance 2008.

[5] Li, G.D., D. Yamaguchi and M. Nagai (2006). Non-equidistance Grey Model Based on Grey Interval Weighting Accumulated Generating Operation. *Proc. Int Conf on Machine Learning, Models, Technologies Applications* 818–824.

[6] Li, G.D., D. Yamaguchi and M. Nagai (2007). A GM (1,1)–Markov Chain Combined Model with an Application to Predict the Number of Chinese International Airlines, *Technological Forecasting & Social Change* 74(8): 1465-1481.

[7] Zhang, SJ. And Y. He (2001). A Grey–Markov Forecasting Model for Forecasting the Total Power Requirement of Agricultural Machinery in Shangxi Province. J Shanxi Agric Univ (Nat Sci Edi) 21(3): 299–302.

پیش‌بینی تقاضای انرژی بخش صنعت در ایران با استفاده از مدل زنجیره مارکوف خاکستری

عالیه کاظمی[1]، محمد مدرس[2]، محمدرضا مهرگان[3]

هدف از این مقاله، توسعه مدل پیش‌بینی تقاضای انرژی بخش صنعت در ایران است. مدل زنجیره مارکوف خاکستری برای پیش‌بینی تقاضای انرژی این بخش پیشنهاد شده است. نتایج حاصل از پیش‌بینی با مدل مذکور با نتایج حاصل از پیش‌بینی با مدل خاکستری و مدل رگرسیون مورد مقایسه قرارگرفته است. این مقایسه نشان می‌دهد مدل پیش‌بینی زنجیره مارکوف خاکستری دارای دقت بالاتری نسبت به مدل‌های پیش بینی خاکستری و رگرسیون است. سپس تقاضای انرژی بخش صنعت در ایران تا سال ۲۰۲۰ با استفاده از مدل زنجیره مارکوف خاکستری پیش‌بینی شده است. نتایج حاصل، اساسی علمی برای توسعه برنامه‌های مربوط به عرضه انرژی در بخش صنعت در ایران فراهم می‌آورد.

واژگان کلیدی: تقاضای انرژی، بخش صنعت، پیش‌بینی، مدل خاکستری، مدل زنجیره مارکوف خاکستری، مدل رگرسیون.

۱. استادیار گروه مدیریت صنعتی دانشکده مدیریت دانشگاه تهران.

۲. استاد دانشکده مهندسی صنایع دانشگاه صنعتی شریف.

۳. استاد گروه مدیریت صنعتی دانشکده مدیریت دانشگاه تهران.

12