



Implementation of Geometric Brownian Motion to Predict Crude Oil Prices

Feby Seru¹, Christian Dwi Suhendra², Agung Dwi Saputro³, Gautam Makwana⁴, H Elizabeth⁵

^{1,3} Department of Mathematics Universitas Cenderawasih, Indonesia

² Department of Informatic Engineering Universitas Papua, Indonesia

^{4,5} Department of Social Work, Mizoram University, India

Correspondence: ✉ febyseru.math@gmail.com

Article Info	Abstract
Article History Received: 13-09-2022 Revised: 04-11-2022 Accepted: 12-11-2022	Crude oil has a vital role in the economic growth of a country because crude oil is a source of energy driving the economy. To maintain economic stability, the price of crude oil in the coming period needs to be anticipated by making predictions on world crude oil commodity prices. One of the models that can be used to predict crude oil prices in the short term is Geometric Brownian Motion (GBM). This study aims to implement the GBM model to predict crude oil prices during the Covid-19 pandemic and measure the model's accuracy. This study made crude oil price predictions with several iterations of 50, 100, and 1000. The results showed that the smallest MAPE value was carried out 1000 times in iterations, namely 2.13%. Based on the MAPE value, it can be concluded that the results of crude oil price predictions using GBM have a high level of accuracy.
Keywords: Crude Oil; Geometric Brownian Motion; Predictions	

INTRODUCTION

Crude oil price movements are always an exciting thing to observe. Because OF crude oil is an essential factor, especially in the economic activities of a country. The largest energy source is processed crude oil, such as Liquified Petroleum Gas (LPG), gasoline, diesel, lubricating oil, fuel oil, and others [1]. In addition, changes in crude oil prices also affect changes in other commodity prices [2]. World crude oil prices are measured using the world oil market prices. Standard world oil prices generally refer to *West Texas Intermediate* (WTI). WTI is a high-quality petroleum produced in Texas [3].

Fluctuations in oil prices have always been considered a barometer of the economy worldwide, so any erratic movement is of particular concern and is always a hot issue to be discussed in political and economic circles in every country. Oil price fluctuations are sensitive to each country's economic conditions and growth [1].

Three factors affect world crude oil prices: fundamental factors, non-fundamentals, and the influence of OPEC's supply policy. Fundamental factors include oil demand, supply, stock, world reserve production capacity, and world refinery capability. At the same time, non-fundamental factors include geopolitics, government policies, weather, natural disasters, strikes, damage to production chain installations, weakening of the dollar, and speculation [4].

The stability of the world oil supply and prices is needed to boost economic growth. Therefore, so that all activities can go according to plan, predictions of world oil prices are needed by many parties, both from the government sector, business entities, and investors. Prediction of crude oil prices is intended to reduce the impact of price fluctuations and assist investors and individuals in making decisions related to the energy market.

Geometric Brownian Motion (GBM) is a model that can be used to predict stock and commodity prices. Several researchers have researched stock price prediction using GBM, including Trimono [5], Farida Agustini, et al. [6], Suganthi and Jayalalitha [7], Maulidya et al. [8], Putri and Hasibuan [9], Rosita [10] and Bhakti [11]. The results obtained from these studies are that the GBM model can be used to predict stock prices in the future for a short period. Research on the accuracy of GBM in predicting stock prices during the Covid-19 pandemic has also been carried out by several researchers, including Fitria et al. [12] and Edriani et al. [13] show that the results of stock price predictions during the Covid-19 pandemic using GBM are inaccurate. According to Fitria et al. [12]. This is due to the inappropriate use of historical data when modeling stock prices. At the same time, the research conducted by Edriani et al. [13] shows that the GBM model successfully represents stock price movements so that the model can be used to predict stock prices during the Covid-19 pandemic.

