The Application of a Grey Markov Model in Forecasting the Errors of EIA's Projections in Gas Production and Energy Intensity

Seyed Hossein Iranmanesh, PhD
Department of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran
hiranmanesh@ut.ac.ir

Hamidreza Mostafaei, PhD
Department of statistics Islamic Azad University, North Tehran Branch &
Department of Economics Energy, Institute for International Energy Studies (Affiliated to Ministry of Petroleum)
h_mostafaei@iau-tnb.ac.ir

Shaghayegh Kordnoori, PhD candidate
Statistics expert of Research Institute for ICT, Tehran, Iran
sh_kordnourie@yahoo.com

Abstract: G*rey system theory looks for realistic patterns based on modeling with a few available data. In this paper, a Grey-Markov prediction model which is the combination of the GM(1,1) and Markov model was studied; Moreover, its applications in energy system were presented. The average errors of Energy Information Administration's predictions for Natural Gas production and Energy intensity from 1985 to 2008 and 1985 to 2007 respectively were used as two forecasted examples. Comparing with GM(1,1) prediction model, we showed that the Grey- Markov prediction model improves the forecast accuracy.*

Key words: Grey Theory, Grey-Markov, EIA, Gas Production, Energy Intensity

JEL Classification: C15,C53

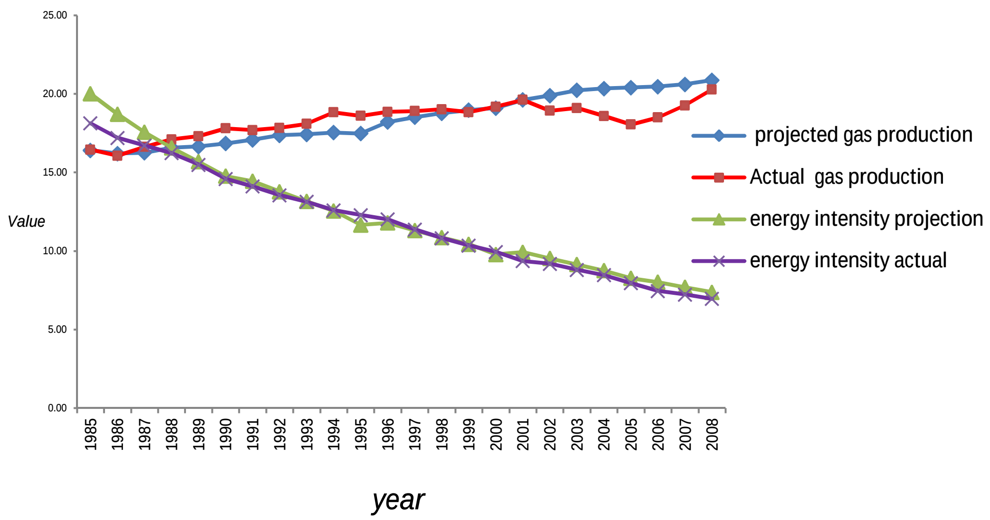
1. Introduction

In prediction field, predictions should be accurate and reliable. Forecasting methods are composed of quality and quantitative ones. Qualitive forecasting methods have Delphi method, trend method, prediction method, market research and the expert system etc. while the quantitative forecasting methods consist of Grey model (GM), neural networks, time series, econometric models, regression models, statistical methods and casual model etc. Finding reliable patterns according to modeling by poor information is the main function of GM. The advantage of the Grey models is their ability in evaluating the performance of unexplored systems with a restricted amount of data which is at least only 4 but the regression method, simple exponential and box Jenkins require at least 10 to 20, 5 to 10 and 50 data, respectively. The Grey model can be applied for both equal and not equal gaps while simple exponential method and box Jenkins only be applied for equal gaps, regression model for same and regular trend data and time series method for mixture and regular data. Another advantage of the GM is its usage in short, mid and long term predictions while the regression model and time series can be used only for short and middle term predictions. GM is the core of Grey system theory. Random variables in the Grey theory are considered as inconstant numbers that change with time parameters; The level of uncertainty is denoted by "color" in this model. The order series is transformed to the differential equation in the GM. The application of the Grey models have been very broad: In predicting gross national product, inflation index and stock price (Kotil et al.,2005; Ma and Zhand,2009), business failure(Cheng and Chen,2009),energy issues(Lin & Yang ,2003; Li, 2006; Zheng & Wang, 2011; Ma and Zhand,2009; Kordnoori & Mostafaei,2011;Kumar & Jain,2011), voyaging systems and networks (Zhang et al,2008;Xue et al.,2011), output values of industries (Ficherr et al.,2008).

