Modeling the Errors of EIA's Oil Prices and Production Forecasts By the Grey Markov Model

Hamidreza Mostafaei* and Shaghayegh Kordnoori

1. Department of Economics Energy, Institute for International Energy Studies (IIES), (Affiliated to Ministry of Petroleum)& Department of Mathematics, Faculty of Science, Shahid Rajaei Teacher Training University, Tehran, Iran.Email: ***hrmostafaei@yahoo.com***

2. M.Sc. Statistics of Tehran North Branch, Islamic Azad University & statistics expert of Research Institute for ICT,Tehran, ISLAMIC REPUBLIC OF IRAN. Email: ***sh_kordnourie@yahoo.com***

** Corresponding author: Institute for International Energy Studies (I.I.E.S) No.65, Sayeh St., Vali-e-asr Ave., Tehran 1967743711 I.R.Iran Tel: +982127644197 Fax: +982122054853*

ABSTRACT

Grey theory is about systematic analysis of limited information. The Grey-Markov model can improve the accuracy of forecast range in the random fluctuating data sequence. In this paper, we employed this model in energy system. The average errors of Energy Information Administrations predictions for world oil price and domestic crude oil production from 1982 to 2007 and 1985 to 2008 respectively were used as two forecasted examples. We showed that the proposed Grey-Markov model can improve the forecast accuracy of original Grey forecast model.

Key words: Grey theory, Grey Markov model, EIA, Oil

JEL Classifications: C15, C53, C63

1. INTRODUCTION

Probability and statistics, fuzzy mathematics, and Grey systems theory have been the three most-often applied theories and methods employed in studies of non-deterministic system [7].In recent years, researchers have paid close attention to the intelligent schemes, such fuzzy logic, neural network, genetic algorithm, and Grey system [9]. Grey system theory was founded at the beginning of 1980s. Grey theory deals with model uncertainty and information inadequacy in analysis systems by studying conditional analysis, prediction, control and decision making. The Grey prediction model is a qualified method for predictions with small data set; Moreover it has many applications in different fields.Y.Zhang andet al(2009)[13], applied a forecasting method combining Grey Model(GM) with partial least squares regression to Dalian city terminal energy consumption. Y.Hao and et al (2007) [12], used the Grey system model to simulate the long-term trend of economic development and obtained the periodic variation from analysis of residual model. H.Wu and F. Chen(2011)[5], proposed a GM(1,1) model for predicting short-term changes in Chinese exchange rate.

Markov Chain forecasting model is appropriate for systems with high random fluctuation, which can improve the accuracy of GM(1,1) model; therefore the Grey-Markov forecasting model holds the advantages of both Grey model and Markov transition probability matrix model.

H. Ma and D. Zhang(2009)[4], used the Grey- Markov model to forecast the coal production and consumption in chain. U. Kumar and V.K. Jain(2010), applied Grey Markov model to forecast the consumption of conventional energy in India and compared it with Grey model with rolling mechanism and singular spectrum analysis. Y. Zhang(2010)[10], used Grey Markov forecasting model for a highwayبtraffic volume.

Recently,energy systems were studied by many forecasting models. In this paper, we applied the Grey Markov model for the errors of Energy Information Administration (EIA) projections.The EIA is a creditable statistical agency for proficient competence, political independence. Several scientific analyses studied the precision of EIA projections. M. Auffhammer(2005)[8], studied test of rationality of published EIA forecasts under symmetric and asymmetric loss. D. Sakava (2005) [2], evaluated accuracy of annual energy outlook reports from 1982 to 2003 by error decomposition and regression. B. Neill and m. Desai(2005)[1], assessed and analyzed the accuracy of projections of US energy consumption produced by EIA over the period 1982-2000. About 35 percent of the worldبprimary energy consumption is supplied by oil[3]. In this paper, we introduce an improved GM(1,1) model which combined the Grey System and the Markov Chain model. Then we predict and analyze the errors of EIA predictions for world oil price and domestic crude oil production by the Grey-Markov forecasting model.