Research about prediction price commodities, specifically oil raw using GBM, has been conducted by Zakia [14] on predicting the price of *West Texas Intermediate* crude oil (WTI) using GBM. The results indicate that the prediction of crude oil prices using GBM is highly accurate. At the same time, the research conducted by Bahar et al. [15] shows that the results of crude oil price predictions using GBM are very accurate for a short period. Research conducted by Zakia [14] was conducted when economic conditions were normal. Therefore, it is essential to know whether the GBM model can predict crude oil prices during abnormal conditions, especially during the Covid-19 pandemic.

Based on the background above, this study aims to measure the accuracy of the GBM model in predicting crude oil prices during the Covid-19 pandemic. Forecasting accuracy is measured using the *Mean Absolute Percentage Error* (MAPE).

METHOD

This section explains the steps for constructing the GBM model on crude oil price data and using the model to predict crude oil prices. The data used in this research is the daily closing price of *West Texas Intermediate* crude oil (WTI) from June 2020 - June 2021. This data is secondary data obtained from <https://id.investing.com>. The steps taken to predict the price of crude oil using GBM are as follows:

1. Compile the daily closing price of crude oil
2. Determine *in-sample* data and *out-sample* data
3. Calculating stock *return* value from *in-sample* data
4. *in-sample* data normality test using Kolmogorov-Smirnov. test
5. *drift* parameter estimates (μ) and volatility (σ)
6. Modeling crude oil prices using GBM
7. Making crude oil price predictions
8. Validate the model from the *out-sample* data by calculating the MAPE value.

Return

The return value of crude oil in the period t , is defined as [16]:

$$R_t = \ln \left(\frac{S_t}{S_{t-1}} \right) \quad (1)$$

with:

R_t : return value in period t

S_t : the price of crude oil in the period t

S_{t-1} : the price of crude oil in the period $t - 1$

Normality test

Kolmogorov–Smirnov is one of the normality tests that can be done to determine the normality of data. Kolmogorov's test was carried out by comparing D_{count} and $D_{\alpha,n}$. It is obtained from the table [17].

Hypothesis:

H_0 : Data return is normally distributed

H_1 : Data return is not normally distributed

Test Statistics:

$$D_{hitung} = \max |F_0(x) - S_N(x)| \quad (2)$$

with:

$F_0(x)$: cumulative distribution function of a normal distribution

$S_N(x)$: cumulative distribution function of a sample data

Test Criteria:

If $D_{hitung} < D_{\alpha,n}$ it is H_0 Accepted, it means that the data is normally distributed. The Kolmogorov-Smirnov test can be carried out using SPSS with the test criteria, if $P_{value} > \alpha$ then it H_0 . It is accepted, which means that the data is normally distributed.

Brownian Motion

A stochastic process is $\{X_t, t \geq 0\}$ called *Brownian Motion* if it fulfills [18]:

- $X_0 = 0$
- $\{X_t, t \geq 0\}$ have stationary increments and are independent of each other
- X_t normal distribution with a mean of 0 and a variance $\sigma^2 t$ for each $t > 0$

At time $\sigma = 1$, then the above process is called standard *Brownian Motion*. The stochastic process is $\{X_t, t \geq 0\}$ called *Brownian Motion* with drift if it fulfills the following conditions [18]:

- $X_0 = 0$
- $\{X_t, t \geq 0\}$ have a stationary and independent increment
- X_t Normal distribution with mean μt and variance $\sigma^2 t$ for each $t > 0$

Based on these assumptions, the *Brownian Motion* with *drift* can be written as follows:

$$X_t = \mu t + \sigma B_t \quad (3)$$

with B_t is *Brownian standard motion*.

Geometric Brownian Motion

Suppose the stochastic process $\{X_t, t \geq 0\}$ is *Brownian Motion* with *drift* as in Equation (3), then the process $\{S_t, t \geq 0\}$ is called *Geometric Brownian Motion* and is defined as [18]:

$$S_t = e^{X_t} \quad (4)$$

The GBM model can be used to model the process of crude oil price movements [11]. GMB assumes that the oil return data in Century is then normally. In general, the GBM model can be expressed as [19]:

$$dS_t = \mu S_t dt + \sigma S_t dB_t \quad (5)$$

with:

S_t : Oil price at the time t

dS_t : Changes in oil prices at the time of t

dB_t : Changes in the Wiener process (*Brownian Motion*).

