COMPARISON OF RANDOM FOREST AND SINGLE EXPONENTIAL SMOOTHING METHODS IN THE PREDICTION OF GOLD DEMAND FOR INDONESIAN JEWELRY PERIOD 2010-2021

Seni Eriani[1], Wiwit Pura Nurmayanti[1*], Harista Almiatus Soleha[1], Muhammad Gazali[1], Kertanah[1], Siti Hadijah Hasanah[2]

[1]Department of Statistics, Universitas Hamzanwadi (INDONESIA)
[2]Study Program of Statistics, Universitas Terbuka (INDONESIA)

**Corresponding author: wiwit.adiwinata3@gmail.com*

Abstract

Forecasting is an activity to predict something that will happen in the future. In forecasting, there are several methods, including Random Forest and Single Exponential Smoothing (SES). Random Forest has the advantage that it does not require assumptions, while SES requires assumptions as data that must be stationary and used for short-term forecasting. We can apply Random Forest and SES in various fields, one of which is economics, which focuses on the demand for gold jewellery. The importance of analysing data on the number of requests for gold jewellery is because there are many enthusiasts from gold itself. This study compares the two methods in predicting the demand for gold jewellery in Indonesia. Based on the results of the analysis, we found Random Forest is better than SES, we can see it from the forecast error value of Random Forest, which is smaller than SES. The Random Forest method got the best model with n-tree 50, resulting in an MAPE value of 12.87%. Meanwhile, the best model for SES with an alpha of 0.4 produces an MAPE value of 20.63%.

Keywords: Forecasting, Random Forest, Single Exponential Smoothing, gold jewellery, MAPE

1 INTRODUCTION

Forecasting is an activity with the aim of predicting what will happen in the future [1][2]. There are several forecasting methods such as Moving Average (MA), Exponential Smoothing (ES), Autoregressive Integrated Moving Average (ARIMA) to Random Forest (RF). Single Exponential Smoothing method is a method used in short-term forecasting by assuming that the data fluctuates around a fixed average value without a trend or consistent growth pattern [3]. Random forest was first introduced by Breiman in 2001, the Random Forest method works by building a model using several decision trees at random and combining the predictions of each tree to get prediction results. Random Forest is a forecasting method that does not require any assumptions [4].

This forecasting method can be applied in various fields, one of which is the economic field, namely, to find out future gold demand. Gold is a precious metal that is in high demand, both for investment and as jewellery [5]. Gold has many enthusiasts because the value of gold tends to be stable, easy to reach, easy to melt and of course promises quite large profits, especially in Indonesia. Indonesia is one of the countries in Southeast Asia that is included in the category of high demand for gold. This is evidenced based on data from the World Gold Council (WGC) that demand for gold jewellery in Indonesia in the fourth quarter of 2015 experienced an annual growth of 16.88% from 7.7 tons to 9 tons. Throughout 2015, the demand figure reached 38.9 tons [6]. From 2016 to 2019, the demand for gold in Indonesian jewellery continues to increase.

There are several previous studies that used the Random Forest and Single Exponential Smoothing methods for forecasting, some of which were carried out by Setyowati [7] namely comparing Exponential Smoothing and Moving Average methods to predict motor vehicle testing levies. The result of this study is that the Single Exponential Smoothing method is better used in predicting the retribution for testing motor vehicles with a MAPE value of 0.12%. Other research was also conducted by Supriyanto [8] namely comparing the KNN,

SVM, Decision Tree, Random Forest and Linear Regression methods. The results of this research indicate that Random Forest is the most optimal method with a MAPE value of 12%. Similar research was also carried out by Rianto and Yunis [9] to forecast the number of new students using the Random Forest method. The result of this research is that the MSE and MAE values are 0.02%, with an accuracy of 99.8%. Based on several previous studies, the Random Forest and Single Exponential Smoothing methods are more optimal than the comparison methods, so the purpose of this study is to compare the performance of these two methods in forecasting the demand for gold for Indonesian jewellery.

2 METHODOLOGY

The data used in this study is secondary data, namely data on the demand for gold for Indonesian jewellery for the period 2010 to 2021. The data is obtained from the World Gold Council (WGC) website which is accessed via https://gold.org. This study uses 2 variables, namely the time variable (quarterly) and the gold jewellery demand variable (tons). Analysis of the data in this study using the Random Forest and Single Exponential Smoothing methods using the R software.

2.1 Random Forest

Random Forest is a method that consists of a structured set of trees, each of which casts a unit of votes for the class and the results obtained according to the most decisions. The basic technique of Random Forest is the decision tree. Random Forest is a set of decision trees that are used for classifying and predicting data by entering input into the roots at the top and then down to the leaves at the bottom [10]. The results of the Random Forest analysis for classification are the mode of each tree of the built forest, while the prediction results are obtained from the average value of each tree [11].