2. THE MATHEMATICAL METHOD

The procedure of our mathematical prediction model can be summarized as follows:

Suppose that $X^{(0)}(k) = \left\{X^{(0)}(1), X^{(0)}(2), \ldots, X^{(0)}(n)\right\}$ is the original data. A new sequence $X^{(1)}$ is set up through accumulated generating as follows:

$$X^{(1)}(k) = (\,X^{(1)}(1), X^{(1)}(2), \ldots, X^{(1)}(n)\,) \qquad (1)$$

Where

$$X^{(1)}(k) = \sum_{i=1}^{k} X^{(0)}(i)\ , k = 1,2, \ldots, n.$$

The Grey prediction GM(1,1) model can be expressed by one variable and first order differential equation

$$\frac{dX^{(1)}(k)}{dk} + aX^{(1)}(k) = b(2)$$

The solution of (2) is:

$$\hat{x}^{(1)}(k) = \left(x^{(\circ)}(1) - \frac{\hat{b}}{\hat{a}}\right) e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}} \qquad (3)$$

Where

$$\binom{a}{b} = (B^T B)^{-1} B^T Y_n \ (4)$$

And

$$B = \begin{pmatrix} -\frac{1}{2}\left(X^{(1)}(1) + X^{(1)}(2)\right) & 1 \\ -\frac{1}{2}\left(X^{(1)}(2) + X^{(1)}(3)\right) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}\left(X^{(1)}(n-1) + X^{(1)}(n)\right) & 1 \end{pmatrix} (5)$$

$$Y_n = \begin{bmatrix} X^{(\circ)}(2) \\ X^{(\circ)}(3) \\ \vdots \\ X^{(\circ)}(n) \end{bmatrix} (6)$$

By inverse accumulative generating operation, the predicted equation is:

$$\hat{X}^{(1)}(k) = \left(X^{(\circ)}(1) - \frac{b}{a}\right)(1 - e^{a})e^{-ak} (7)$$

Suppose

$$\hat{Y}(k) = \hat{X}^{(0)}(k+1)\ (8)$$

Since $\hat{Y}$ is a Markov Chain , we can divided it into several zones which are parallel to the regulation curve in accordance with particular circumstance and any zone H_i can be stated as :

$$H_i = [\hat{H}_{1i}, \hat{H}_{2i}] \quad i=1, 2, 3,..., n (9)$$

Where

$$\hat{H}_{1i} = \hat{X}^{(0)}(k+1) + A_i \quad i=1, 2,3,\ldots,n \qquad (10)$$

$$\hat{H}_{2i} = \hat{X}^{(0)}(k+1) + B_i \quad i=1,2,3,\ldots,n \qquad (11)$$

Here A_i,B_i are constant, which can be obtained by the difference between the raw data and forecasting regulation curve. We defined the borderlines of the zones above the regulation curve as upper borderlines and the ones under as lower borderlines. The upper and lower borderlines are assumed as $\hat{X}^{(0)}(k+1) + A$ and $\hat{X}^{(0)}(k+1) - B$, respectively. A and B are obtained by using the least square method as

$$A = \sum_H X^{(0)}(H+1) - \sum_H \hat{X}^{(0)}(H+1)/p \quad (12)$$

$$B = \sum_L X^{(0)}(L+1) - \sum_L \hat{X}^{(0)}(L+1)/q \quad (13)$$

Where $X^{(0)}(H+1)$ is the observed data above the regulation curve, p is the number of these data, $X^{(0)}(L+1)$ denotes the observed data below the regulation curve and q *is* the number of these lower data. Let $\hat{X}^{(0)}(k+1) + C$ and $\hat{X}^{(0)}(k+1) - D$ as the top and bottom borderlines, respectively where

$$C=max\ \{ X^{(0)}(k+1)\text{-}\hat{X}^{(0)}(k+1) \} \quad (14)$$

$$D=max\ \{\hat{X}^{(0)}(k+1)\text{-}X^{(0)}(k+1)\} \quad (15)$$

By getting *A*, *B*, *C*, and *D*, we have four zones as follows:

$$H_1 = [\hat{X}^{(0)}(k+1)+\mathrm{A}, \hat{X}^{(0)}(k+1) + C]$$

$$H_2 = [\hat{X}^{(0)}(k+1), \hat{X}^{(0)}(k+1)+\mathrm{A}]$$

$$H_3 = [\hat{X}^{(0)}(k+1) - B, \hat{X}^{(0)}(k+1)]$$

$$H_4 = [\hat{X}^{(0)}(k+1) - D, \hat{X}^{(0)}(k+1) - B] (16)$$

More subzones can be divided in each zone above with the same method.