B_t : *Brownian Motion* standard

μ : percentage of *drift*

σ : percentage of the volatility

dt : time intervals

Equation (5) can be solved using the following formula:

$$dF(S_t, t) = \left(\frac{\partial F(S_t, t)}{\partial S_t} \mu S_t + \frac{\partial F(S_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 F(S_t, t)}{\partial S_t^2} \sigma^2 S_t^2 \right) dt + \left(\frac{\partial F(S_t, t)}{\partial S_t} \sigma S_t \right) dB_t \quad (6)$$

with $F(S_t, t)$ is a function of the variables S_t and t . Equation (6) is known as Ito's Lemma.

Suppose a function $F(S_t, t) = \ln S_t$ With:

$$\frac{\partial F(S_t, t)}{\partial S_t} = \frac{1}{S_t} \quad (7)$$

$$\frac{\partial F(S_t, t)}{\partial t} = 0 \quad (8)$$

$$\frac{\partial^2 F(S_t, t)}{\partial S_t^2} = -\frac{1}{S_t^2} \quad (9)$$

Then Equation (6) becomes:

$$dF(S_t, t) = \left(\frac{1}{S_t} \mu S_t + 0 + \frac{1}{2} \left(-\frac{1}{S_t^2} \right) \sigma^2 S_t^2 \right) dt + \left(\frac{1}{S_t} \sigma S_t \right) dB_t$$

$$dF(S_t, t) = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t$$

$$\ln S_t - \ln S_{t-1} = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dB_t$$

$$\begin{aligned}\ln S_t &= \ln S_{t-1} + \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dB_t \\ S_t &= e^{\ln S_{t-1} + \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dB_t} \\ S_t &= S_{t-1} e^{\left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dB_t}\end{aligned}\quad (10)$$

For example $B_t = \epsilon\sqrt{dt}$, then Equation (10) becomes:

$$S_t = S_{t-1} e^{\left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\epsilon\sqrt{dt}} \quad (11)$$

With:

- S_t : The price of oil at the moment t
- S_{t-1} : The price of oil at the moment $t - 1$
- S_0 : Initial price of oil
- μ : drift parameters
- σ : Volatility parameters
- ϵ : standard normal distribution

Parameter Estimation

Volatility is the rate of oil price movement, while drift is an expectation from the rate of oil price movements. Volatility Parameters estimated use equality [20]:

$$\sigma = s/\sqrt{\tau} \quad (12)$$

with

$$\bar{R} = \frac{1}{n} \sum_{t=1}^n R_t \quad (13)$$

$$s = \sqrt{\frac{1}{n-1} \sum_{t=1}^n (R_t - \bar{R})^2} \quad (14)$$

Moreover, estimated drift parameters use:

$$\mu = \frac{\bar{R}}{\tau} + \frac{1}{2} \sigma^2 \quad (15)$$

with:

- σ : volatility value
- s : standard deviation return
- \bar{R} : average return
- μ : drift parameter value
- τ : interval of time $t_i - t_{i-1}$
- n : lots of data
- R : the *return* of oil price

Mean Absolute Percentage Error (MAPE)

MAPE is the average absolute percentage of prediction error. MAPE is an essential factor in measuring the level of forecasting accuracy. The MAPE formula is defined as [21]:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|S_t - F_t|}{S_t} \cdot 100\% \quad (16)$$

with:

S_t : actual value
 F_t : oil price prediction
 N : the amount of data

The following is a table of the level of forecasting accuracy:

Table 1. MAPE Value Scale for Forecasting Accuracy Level

MAPE Percentage	Level of accuracy
< 10%	High forecasting accuracy
10% – 20%	Good forecasting accuracy
21% – 50%	Ordinary forecasting accuracy
> 50%	Inaccurate forecast

RESULTS AND DISCUSSION

The closing crude oil price data used as historical data in this study is from June 2020 to June 2021, totaling 270 data. The selection of sample data will affect the prediction results obtained. Fitria et al. [12] used data before the pandemic as historical data, resulting in inaccurate prediction results. Therefore in this study, the historical data used was crude oil price data at the time of the Covid-19 pandemic.