A probability model which is a sequence of independent random variables is a Markov chain. It's advantage is the ability of modeling the uncertainties in systems which fluctuate dynamically in time. The transition probabilities demonstrate the importance of all random aspects. In Grey-Markov model, the Markov model is applied to predict the stochastic alternations and the GM for predicting the trend of data sequence. A stationary and non-stationary time sequence can be forecasted by the Grey-Markov model.

Energy is one of the indispensable factors for continuous development and economic growth; Moreover Energy predictions have a fundamental influence on improving energy and environmental plans. The growth and development of the economic activities depend on energy fore castings. Therefore it is vital for governments to plan according to the reliable forecasts. There are some organizations which regularly predict energy statistics. Their predictions are applied in budget forecasting, energy planning's, investing procedures, political decisions, economic activities and etc. The most important agencies are EIA, BP, IEA, OPEC and Exxon-Mobil which publish and forecast energy data. As these energy forecasts play important roles in world economic, politics, governmental policies, it is important to evaluate the accuracy of these predictions. By choosing and comparing the actual and forecasted values of EIA errors, we recognize that some errors exist in their predictions. As seen in Figure 1, there are some errors in predicting and actual values of EIA. We display these errors only for natural gas production and energy intensity in this figure as they are the cases which we analyze in this research. In all energy forecasting these errors exist. Therefore modeling and predicting these errors are essential.

Fig.1: The differences between actual and EIA's projection values for natural gas production and energy intensity



-Unit of Gas production is trillion cubic feet;unit of Energy intensity is quadrillion Btu/$ Billion Nominal(GDP)

The researchers pay much attention to various energy predicting models in recent years. There are some related works which analyzed the accuracy and the errors of energy forecasts (Soldo, 2012 ; Sanders et al., 2008 ; Joutz & Trost , 1992; Huntington, 1994; Linderoth, 2002; Craige et al.,2002; Smil,2000;Asher,1978).More specifically, Several independent analyses have been published over the past years which examine the accuracy of Energy Information Administration (EIA)'s projections. Neil and Desai analyzed and assessed the EIA projections of U.S. energy consumption (Neill & Desai, 2005; Neill & Desai,2003). Fischer et al.(2008) investigated the potential for systematic errors in the EIA's widely used annual energy outlook. Valex Lekat and Larry Dale (2005) evaluated the accuracy of AEO's forecasted price and Henry Hub compared to U.S. wellhead future price. Winebrake and Sakva (2006) explored U.S. energy forecasts through the EIA and its predecessors in order to uncover potential systematic errors in U.S. forecasting model. Test of rationality of EIA's projections was studied by Auffhammer (2005).Shlyakhter and et al.(1994) found the distribution of error of EIA's projections. Cohen and et al.(1995) studied the EIA's projections.

In this paper first we introduce the GM(1,1) and Grey-Markov model. In section 3 we apply these methods in modeling the errors of EIA's projections for natural gas production and energy intensity. We compare the results of these models and forecast the errors of EIA's projections for future. In section 4 we end this paper with our conclusions and future works.

2. The Mathematical Method

In this part we review the procedure of our mathematical prediction model. Suppose we have $X^{(0)}(k) = \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)\}$ as the initial data sequence. A new series $X^{(1)}$ is set up through accumulated generating as follows:

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

Where

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

The first-order difference equation of GM(1,1) is defined as:

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

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

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

and

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

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

By inverse accumulative generating operation, the predicted equation is:

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

Suppose

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

We consider the states of a Markov chain $\hat{Y}$ which are alongside the regulation curve as follows:

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

Where

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

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

where A_i and B_i are the differences between the original data and predicting curve. The top 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 $X^{(0)}(H+1)$ and $X^{(0)}(L+1)$ are the data above and below the forecasting curve and p ,q correspond to the number of such data respectively. Let $\hat{X}^{(0)}(k+1) + C$ and $\hat{X}^{(0)}(k+1) - D$ as the top and bottom borderlines, respectively where

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

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

The states of our Markov model are obtained as follows:

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

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

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

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

Similarly, each zone can be classified into more subzones.