The analysis stages using the Random Forest method are as follows:

1. The first stage in the Random Forest Analysis is to input quarterly data on gold jewellery demand from 2010 to 2021 into the R Software.
2. Splitting data into training and testing data. Identification of the Random Forest model with a predetermined ntree value (the number of trees) using training data, data testing is used to see the error rate of the model made.
3. Identify the model until you get the model with the smallest error.
4. Evaluation of the model and forecasting gold jewellery demand for data in the sample and for the next year's data using the best model with the smallest error.

2.2 Exponential Smoothing

Exponential Smoothing method is a moving average forecasting technique with weights where the data is weighted by an exponential function. Exponential smoothing is a moving average forecasting method with advanced weighting, but still easy to use [12]. This method has very little recording of past data or in other words this method pays more attention to the value of the most recent observations [7]. The analysis stages using the Single Exponential Smoothing method are as follows:

1. The first stage in the Single Exponential Smoothing Analysis is to input quarterly data on demand for gold jewellery from 2010 to 2021 into R Software.
2. The SES method uses data with a stationary pattern. A data can be said to be stationary if there is no growth, decline, and the data pattern is around a fixed average value $E(X_t) = \mu$ and the variance around the average remains $Var(X_t) = E(X_t - \mu_t)^2 = \sigma^2$ for a certain time [13].
3. It takes an alpha value, then it is tested for several alpha values, namely (alpha = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9) and the smallest error value is obtained using an alpha value of 0.4.
4. Forecasting the amount of gold demand for data in the sample and data for the next period using the best model (alpha = 0.4).

Single Exponential Smoothing calculation can be written in Equation 1, where F_{t+1} = Forecast for period *t+1*, X_t = Real value of period-*t,* α = The weight representing the smoothing constant (0< α<1), F_{t-1} = forecast for period *t-1*

$$F_{t+1} = \alpha X_t + (1-\alpha)F_{t-1} \quad (1)$$

2.3 Forecasting Error Measurement

Measurement of forecasting errors is done by comparing the results of forecasting with the reality that happened. A model is said to have good and accurate forecasting results if it generates a small error value [14]. The equation formula for finding the MAPE error value is in Equation 2, where Y_t= Actual data in period-*t,* F_t' = Forecasting data in period-*t,* n = Number of data used.

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - F_t'|}{Y_t} x100\% \quad (2)$$

3 RESULT

3.1 Descriptive Analysis

Based on ***Figure. 1***, Indonesia's gold jewellery demand in 2010-2020, where the highest demand was in 2018 with a value of 41.9 tons, an increase of 3.3 tons from the previous year (in 2017). According to data from the world gold council, Indonesia's gold demand in 2018 was ranked 6th as the country with the world's largest gold demand after Russia. Meanwhile, the lowest gold demand occurred in 2020, which was 20.9 tons. According to data from the World Gold Council, the decline in demand for gold in the jewellery sector experienced a significant decline in the 1st quarter of 2020. The decline in demand for gold jewellery in 2020 is not only in Indonesia but also in several other countries. This is because local gold prices have skyrocketed so that consumers are unable to absorb the existing supply of jewellery. And the main reason is the Covid-19 pandemic [15].

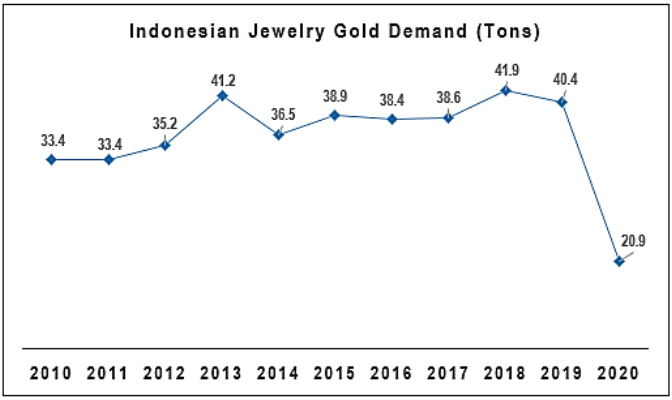

Figure 1. *Indonesian jewellery gold demand chart (tons) for the period 2010-2020.*

3.2 Forecasting Using Random Forest

Forecasting analysis using the Random Forest method requires training data and testing data. Training data is used to create a model, while testing data is used to see the accuracy of the model created. ***Table. 1.*** below shows the percentage as well as the amount of training and testing data after distribution.

Table 1. *Distribution of training and testing data*

Data	Percentage	Amount
Training	75%	35
Testing	25%	12

In forecasting using the Random Forest method, the training data is used to compare several tree values (the number of trees) in modelling. This is done before further forecasting analysis with the aim of seeing the best model based on the smallest error value which is then used in further analysis. Determination of the number of trees in this study based on the provisions of the researcher. ***Table. 2.*** is a comparison of error values based on several trees.