In the processing, historical data is divided into two. Namely, 80% of the total sample (216 data) is used as *in-sample data*, and 20% of the total sample (54 data) is used as *out-sample data*. *In-sample data* is used to build the model, while *out-sample data* is used for model validation [5]. The following is a plot for the *in-sample data*.

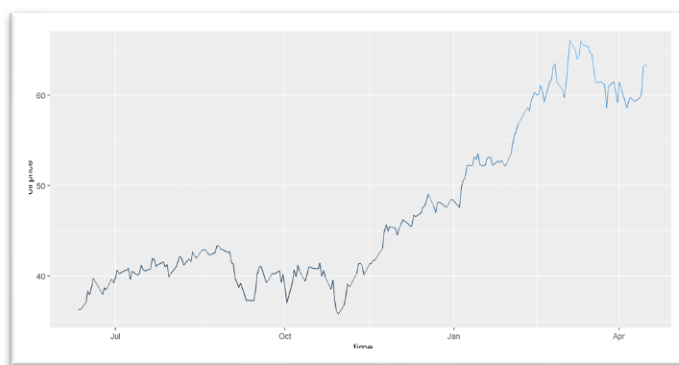


Figure 1. The plot of WTI crude oil price data

Figure 1 shows that the trend of oil price fluctuations generally tends to increase even though it has decreased at certain times. Furthermore, calculating the return value from the *in-sample data* is carried out using Equation (1). The plot of return on crude oil prices is as follows:

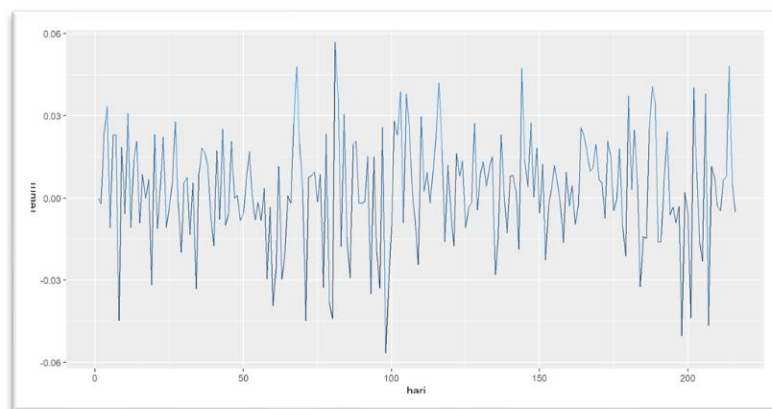


Figure 2. The plot of return of crude oil price data

In order to model oil prices using GBM, the normality test of the crude oil return data was first tested using the Kolmogorov-Smirnov test. Based on the Kolmogorov-Smirnov test using SPSS, the *p-value* = 0,2 > 0,05, so it H_0 is accepted, which means the raw oil return data is normally distributed and meets the GBM model's assumptions.

Drift and volatility parameters use Equations (12) and (15) and thus obtained:

Table 2. Parameter Estimation Value

Parameter	Score
$\hat{\mu}$	0,002778
$\hat{\sigma}$	0,020482

So the GBM model for crude oil price data is:

$$\hat{S}_t = \hat{S}_{t-1} e^{\left(0,002778 - \frac{0,000420}{2}\right) dt + 0,020482 \epsilon \sqrt{dt}} \quad (17)$$

After obtaining the GBM model for crude oil prices, crude oil price predictions for *out-sample data* are then carried out using Equation (17). In Equation (17), the value ϵ is obtained from random data standard normal distribution, so in this study, the predictive value was calculated in several iterations to produce an accurate predictive value.