If $M_{ij}(m)$ is the data number of raw series which transfer *m* step from H_i to H_j and M_i is the number of data that is in the zone H_i, then we call

$$p_{ij} = \frac{M_{ij}(m)}{M_i} \quad i,j=1,2,3,\ldots,n \qquad (17)$$

the *m*th step transitionprobability. The transition matrix *P (m)* is as follow:

$$P(m)=\begin{bmatrix} p_{11}(m) & p_{12}(m) & \cdots & p_{1n}(m) \\ p_{21}(m) & p_{22}(m) & \cdots & p_{2n}(m) \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1}(m) & p_{n2}(m) & \cdots & p_{nn}(m) \end{bmatrix} \quad (18)$$

P(m) reflects the transition regulation between different states and is the foundation of the forecast model of Grey Markov. We can predict the future trend of the systems by studying the stochastic transition matrix *P(m).* If *P(1)* has more than two lines whose probability values are same alike or close to other and it is difficult to decide the next direction of the system with certain, it is needed to study and check the matrix *P(2)* or $P(m)(m \geq 3)$. At the same time, it can be decide the transition of the system by checking P*(1)* or $P(m)(m \geq 2)$.At last the eventual forecast value can be obtained as

$$\hat{Y}'(k) = \frac{1}{2}(\hat{H}_{1l} + \hat{H}_{2l}) \qquad (19)$$

Applying (10), (11) and since the forecast is most probably in zone H_l, then $\hat{Y}'(k)$ can be expressed as

$$\hat{Y}'(k) = \hat{X}^{(0)}(k+1) + \frac{1}{2}(A_l + B_l) \qquad (20)$$

3. APPLICATIONS

The office of integrated analysis and forecasting of the Energy Information Administration (EIA) has produced annual evaluation of accuracy of annual energy outlook (AEO) since 1996 [3]. The projections in the AEO are not statements of what will happen but of what might happen, given the assumptions and methodologies [3].This paper deals with AEOبی oil price and production projections. Figure 1 shows the EIAs oil price forecast history.

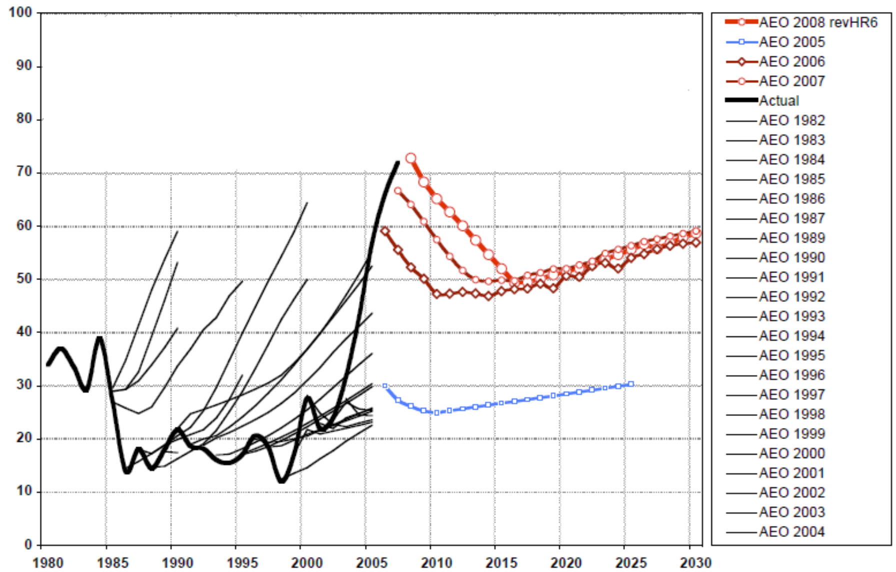

Figure 1.EIA crude oil price forecast history (1982-2008) [14].

