# การพยากรณ์ราคาทองคำโดยการปรับปรุงตัวแบบ GM (1,1) ด้วยลูกโซ่มาร์คอฟโดยใช้ค่าเฉลี่ยของจุดกึ่งกลาง

# Gold Price Forecasting Based on the Improved GM(1,1) Model with Markov Chain by Average of Middle Points

*SoPheap Sous[1], Thotsaphon Thongjunthug[1] and Watcharin Klongdee[1*]*


## บทคัดย่อ

งานวิจัยนี้ทำการพยากรณ์ราคาทองคำด้วยตัวแบบ MCGM(1,1) ซึ่งเกิดจากการนำตัวแบบมาร์คอฟ และตัวแบบ GM(1,1) ประกอบเข้าด้วยกัน ซึ่งความน่าจะเป็นของการเปลี่ยนสถานะของลูกโซ่มาร์คอฟคำนวณได้จากความคลาดเคลื่อนระหว่างตัวแบบ GM(1,1) กับข้อมูลการสอนซึ่งเป็นข้อมูลของราคาทองคำจาก เดือนมกราคม พ.ศ. 2533 ถึง เดือนธันวาคม พ.ศ. 2554 หลังจากนั้นนำค่าพยากรณ์ที่ได้ไปทดสอบความแม่นยำกับข้อมูลการทดสอบ ซึ่งเป็นข้อมูลของราคาทองคำจาก เดือนมกราคม 2555 ถึง เดือนมิถุนายน 2557 ผลการวิจัยพบว่า ตัวแบบ MCGM(1,1) มีความแม่นยำสูงกว่า ตัวแบบ GM(1,1) โดยพิจารณาจาก 3 เกณฑ์ ดังนี้ ค่าคลาดเคลื่อนกำลังสองเฉลี่ย ค่าคลาดเคลื่อนสมบูรณ์เฉลี่ย และ ค่าคลาดเคลื่อนสัมพัทธ์สมบูรณ์



## ABSTRACT

In this paper, we combine Markov chain and GM(1,1) model, which will be called MCGM(1,1), to forecast the gold price in London. The transition probabilities of Markov chain are constructed by the error of GM(1,1) and original data from January 1990 to December 2011. We compare the accuracy of the prediction using the testing data from January 2012 to June 2014. We find that MCGM(1,1) is better than GM(1,1) in terms of three criterions, namely, mean square error, absolute mean error and absolute relative error.


**คำสำคัญ:** ตัวแบบเกรย์ ลูกโซ่มาร์คอฟ ราคาทองคำ การพยากรณ์

**Keywords:** Grey model, Markov chain, Gold price, Forecasting

[1]Risk and Insurance Research Group, Department of Mathematics, Faculty of Science, Khon Kaen University, Thailand 40002

[*]Corresponding Author, E-mail: kwatch@kku.ac.th

## 1. Introduction

Gold is a precious and the most popular metal for investment. Investors hold gold to hedge against economic and political uncertainly, or currency crisis. However, gold price is fluctuated daily or monthly. For example, the gold price increased rapidly in 2012, whereas it fell unexpectedly in 2013. The main interest from the point of view of an investor is the prediction or the forecast of the gold price for deciding whether to invest in the gold market. In 1982, Deng (1982) proposed the theory of grey systems whose partial information is known, and studied the uncertainty of such systems. Grey forecasting model (denoted by GM(1,1)) uses a first order differential equation of one variable to simulate an unknown information, which is suitable for forecasting in the competitive environment where decision makers can refer only to limited historical data. However, the forecasting precision for data sequences with large random fluctuation is low. On the other hand, Markov chain model can be applied to forecast the system with randomly varying time series by using the transition probabilities between states. Thus, there are many papers studying the effect of Markov chain to improve the accuracy for the forecasting. For example, Li et al. (2007) developed the prediction of a combined GM(1,1) and Markov chain model with an application to the prediction of the number of Chinese international airline. Hung et al. (2007) presented a grey Markov forecasting model and its analysis to forecast the electric power demand in China. Kazemi et al. (2011) developed the prediction model of energy demand of industry sector in Iran as grey-Markov chain model forecasting.

In this paper, we will consider the gold price from January 1990 to June 2014 and study the combination model of Markov chain and GM(1,1). The transition probabilities of Markov chain will be constructed and trained by the error of GM(1,1) and the original data from January 1990 to December 2011. Then we will compare the accuracy of the prediction by using the testing data from January 2012 to June 2014.