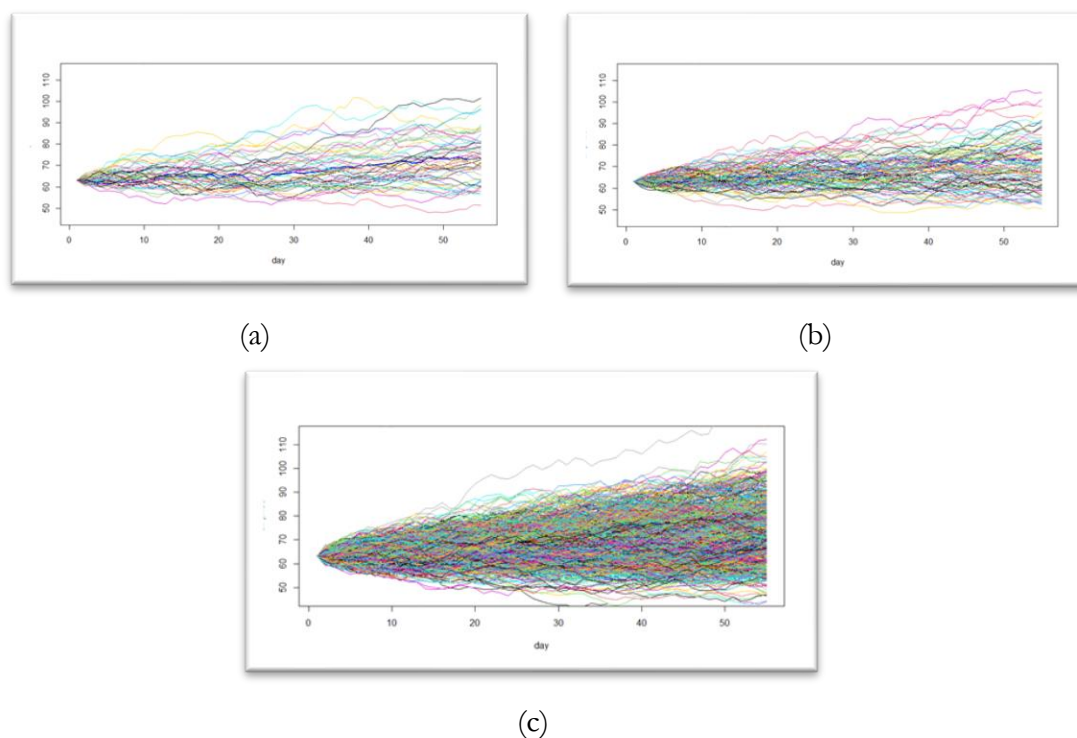


Figure 3. Prediction of crude oil prices with (a) iterations of 50, (b) 100, and (c) 1000

Figure 3 is the result of crude oil price prediction using GBM, which was carried out with several iterations of 50, 100, and 1000. In Figure 3(a), there are 50 prediction results represented by 50 colors, while in Figure 3(b), there are 100 prediction results represented by 100 colors, and in Figure 3(c), there are 1000 prediction results represented by 1000 colors. Each number of iterations produces a different predictive value due to random data. Random data selection will affect the prediction results obtained. To find out the best prediction results from each iteration, the MAPE value is calculated using Equation (16). The smallest MAPE value for each iteration is presented in the following table:

Table 3. The smallest MAPE value for each iteration

No	Number of Iterations	MAPE (%)
1	50	3.121317
2	100	2.289672
3	1000	2.133447

Table 3 shows that the smallest MAPE value is carried out 1000 times in iterations, namely 2.13%. These results are from Bhakti et al. [11] that the more iterations carried out, the smaller the resulting *error value*. The MAPE value obtained in Table 3 is less than 10%, so based on the MAPE value scale in Table 1, it can be concluded that the prediction results of crude oil prices during the Covid-19 pandemic, which were carried out using GBM, had a high level of accuracy.