The time series of average absolute error of world oil prices and domestic crude oil production of the AEOبی projections show random fluctuation. In this section we predict and analyze these projections errors by the Grey Markov forecasting model. The data of world oil price errors from 1989 to 2007 and domestic crude oil errors from 1985 to 2008 are listed in table 1 and 2, respectively[15].

Year	1982	1983	1984	1985	1986	1987	1988
Oil price error	25.54	28.65	20.66	16.81	16.67	7.59	11.64
Year	1989	1990	1991	1992	1993	1994	1995
Oil price error	10.55	11.69	10.64	9.59	9.75	11.84	12.84
Year	1996	1997	1998	1999	2000	2001	2002
Oil price error	15.84	17.29	20.99	21.12	22.51	23.06	25.53
Year	2003	2004	2005	2006	2007		
Oil price error	29.81	28.85	12.84	10.10	11.17		

Table 1. Average absolute errors of world oil price of EIAبـprojections.

Year	1985	1986	1987	1988	1989	1990
Oil production error	0.16	0.12	0.34	0.33	0.59	0.50
Year	1991	1992	1993	1994	1995	1996
Oil production error	0.24	0.17	0.17	0.19	0.34	0.29
Year	1997	1998	1999	2000	2001	2002
Oil production error	0.40	0.36	0.27	0.25	0.29	0.31
Year	2003	2004	2005	2006	2007	2008
Oil production error	0.28	0.28	0.37	0.44	0.47	0.53

Table 2. Average absolute residual errors of domestic crude oil production of EIAبـprojections.

We predict the world oil price error of 2008 and domestic crude oil production error of 2009 by applying the Grey-Markov model. According to our method we obtained:

World oil price error: $\hat{X}^{(0)}(k+1)$ =14.2972$e^{0.011741k}$

Domestic crude oil production error: $\hat{X}^{(0)}(k+1)$ =0.2653$e^{0.016869k}$

By (12) to (15), it follows that

World oil price error: *A*=6.3161, *B*=5.4385, *C*=14.18 and *D*=8.85

Domestic crud oil production error: *A*=0.107, *B*=0.0821, *C*=0.31 and *D*=0.15

Therefore four zones are compartmentalized as follows:

World oil price error:

$$H_1= [\hat{X}^{(0)}(k+1)+6.3161, \hat{X}^{(0)}(k+1)+14.18]$$

$$H_2= [\hat{X}^{(0)}(k+1), \hat{X}^{(0)}(k+1)+6.3161]$$

$$H_3=[\hat{X}^{(0)}(k+1)-5.4385\,, \hat{X}^{(0)}(k+1)]$$

$$H_4=[\hat{X}^{(0)}(k+1)-8.85\,, \hat{X}^{(0)}(k+1)-5.4385]$$

Domestic crud oil production error:

$$H_1= [\hat{X}^{(0)}(k+1)+0.107, \hat{X}^{(0)}(k+1)+0.31]$$

$$H_2= [\hat{X}^{(0)}(k+1), \hat{X}^{(0)}(k+1)+0.107]$$

$$H_3=[\hat{X}^{(0)}(k+1)-0.0821\,, \hat{X}^{(0)}(k+1)]$$

$$H_4=[\hat{X}^{(0)}(k+1)-0.15\,, \hat{X}^{(0)}(k+1)-0.0821]$$

Figure 2 shows these four zonesH_1,H_2,H_3, H_4 from the top down and their borderlines for the crude oil production. We obtained for these errors that M_1=5,M_2=8,M_3=7 and M_4=5 and the number of the raw data from H_4 to H_1, H_2, H_3 and H_4 respectively by one step is 0,0,2,3. Thus from (17) and with the same way the transition probability matrix of one step is as follow:

$$P(1)=\begin{bmatrix} 3/5 & 1/5 & 0 & 1/5 \\ 1/8 & 3/4 & 0 & 1/8 \\ 0 & 1/7 & 5/7 & 1/7 \\ 0 & 0 & 2/5 & 3/5 \end{bmatrix}$$