A Markov chain $\{X_n; n \geq 0\}$ is a stochastic process with the property that for all *I,j,k,l* in state space

$$p_{ij} = P[x_{n+1} = j \urcorner X_n = i] = P[x_{n+1} = j \urcorner X_n = i, X_{n-1} = k \ldots X_0 = l] \quad (17)$$

The transition probabilities p_{ij} for all ***i*** ,***j*** in state space satisfy in:

$$p_{ij} \geq 0, \sum_{j=o}^{m} p_{ij} = 1 \quad (18)$$

The future trend of systems can be forecasted by the transition probability matrix. We can get this matrix in ***m*** th step as :

$$P(m)=\begin{bmatrix} p_{11}(m) & p_{12}(m) & \cdots & p_{1n}(m) \\ p_{21}(m) & p_{22}(m) & \cdots & p_{2n}(m) \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1}(m) & p_{n2}(m) & \cdots & p_{nn}(m) \end{bmatrix} \quad (19)$$

Where

$$p_{ij} = \frac{M_{ij}(m)}{M_i} \ \ i,j=1,2,3,\ldots,n \quad (20)$$

When we cannot certainly distinguish the next path of the system, the matrix *P(2)* or *P(m)*$(m \geq 3)$ must be calculated. At last the final predicted value can be obtained as

$$\hat{Y}'(k) = \frac{1}{2}(\hat{H}_{1l} + \hat{H}_{2l}) \quad (21)$$

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

$$\hat{Y}'(k) = \hat{X}^{(0)}(k+1) + \frac{1}{2}(A_l + B_l) \qquad (22)$$

For evaluating the accuracy of our model we compute the relative percentage error by

$$RPE = \frac{\left|x^{(0)}(k) - \hat{x}^{(0)}(k)\right|}{x^{(0)}(k)} \qquad (23)$$

and whence find the precision by *(1-RPE)×100* . If the precision is more than 90% we can conclude that the model is reliable and accurate.

3. Applications

A superior statistical analytical organization which presents helpful energy information is Energy Information Administration (EIA). Developing history and activities of EIA was studied (Kent,1993).This paper deals with AEO's natural Gas production and energy intensity projection errors. Natural gas is one of the most abundant energy sources in the world and a major energy source of industrial and electrical section; EIA forecasted that the greatest increase in gas production up to 2035 is for Middle East; Furthermore, energy intensity is used an energy conservation index for a country. It is calculated as units of energy per unit of GDP. The average absolute differences between the AEO's projections and actual for natural Gas production and energy intensity from 1985 to 2008 and 1985 to 2007 (table1 and 2)(Department of Energy Washington,2010), are applied here. By applying the Grey - Markov predicting model the projection errors of EIA are modeled and forecasted.

Table 1: Average absolute differences (errors) of EIA's natural Gas production projections

Year	1985	1986	1987	1988	1989	1990
Gas production error	0.82	0.86	0.80	0.73	0.73	0.96
Year	1991	1992	1993	1994	1995	1996
Gas production error	0.61	0.73	0.70	1.28	1.20	0.77
Year	1997	1998	1999	2000	2001	2002
Gas production error	0.69	0.63	0.78	0.82	0.52	1.03
Year	2003	2004	2005	2006	2007	2008
Gas production error	1.16	1.75	2.35	2.01	1.53	1.23