## 2. Research Methodology

### 2.1 GM(1,1) model

GM(1,1) is a type of grey model which is the most widely used in literature. The steps of GM(1,1) are shown as followings;

**Step 1:** Let $X^{(0)}$ be the original series with $n$ samples given by

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

The accumulating generating operator (AGO) is defined by

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

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)\,, \quad k = 1, 2, 3, \dots, n.$$

Obviously, we obtain that

$$x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1), k = 2, 3, \dots, n$$

where $x^{(0)}(1) = x^{(1)}(1)$.

**Step 2:** The GM(1,1) model is given by

$$x^{(0)}(k) + a\, z^{(1)}(k) = b,$$

where

$$z^{(1)}(k) = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right).$$

The parameters $a, b$ are calculated by

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B'B)^{-1}B'Y,$$

where

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \text{and } B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}.$$

**Step 3:** The time response sequence of GM(1,1) is obtained by

$$\hat{x}^{(0)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)(1 - e^{a})e^{-ak}, k = 1,2,3, \dots . \quad (2.1)$$

Consequently, we have two data series, namely, a GM(1,1) training data $\{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n)\}$ and a GM(1,1) prediction series $\{\hat{x}^{(0)}(n+1), \hat{x}^{(0)}(n+2), \dots, \hat{x}^{(0)}(n+k), \dots\}, k = 1,2,3, \dots$.

**2.2 MCGM (1,1) model**

We combine Markov chain and GM(1,1) in order to improve the forecasting. The Markov chain rule of transition probability matrix is constructed by error series of original data and training series of grey model as in the previous section, that is,

$$e(i) = x^{(0)}(i) - \hat{x}^{(0)}(i), \qquad i = 1,2,3, \dots, n. \quad (2.2)$$

Let $r$ be the number of states, $s_0 = \min\{e(i),\ i = 1,2,3, \dots n\}$, and $s_1 = \max\{e(i), i = 1,2,3, \dots, n\}$. Set the error $e_0$ in the state $j^{th}$ (where $1 \le j \le r$) when

$$e_0 \in \left(s_0 + (j-1)\frac{(s_1 - s_0)}{r}, s_0 + j\frac{(s_1 - s_0)}{r}\right]$$

(observe that $e_0$ is in the first state when $e_0 = s_0$). Thus, we obtain the middle point of the $j^{th}$ state as

$$v_j = s_0 + (2j-1)\frac{(s_1 - s_0)}{2r} \text{ for all } j = 1,2,3,\dots,r.$$

Next, we let $R = [p_{ij}]_{r\times r}$ be the transition matrix, where $p_{ij}$ is a transition probability of moving one step from the $i^{th}$ state to the $j^{th}$ state for all $i,j = 1,2,\dots,r$. This leads to the property that $A_{M+1} = A_M R$ where $A_M = [a_1(M)\ a_2(M)\ \dots\ a_r(M)]$ and $a_i(M)$ is the opportunity of event that will be in the $i^{th}$ state at time step $M$ for $M = 1,2,3,\dots$ . Therefore, MCGM(1,1) is given by

$$\tilde{x}^{(0)}(M+k) = \hat{x}^{(0)}(M+k) + (A_{M+k-1}R)V, \qquad k = 1,2,3,\dots,$$

where $V = [v_1\ v_2\ \dots\ v_r]^{'}$ is the matrix of all middle points. If the original series have $n$ samples and $M = 1$, then

$$\tilde{x}^{(0)}(k+1) = \hat{x}^{(0)}(k+1) + (A_k R)V$$

where $\{\tilde{x}^{(0)}(1), \tilde{x}^{(0)}(2), \dots, \tilde{x}^{(0)}(n)\}$ is called an MCGM(1,1) training series, and $\{\tilde{x}^{(0)}(n+1), \tilde{x}^{(0)}(n+2), \dots, \tilde{x}^{(0)}(n+k), \dots\}$ is called an MCGM(1,1) prediction series.

## 3. Main Results

We separate the data series of gold price from January 1990 to June 2014 into two parts. The first part, which is the data series from January 1990 to December 2011, is considered for training. The second part, which is the data series from January 2012 to June 2014, is considered for testing the accuracy of GM(1,1) and MCGM(1,1).