The results of the prediction of crude oil prices, which have the smallest MAPE value in each iteration, compared to the actual data obtained from the out-sample data, are as follows:

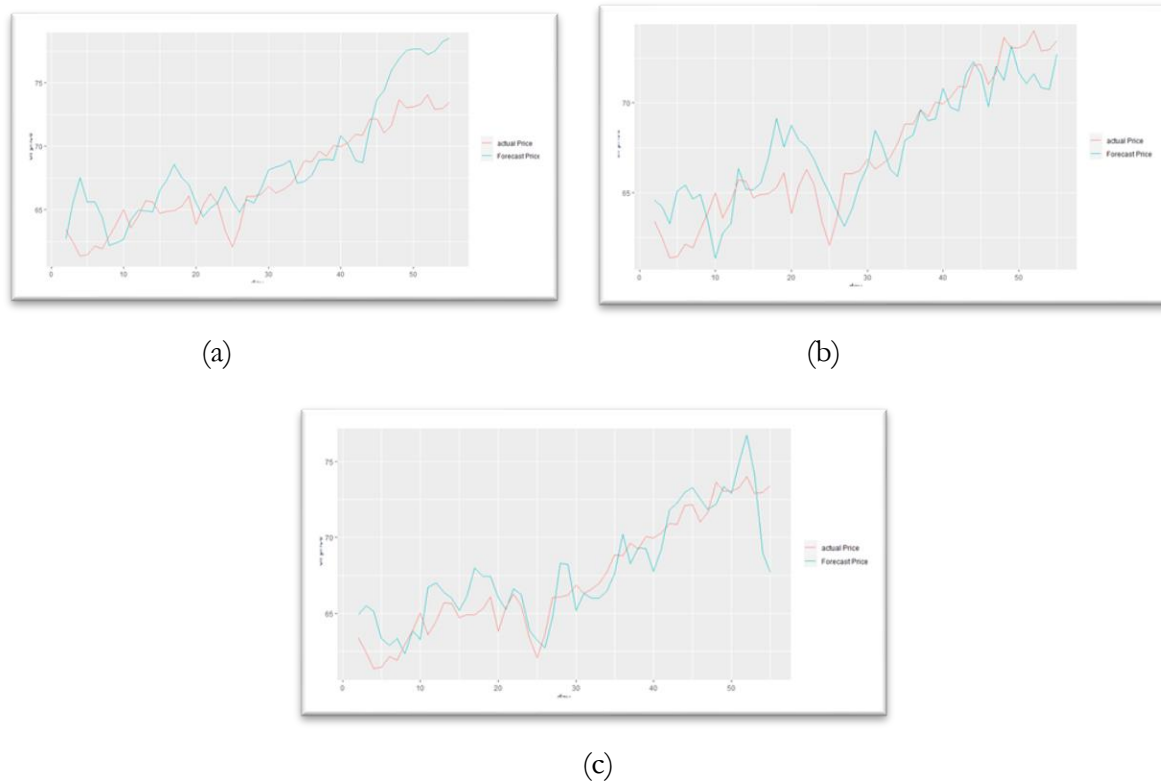


Figure 4. Oil price prediction with the smallest MAPE at (a) 50, (b) 100, and (c) 1000 iterations

Figure 4 shows a graph between the predicted value of crude oil prices and the actual data for each iteration. In Figure 4 (a), there are several significant gaps between the predicted and actual values. This is also reinforced by the significant MAPE value of 3.12%. In Figures 4 (b) and 4 (c), the gap between the predicted value and the actual value decreases, which means that the predicted value is almost close to the actual value.

The following are the results of crude oil price predictions with the smallest MAPE value in 1000 iterations.