Table 2: Average absolute difference errors of EIA energy intensity projection

Year	1985	1986	1987	1988	1989	1990
Energy intensity	1.87	1.51	0.85	0.54	0.57	0.65
Year	1991	1992	1993	1994	1995	1996
Energy intensity	0.47	0.59	0.71	0.81	1.08	0.61
Year	1997	1998	1999	2000	2001	2002
Energy intensity	0.68	0.72	0.75	0.83	0.58	0.45
Year	2003	2004	2005	2006	2007	
Energy intensity	0.46	0.46	0.60	0.64	0.59	

We forecast the natural Gas production error of 2009 and energy intensity error of 2008 by the Grey Markov model. According to our method we obtained:

Natural Gas production error: $\hat{X}^{(0)}(k+1) = 0.5615\, e^{0.046024k}$

Energy intensity error: $\hat{X}^{(0)}(k+1) = 0.8902\; e^{-0.023171k}$

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

Natural Gas production error: *A*=0.3373, *B*=0.2392, *C*=0.94 and *D*=0.65

Energy intensity error: *A*=0.2208, *B*=0.1518, *C*=9798 and *D*=0.3046

Therefore four zones are compartmentalized as follows:

Natural Gas production error:

$$H_1= [\hat{X}^{(0)}(k + 1)+0.3373,\hat{X}^{(0)}(k + 1) + 0.94]$$
$$H_2= [\hat{X}^{(0)}(k + 1),\hat{X}^{(0)}(k + 1)+0.3373]$$
$$H_3=[\hat{X}^{(0)}(k + 1) - 0.2392 ,\hat{X}^{(0)}(k + 1)]$$
$$H_4=[\hat{X}^{(0)}(k + 1) - 0.65 , \hat{X}^{(0)}(k + 1) - 0.2392]$$

Energy intensity error:

$$H_1= [\hat{X}^{(0)}(k + 1)+0.2208,\hat{X}^{(0)}(k + 1) + 0.9798]$$
$$H_2= [\hat{X}^{(0)}(k + 1),\hat{X}^{(0)}(k + 1)+0.2208]$$
$$H_3=[\hat{X}^{(0)}(k + 1) - 0.1515 ,\hat{X}^{(0)}(k + 1)]$$
$$H_4=[\hat{X}^{(0)}(k + 1) - 0.3046 , \hat{X}^{(0)}(k + 1) - 0.1518]$$

Figure 2 shows these four zones H_1,H_2,H_3, H_4 from the top down and their border lines for the natural Gas production. We find that for these errors M_1=4, M_2=7, M_3=7 and M_4=5 and the number of the original data by one step from H_4 to H_1,H_2,H_3 and H_4 respectively is 0,0,1 and 5. Therefore, the one step transition probability matrix is calculated as:

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

All the fluctuation and transitions of gas production error forecasting's can be seen in figure 2. By recognizing the next state from this representation and find the maximum probability of transitions of that state we can forecast the next EIA's prediction error of gas production. So From this figure, we can see that the error of natural Gas production of 2008 is in H_4.Therefore By examining the fourth line of $P(1)$ we realize that p_{44} is the maximum probability, so the most probable state which the system may transfer to is from $H_4$to H_4. Finally, the error of EIA's natural gas production projection for 2009 can be obtained as follow:

$$\hat{Y}'(24)=\frac{1}{2}(\hat{H}_{14} + \hat{H}_{24})= \hat{X}^{(0)}(25)-\frac{1}{2}(\text{ B+D}) =1.25$$

Figure 3 show the four zones H_1,H_2,H_3, H_4 from the top down and their borderlines for error of energy intensity. We conclude that for these errors M_1=3,M_2=8,M_3=7 and M_4=4 and the raw number of data from H_2 to H_1,H_2,H_3 and H_4 by one step is 1,5,1,1, respectively. Hence, the one step transition probability matrix is:

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

The upper and lower borderlines and transitions of energy intensity are shown in this figure. Therefore, We realize that the error of energy intensity of 2008 is in H_2, As a result we examine the second line of $P(1)$ and see that p_{22} is the maximum probability. Hence most likely the state which the system may transfer to is from H_2 to H_2. The error of EIA's energy intensity projection for 2008 can be obtained as follow:

$$\hat{Y}'(23)=\frac{1}{2}(\hat{H}_{12} + \hat{H}_{22})= \hat{X}^{(0)}(24) + \frac{1}{2}(\text{ A}) =0.6328$$