### 3.1 The accuracy of GM(1,1) and MCGM(1,1)

Based on the original data series from January 1990 to December 2011, we obtain the parameters of GM(1,1) as $[a\ \ b]^{'} = [-0.00868\ \ 83.8843]$, that is, the GM(1,1) model is given by

$$\hat{x}^{(0)}(k+1) = 87.066\exp(0.00868k),\ \ k = 0,1,2,\dots . \qquad (3.1)$$

Based on the GM (1,1) prediction series, we classify the error series as mentioned in (2.2) into 10 states: each state is considered as in Section 2.2. Therefore, the transition matrix $R_T$ of the error series is obtained as

$$R_T = \begin{bmatrix} 0.99 & 0.01 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.03 & 0.88 & 0.09 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.05 & 0.87 & 0.08 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.73 & 0.07 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.11 & 0.67 & 0.22 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.5 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.6 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.67 & 0.33 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.25 & 0 & 0.5 & 0.25 \end{bmatrix}.$$

The middle point matrix is obtained as

$$V_T = [22.77\ \ 119.32\ \ 15.80\ \ 312.32\ \ 408.80\ \ 505.00\ \ 601.90\ \ 698.39\ \ 794.91\ \ 891.40]'$$

Therefore, MCGM(1,1) model is given by

$$\tilde{x}^{(0)}(k+1) = 87.066\exp(0.00868k) + A_k R_T V_T, \quad k = 0,1,2,\dots \qquad (3.2)$$

Next, we shall compare the accuracy of GM(1,1) and MCGM(1,1) by using the testing data from January 2012 to June 2014. We use three criterions: mean square error (MSE), mean absolute error (MAE) and absolute relative error (ARE). To be precise,

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left|X(i) - \hat{X}(i)\right|^2, AME = \frac{1}{n}\sum_{i=1}^{n}\left|X(i) - \hat{X}(i)\right|, \text{and } ARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X(i) - \hat{X}(i)}{X(i)}\right|$$

The results are shown in Table 3.1.

**Table 3.1** Comparison of the forecasting results in terms of MSE, AME, and ARE

| Model | Criterions | | |
|---|---|---|---|
| | MSE | AME | ARE |
| GM | 321,974.99 | 510.22 | 0.3270 |
| MCGM | 68,273.95 | 239.20 | 0.1689 |

From Table 3.1, for each criterion, we see that MCGM(1,1) forecasting is better than GM(1,1) for predicting the testing data.

**3.2 Forecasting model**

From Table 3.1, we therefore choose the MCGM(1,1) to be the forecasting model of this research in order to predict the gold price in the next period.

From the original data series from January 1990 to June 2014, we obtain the parameters of GM(1,1) as $[a\ \ b]' = [-0.008502\ \ 94.80305]'$, that is, the GM(1,1) model is given by

$$\hat{x}^{(0)}(k+1) = 97.87\exp(0.0085k), \qquad k = 0,1,2,\dots.$$

By using the original data series from January 1990 to June 2014, we classify the error series into 10 states. Therefore, the transition probability matrix $R_F$ of the error series is obtained as

$$R_F = \begin{bmatrix} 0.975 & 0.025 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.043 & 0.869 & 0.087 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.077 & 0.785 & 0.138 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.175 & 0.789 & 0.035 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.222 & 0.556 & 0.222 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.4 & 0.4 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.143 & 0.571 & 0.143 & 0 & 0.143 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.111 & 0.667 & 0.222 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.222 & 0.444 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0.6 & 0.2 \end{bmatrix}.$$

The middle point matrix is obtained as

$$V_F = [-7.59\ \ 85.75\ \ 179.10\ \ 272.46\ \ 365.81\ \ 459.16\ \ 552.51\ \ 645.86\ \ 739.21\ \ 832.56]^T.$$

Therefore, MCGM(1,1) model is given by

$$\tilde{x}^{(0)}(k+1) = 97.87\exp(0.0085k) + A_k R_F V_F, \quad k = 0,1,2,\ldots.$$

Thus we can forecast gold price by using MCGM(1,1) model from July 2014 to June 2017.