From Fig.2, we can see that the error of world oil price of 2007 is in H_4 , so we examine the fourth line of *P* (1). The maximum probability of this line is p_{44} . Hence the next state of the system may transfer from H_4 to H_4 . We predict that the error of world oil price of 2008 is probably inH_4. At last the forecast error of 2008 can be obtained as follow:

$$\hat{Y}'(26)=\frac{1}{2}(\widehat{H}_{14}+\widehat{H}_{24})= \hat{X}^{(0)}(27)-\frac{1}{2}(\text{ B+D}) =12.257$$

Table 3 shows the forecast value and the precision with GM(1,1) model and the Grey Markov model. By comparing the results we can deduce that the forecast value of the Grey Markov forecasting model is more precise and reliable than the original GM (1, 1).

Figure 3 shows the four zonesH_1,H_2,H_3, H_4 from the top down and their borderlines for error of crude oil production. We obtained for these errors that M_1=2,M_2=7,M_3=7 and M_4=7 and the raw number of data from H_1 to H_1, H_2, H_3 and $H_4$by one step is 1,0,1,0, respectively . Therefore, the transition probability matrixes of one and two step are:

$$P\ (1) =\begin{bmatrix} 1/2 & 0 & 1/2 & 0 \\ 2/7 & 3/7 & 2/7 & 0 \\ 0 & 2/7 & 2/7 & 3/7 \\ 0 & 2/7 & 2/7 & 3/7 \end{bmatrix} \text{ and } P\ (2) =\begin{bmatrix} 1/4 & 1/7 & 11/28 & 3/14 \\ 13/49 & 13/49 & 17/49 & 6/49 \\ 4/49 & 16/49 & 2/7 & 15/49 \\ 4/49 & 16/49 & 2/7 & 15/49 \end{bmatrix}$$

From Fig.3, we see that the error of oil production projection of 2008 is in H_1 , so we investigate the first line of $P(1)$ and since p_{11} and p_{13} are equal, we examine the $P(2)$ and see that p_{13} is the maximum probability. Therefore the next state of the system may transfer from H_1 to H_3 . The error of crude oil production of 2009 can be obtained as follow:

$$\hat{Y}'(24)=\frac{1}{2}(\widehat{H}_{13}+\widehat{H}_{23}) = \hat{X}^{(0)}(25) - \frac{1}{2}\text{B} =0.3567$$

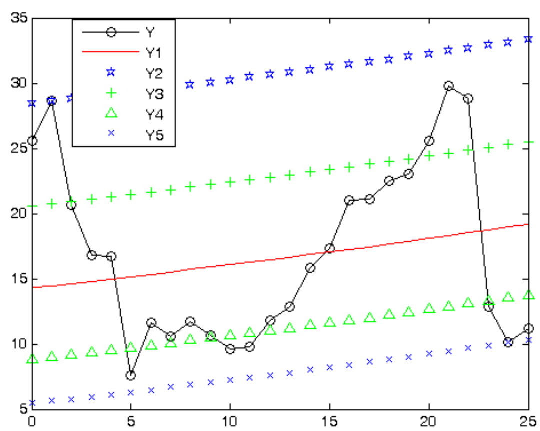

Fig.2 Forecast curve of the errors of world oil price predictions for EIA during 1982 to 2007

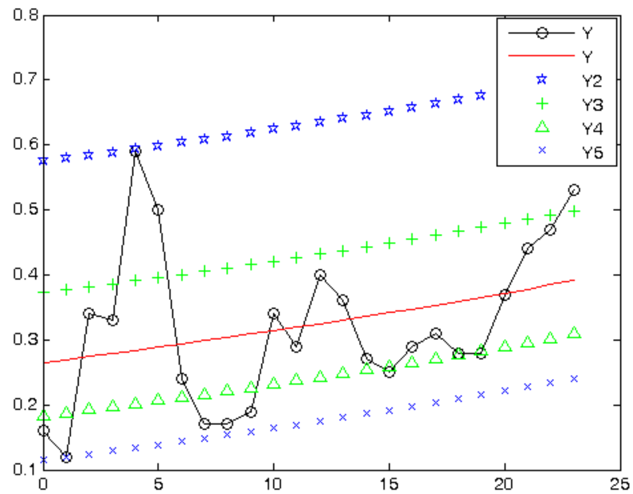

Fig. 3. Forecast curve of the errors of domestic crude oil productions for EIA during 1985 to 2008.