Fig. 2: Four zones and forecasting regulation curve of EIA's error projections for Natural Gas production during 1985 to 2008

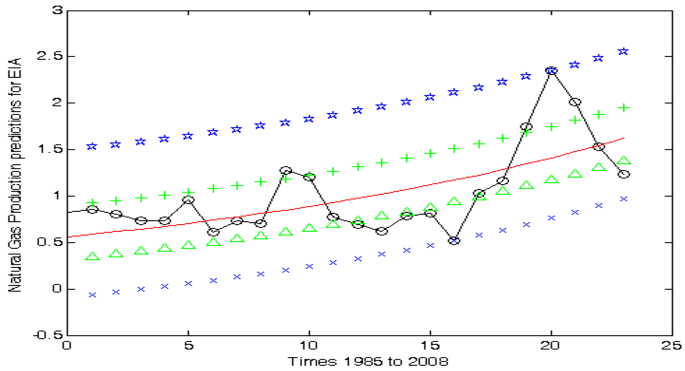


Fig.3: Four zones and forecasting regulation curve of EIA's error projections for energy intensity during 1985 to 2007.

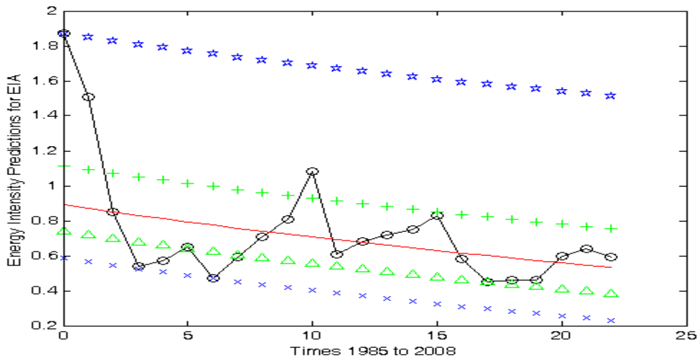


Table 3 shows the forecast value and the precision of EIA's forecasting errors by GM(1,1) and the Grey-Markov model. By comparing the results we conclude that the predicting values obtained by Grey Markov model are more accurate than GM(1,1). All the obtained precisions by the Grey markov model are more than 90% and prove the efficiency of our model.

Table 3:**The resulting forecasts and precisions by GM(1,1) and Grey-Markov model.**

Year	Actual average absolute difference (error) of natural Gas production of EIA	**GM(1,1)**		**Grey-Markov model**	
		Forecast value	**precision**	**Forecast value**	**precision**
2009	1.38	1.69	77.20%	1.25	90.58%
Year	Actual average absolute difference (error) of energy intensity of EIA				
2008	0.60	0.52	87.07%	0.63	94.53%

As a result, the error of all EIA's projections in future for gas production and energy intensity can be obtained by Grey Markov model with high precisions.

4. Conclusion

Some organizations give forecasting energy information. It is important to recognize their prediction errors; moreover, these errors are fluctuated over time. Therefore a Grey-Markov model which is the combination model of GM and Markov chain is suitable for modeling these errors. We applied the GM(1,1) and Grey-Markov model for the absolute average error of EIAs projections for natural Gas production and energy intensity. By comparing the results of these methods, we obtained that the Grey Markov model gives more accurate and certain projections. In future, researchers can model and forecast the errors of other energy organizations such as IEA, BP, Exxon-Mobil and OPEC by the Grey-Markov model and compare the results to realize that which agency's predictions has a lower error and are more realistic; Moreover, the Grey Markov model can be compared to other modified Grey models.