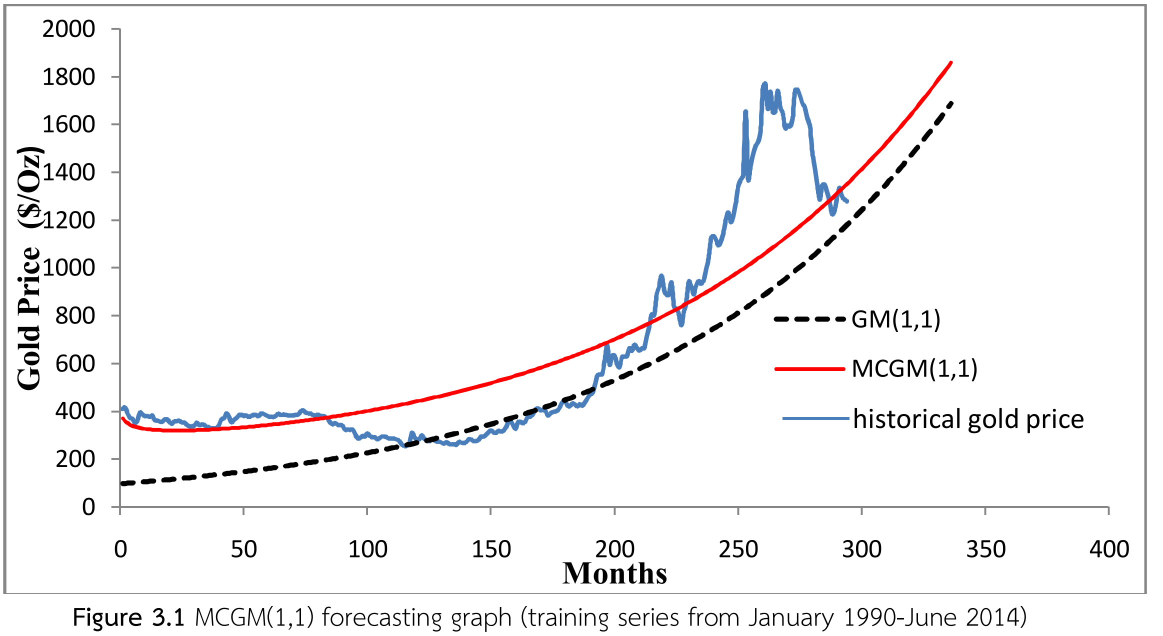


**Figure 3.1** MCGM(1,1) forecasting graph (training series from January 1990-June 2014)

In Figure 3.1, the blue line shows the historical gold price, the red line shows the forecast obtained by GM(1,1), and the black dashed line reveals the MCGM(1,1) forecasting. From simulation, we obtain the forecasting result as shown in Table 3.2

**Table 3.2** Gold price forecasting by the MCGM(1,1) (from July 2014 to June 2017)

| Month/Year | Gold price ($/Oz) | Month/Year | Gold price ($/Oz) | Month/year | Gold price ($/Oz) |
|---|---|---|---|---|---|
| Jul 2014 | 1361 | Jul 2015 | 1489 | Jul 2016 | 1631 |
| Aug 2014 | 1371 | Aug 2015 | 1501 | Aug 2016 | 1644 |
| Sep 2014 | 1381 | Sep 2015 | 1511 | Sep 2016 | 1656 |
| Oct 2014 | 1392 | Oct 2015 | 1523 | Oct 2016 | 1669 |
| Nov 2014 | 1403 | Nov 2015 | 1535 | Nov 2016 | 1682 |
| Dec 2014 | 1413 | Dec 2015 | 1547 | Dec 2016 | 1695 |
| Jan 2015 | 1424 | Jan 2016 | 1558 | Jan 2017 | 1708 |
| Feb 2015 | 1434 | Feb 2016 | 1570 | Feb 2017 | 1721 |
| Mar 2015 | 1445 | Mar 2016 | 1582 | Mar 2017 | 1734 |
| Apr 2015 | 1456 | Apr 2016 | 1594 | Apr 2017 | 1747 |
| May 2015 | 1467 | May 2016 | 1606 | May 2017 | 1761 |
| Jun 2015 | 1478 | Jun 2016 | 1619 | Jun 2017 | 1774 |

## 4. Conclusions

The objective of this paper is to develop the forecasting model for gold price. We use the data of gold price in London from January 1990 to December 2011 for training, and use the data from January 2012 to June 2014 to test the model. In this way we use three criterions, namely, MSE, AME and ARE to measure the accuracy of our model. The result reveals that MCGM(1,1) forecasting model is more accurate than GM(1,1). This model is also suitable for forecasting in human society, such as weather, agriculture, and so on. For the gold price, the forecasting value of MCGM(1,1) in July 2014 is 1361$/Oz.

## 5. References

Huang, M., He, Y. and Cen, H. (2007). Predictive analysis on electric-power supply and demand in China. Renewable Energy 32(7): 1165–1174.

Julong, D. (1982). Control problem of grey system. System Control Letters 1(5): 288-294.

Kazemi, A., Modarres, M., Mehregan, M. R., Neshat, N., Foroughi, A. (Eds). (2011). Proceeding of the 3rd international conference on information and financial engineering: A Markov chain Grey forecasting Model: A case study of energy demand of industry sector In Iran. (pp.13-18). Singapore: IACSIT Press.

Li, G.D., Yamaguchi, D. and Nagai, M. (2007). GM (1,1)-Markov chain combined model an application to predict the number of Chinese international airlines. Technological Forecasting & Social Change 74(8): 1465-1481.

❑❑❑❑❑