We give the forecasted values and their precisions by the GM(1,1) and Grey Markov model in table 3. Obviously we conclude that the Grey Markov model give more precise predictions and improve the GM(1,1) forecasts.

Year	Actual average absolute error of world oil prices for EIA	GM(1,1) model		Grey-Markov model	
		Forecast value	precision	Forecast value	precision
2008	12.50	19.40	44.79%	12.26	98.06%
Year	Actual average absolute residual of domestic crude oil production for EIA				
2009	0.32	0.40	75.70%	0.36	88.34%

Table 3. The comparison between GM(1,1) and Grey-Markov method.

4. CONCLUSION

The model GM(1,1) can be used to forecast the change trend of data sequences, while the Markov model can be applied to decide the vibration regulation of their development , and both can be joined together to become a Grey-Markov forecast model. We applied this model for the average errors of EIAبprojections. By comparing the GM(1,1) and Grey Markov model results, we concluded that the Grey-Markov model give higher and reliable forecastings.

6. REFERENCES

[1]Brain C.O, Neill, MavSami Desai, Accuracy of Past Projections of US Energy Consumption, Energy Policy 33, pp.979-993, 2005.
[2]Denys Sakva, Evaluation of errors in national energy forecasts, Rochester York, A Thesis Submitted to the public policy program of Rochester Institute of Technology,2005.

[3]Energy Information Administration, Annual Energy Outlook 2009, retrospective review: Evaluation of reference case projections in past edition (1982-2009), 2009.
[4]Hong-Wei Ma, Dong-Qing Zhang, A Grey forecasting model for Coal production and consumption, proceeding of 2009 IEEE International Conference on Grey System and Intelligent Services, November 10-12,China ,2009.
[5]Hong Wu , Fuzhong Chen, The Application of Grey System to exchange rate prediction in the past-crisis era, International Journal of Innovative Management Information & Production, Vol.1,No.2, June 2011.
[6]International Energy Agency, World Energy Outlook 2006, Paris, 2006.
[7]J.L. Deng, Introduction to grey system theory, the journal of grey system, Vol.1, No.1, pp.1-24, 1989.
[8]Maximilian Auffhammer, The Rationality of EIA forecasts under symmetric and asymmetric loss, 2005.
[9]Sifeng Liu and Yilin, Grey information, theory and practical applications, Springer,2006.
[10]Ujjwal Kumar, V.K. Jain, Time Series models (Grey-Markov, Grey Model with rolling mechanism and Singular Spectrum Analysis) to forecast energy Consumption in India, Energy 35, 1709-1716, 2010.
[11]Yin peng Zhang, Predicting Model of Traffic Volume based on Grey Markov, Modern Applied Science, Vol.4, No.3, March 2010.
[12]Yonghong Hao, JuanDu,Wei Wang, Xuemeng Wang, A Grey System Model for Simulation of Economic Development, proceeding of 2007 IEEE, International Conference on Grey Systems and Intelligent Services, November 18-20, Nanjing chain,2007.
[13]Yon Zhang, Jing Zhang, Xiaomei Zhang,Shushen Zhang,Suling Liu,Yu Chen, Establishment of Combining Grey Model with partial least squares regression for city Energy Consumption forecasting, IEEE,2009.
[14] www.eia.doe.gov/olaf/aeo/index.htm, Bill Hudson ,A Review of EIAبیAnnual Energy Outlook 2008 ,Revised to Reflect ,Energy Independence & Security Act of 2007,
[15]Energy Information Administration, Annual Energy Review 2009, DOE/EIA-0384(2009) Washington, DC, July 2010.