5. Bibliography

[1] B. Neill, M. Desai (2005) - "Accuracy of past projections of us energy consumption", Energy Policy 33: 979-993

[2] B. Neill, M. Desai (2003) - "The accuracy of past projections of U.S. energy consumption", interim report

[3] B. Soldo (2012) - "Forecasting natural gas consumption", Applied Energy, ,92: 26-37

[4] C. Fischerr, E. Herrnstadt, and R. Morgenstern (2008) - "Understanding error in EIA projections of energy demand"

[5] Ch. Lin, Sh. Yang (2003) - "Forecast of the output value of Taiwan's opto-electronics industry using the Grey forecasting model", Technological Forecasting and Social Change 70: 177-186

[6] C. Kent, (1993) - "An analytical history of EIA", Government Information Quarterly, 10(1): 3-24

[7] C. Li (2006) - "Study on theory of the Grey Markov Chain method and its application", IMACS multi conference on "Computational engineering in systems applications" CESA , China

[8]Cohen Barry, G. Peabody, M. Rodekohr, S. Shaw - "A history of midterm energy projections: a review of energy outlook projections. Washington D.C.: Energy Information Administration, U.S. Department of energy.

[9] D.R. Sanders, M.R. Manfredo, K. Boris (2008) - "Accuracy and efficiency in the U.S. department of energy's short term supply forecasts, Energy economics, 30(3): 1192-1207.

[10] E. Kotil, SaclikCukut, ResulERY, GIT. (2005) - "A Grey model study of forecasting selected Turkish macro economic variables", Economic Yaklasim, city:16, sayi:s6: 34-42.

[11] F. Joutz, R. Trost.(1992), " Using stochastic simulation to test the effect of seasonal adjustment on forecast standard errors of motor gasoline demand", International journal of forecasting, 8(2):219-231.

[12] G. Zheng and T. Wang.(2011) , "A new model for information fusion based on Grey theory , information technology journal ", Gabrielle wong-parodi, 10(1):189-194.

[13] H. Ma and Z .Zhang.(2009), Grey prediction with markov chain for crude oil production and consumption in China , springer , AISc,56: 551-561.

[14] H. Ma , D. Zhang.(2009), "A Grey forecasting model for coal production and consumption", proceedings of 2009 IEEE International conference on Grey systems and intelligent services, China.

[15] H. Huntington.(1994), "Oil price forecasting in the 1980s: What went wrong?", The energy journal, 15(2): 1-22.

[16] H. Linderoth.(2002), "Forecast errors in IEA countries' energy consumption", energy policy, 30:53-61.

[17] J. J. Winebrake , D. Sakva.(2006)," an evaluation of errors in U.S. energy forecasts" : energy policy,vol. 34 (18):3475-3483.

[18] J. Cheng , H. Chen , K. Cheng.(2009), "Business failure prediction model based on Grey prediction and rough set", Wseas Transactions on information science and applications,6(2).

[19] L. Zhang ,Y. Lu, L. Chen, D. Deng.(2008), "Evaluating strategy with Grey theory utilized in wireless sensor networks" , proceeding of the world congress on engineering , vol I.

[20] M. Auffhammer.(2005)," The Rationality of EIA forecasts under symmetric and asymmetric loss", University of California Berekeley.

[21] P. Craig, A. Gadgil, J. Koomy.(2002), "what can history teach us? A retrospective examination of long term energy forecasts for the United States", The annual review of energyand environment,27: 83-118.

[22] Retrospective Review Annual Energy Outlook 2010 U.S. Department of Energy Washington, DC 20585 .

[23] Sh. Kordnoori, H. Mostafaei.(2011), " Grey Markov model for predicting the crude oil production and export in Iran" , international journal of academic research,3(2) .

[24]Shlyakhter,A.I.,D.M.Kammen,C.L.Broido,WILSON(1994),"Quantifying credibility of energy projections from trends in past data", energy policy, 2:119-130.

[25] U. Kumar, V.K. Jain. (2010), "Time series models(Grey markov, Grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India": 1709-1716.

[26] V. Lekat and L. Dale.(2005), "Natural gas prices forecast comparsion AEO U.S. Natural gas markets".

[27] V. Smil.(2000), "Perils of long range energy forecasting: reflection on looking far ahead", Technological forecasting and social change,65: 251-264.

[28] W. Ascher.(1978), "Forecasting: An appraisal for policy makers and planners. Baltimore: John Hopkins university press.

[29] Y. Xue, Z. Cao and L. Xu.(2011), "The application of combination forecasting model in energy consumption system", energy proceida,5, 2566-2603.