Table 4. Oil Price Prediction

No.	T date	Prediction	current	No.	T date	Prediction	current
1	19/04/2021	64.90	63.38	28	26/05/2021	68.21	66.21
2	20/04/2021	65.49	62.44	29	27/05/2021	65.16	66.85
3	21/04/2021	65.15	61.35	30	28/05/2021	66.31	66.32
4	22/04/2021	63.33	61.43	31	30/05/2021	66.01	66.59
5	23/04/2021	62.88	62.14	32	31/05/2021	66.00	66.96
6	26/04/2021	63.33	61.91	33	01/06/2021	66.45	67.72
7	27/04/2021	62.32	62.94	34	02/06/2021	67.62	68.83
8	28/04/2021	63.87	63.86	35	03/06/2021	70.21	68.81
9	29/04/2021	63.28	65.01	36	04/06/2021	68.28	69.62
10	30/04/2021	66.68	63.58	37	07/06/2021	69.32	69.23
11	03/05/2021	67.03	64.49	38	08/06/2021	69.27	70.05
12	04/05/2021	66.40	65.69	39	09/06/2021	67.75	69.96
13	05/05/2021	65.99	65.63	40	10/06/2021	69.21	70.29
14	06/05/2021	65.17	64.71	41	11/06/2021	71.81	70.91
15	07/05/2021	66.13	64.9	42	14/06/2021	72.30	70.88
16	10/05/2021	67.98	64.92	43	15/06/2021	72.97	72.12
17	11/05/2021	67.43	65.28	44	16/06/2021	73.29	72.15
18	12/05/2021	67.43	66.08	45	17/06/2021	72.57	71.04
19	13/05/2021	66.03	63.82	46	18/06/2021	71.86	71.64
20	14/05/2021	65.27	65.37	47	21/06/2021	72.19	73.66
21	17/05/2021	66.62	66.27	48	22/06/2021	73.36	73.06
22	18/05/2021	66.24	65.49	49	23/06/2021	72.89	73.08
23	19/05/2021	63.85	63.36	50	24/06/2021	74.99	73.3 0
24	20/05/2021	63.23	62.05	51	25/06/2021	76.78	74.05
25	21/05/2021	62.72	63.58	52	28/06/2021	74.18	72.91
26	24/05/2021	64.71	66.05	53	29/06/2021	68.98	72.98
27	25/05/2021	68.32	66.07	54	30/06/2021	67.68	73.47

CONCLUSION

Based on the study's results, it can be concluded that the smallest MAPE value is in the 1000 iterations, which is 2.13%. The MAPE value is less than 10%, which means that the results of predicting crude oil prices using GBM have a high accuracy level, so the GBM model can be used to predict crude oil prices during the Covid-19 pandemic.

REFERENCES

- [1] Y. Arifin, "Pengaruh Harga Minyak Dunia, Nilai Tukar dan Inflasi terhadap Pertumbuhan Ekonomi Indonesia," *Economics Development Analysis Journal*, vol. 5, no. 4, pp. 474–483, Mar. 2016, doi: 10.15294/EDAJ.V5I4.22184.
- [2] R. V. Febryo, G. Abdillah, and R. Yuniarti, "Sistem Prediksi Harga Minyak Mentah Menggunakan Jaringan Syaraf Tiruan dengan Backpropagation," *SNIIJA*, pp. 5–8, 2015.
- [3] P. Dewi, "Pengaruh Inflasi, Kurs, dan Harga Minyak Dunia Terhadap Indeks Harga Saham Gabungan di Bursa Efek Indonesia," *Jurnal Ilmu Manajemen*, vol. 17, no. 1, pp. 10–19, Sep. 2020, doi: 10.21831/JIM.V17I1.34772.
- [4] R. A. Fauzannissa, H. Yasin, and D. Ispriyanti, "Peramalan Harga Minyak Mentah Dunia Menggunakan Metode Radial Basis Function Neural Network," *Jurnal Gaussian*, vol. 5, no. 1, pp. 193–202, Jan. 2016, doi: 10.14710/J.GAUSS.V5I1.11049.
- [5] Trimono, D. I. Asih Maruddani, and D. Ispriyanti, "Pemodelan Harga Saham Dengan Geometric Brownian Motion dan Value at Risk PT Ciputra Development Tbk," *Jurnal Gaussian*, vol. 6, no. 2, pp. 261–270, 2017.
- [6] W. Farida Agustini, I. R. Affianti, and E. R. M. Putri, "Stock price prediction using geometric Brownian motion," in *Journal of Physics: Conference Series*, Mar. 2018, vol. 974, no. 1. doi: 10.1088/1742-6596/974/1/012047.
- [7] K. Suganthi and G. Jayalalitha, "Geometric Brownian Motion in Stock Prices," *Journal of Physics: Conference Series*, vol. 1377, no. 1, p. 012016, Nov. 2019, doi: 10.1088/1742-6596/1377/1/012016.
- [8] V. Maulidya, E. Apriliani, and E. R. M. Putri, "Prediksi Harga Saham Menggunakan Geometric Brownian Motion Termodifikasi Kalman Filter dengan Konstrain," *Indonesian Journal of Applied Mathematics*, vol. 1, no. 1, pp. 6–18, 2020.
- [9] D. M. Putri and L. H. Hasibuan, "Penerapan Gerak Brown Geometrik pada Data Saham PT. ANTM," *Mathematics & Applications Journal*, vol. 1, no. 1, pp. 1–10, 2020.
- [10] Y. Rosita, "Prediksi Saham Menggunakan Geometric Brownian Motion," *PRISMAKOM*, vol. 1, no. 19, pp. 19–24, 2021.
- [11] H. D. Bhakti, "Prediksi Harga Saham Subsektor Farmasi Menggunakan Geometric Brownian Motion," *Jurnal Media Informatika*, vol. 6, pp. 395–403, 2022, doi: 10.30865/mib.v6i1.3415.
- [12] I. Fitria, K. N. Handayati, and P. Hasanah, "The Application of Geometric Brownian Motion in Stock Forecasting During The Coronavirus Outbreak in Indonesia," *Journal of Physics: Conference Series*, vol. 1821, no. 1, 2021, doi: 10.1088/1742-6596/1821/1/012008.
- [13] T. S. Edriani, U. S. Pasaribu, Y. S. Afrianti, and N. N. W. Astute, "The Geometric Brownian Motion of Indosat Telecommunications Daily Stock Price During The Covid-19 Pandemic in Indonesia," *Journal of Physics: Conference Series*, vol. 2084, no. 1, 2021, doi: 10.1088/1742-6596/2084/1/012012.
- [14] H. I. Zakia, "Prediksi Harga Komoditas Minyak Mentah Menggunakan Model Geometric Brownian Motion," *Institut Teknologi Sepuluh Nopember Fakultas Matematika*, 2017.
- [15] A. Bahar, N. Mohd Noh, and Z. M. Zainuddin, "Forecasting Model for Crude Oil Price With Structural Break," *Malaysian Journal of Fundamental and Applied Sciences*, vol. 13, no. 4–1, pp. 421–424, 2017, doi: 10.11113/mjfas.v13n4-1.861.

- [16] D. Ruppert and D. S. Matteson, *Statistics and Data Analysis for Financial Engineering with R example Second Edition*. New York: Springer, 2015. doi: 10.1201/9781315171401-4.
- [17] F. J. Massey, "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68–78, 1951, doi: 10.1080/01621459.1951.10500769.
- [18] S. M. Ross, *Introduction to Probability Models Eleven Edition*. California: California: Elsevier, 2014.
- [19] A. S. Agbam, "Stock Prices Prediction Using Geometric Brownian Motion: Analysis of the Nigerian Stock Exchange," *Engineering and Applied Scientific Reserach*, vol. 13, no. 1, pp. 1–34, 2021.
- [20] A. Dmouj, "Stock Price Modelling : Theory and Practice," *Vrije Universiteit Faculty of Sciences Amsterdam*, 2006.
- [21] Paul. Wilmott, *Introduces Quantitative Finance*. America: Jhon Wiley & Sons, 2007.