Table 2. *ntree and error values for Random Forest method*

ntree	MAPE	RMSE	MSE
25	13.280	1.328	1.764
50	12.866	1.287	1.656
100	12.883	1.288	1.659
150	13.000	1.300	1.690
200	13.014	1.301	1.693
250	12.917	1.292	1.669
300	13.136	1.314	1.727
500	13.180	1.318	1.737

Based on ***Table. 2.*** The smallest MAPE, RMSE and MSE values are 12.866, 1.287 and 1.656 respectively using a ntree value of 50. With an error value below 15%, it is included in the good category. Therefore, the Random Forest method can be used to predict the demand for gold in Indonesian jewellery. The following is a graph of forecasting results using the Random Forest method using a ntree value of 50.

Based on the graph in ***Figure. 2*** The data from the forecasting results are quite good in following the data pattern or not much different from the actual data pattern. The graph also shows the results of forecasting for each quarter in quartile IV/2021 to quartile IV/2022. As seen in the graph, the demand for gold for Indonesian state jewellery in the period IV/2021 to the next IV/2022 tends to decline.

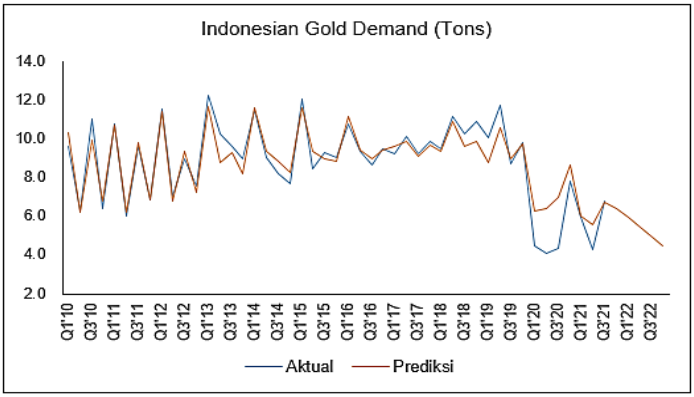

Figure 2. *Random Forest forecasting results graph.*

Table 3. *Forecast results for the next four quarters.*

Period	Gold Jewelery Request (tons)
4th Quarter 2021	6.4
1st Quarter 2022	6.0
2nd Quarter 2022	5.5
3rd Quarter 2022	5.0
4th Quarter 2022	4.5

Based on ***Table. 3*** The results of forecasting gold jewellery demand, gold jewellery demand for the 4th quarter of 2021 to the 4th quarter of 2022 continues to decline.

3.3 Forecasting Using Exponential Smoothing

Single Exponential Smoothing method is used in forecasting analysis on data that has a stationary pattern. In the time series analysis that has been done previously, the data on the number of requests for gold in Indonesian jewellery tends to have a stationary pattern. In performing the analysis using the SES method, first determine the smallest error value using several alpha values that are tested. The following is a table of alpha values used to determine the smallest error value.

Table 4. *The alpha value and error of the Single Exponential Smoothing method.*

ntree	MAPE	RMSE	MSE
0.1	21.953	2.044	4.178
0.2	20.860	1.968	3.873
0.3	20.678	1.949	3.799
0.4	20.626	1.971	3.885
0.5	20.751	2.019	4.076
0.6	21.178	2.090	4.368
0.7	21.885	2.179	4.748
0.8	22.765	2.289	5.240
0.9	23.990	2.420	5.856

Based on ***Table. 4*** obtained the smallest MAPE, RMSE, and MSE values with values of 20.626, 1.971 and 3.885, respectively, using an alpha value of 0.4. The MAPE value of 20.63% is in the descent or sufficient category. Therefore, the SES method can be used in predicting the amount of demand for Indonesian gold jewellery. The following is a graph of forecasting results using the SES method and an alpha value of 0.4.

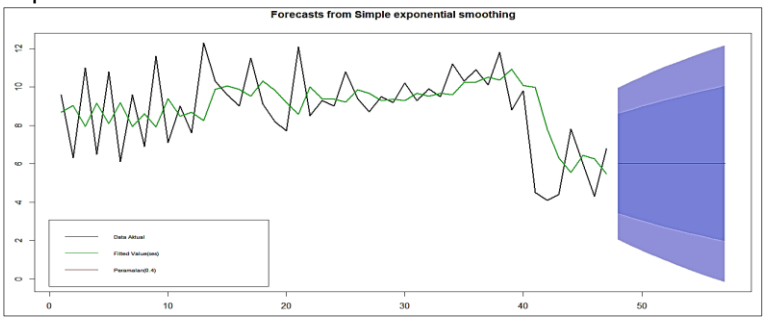

Figure 3*. Graph of forecasting results for Single Exponential Smoothing.*

From ***Figure. 3,*** The black line shows the actual data, and the green line shows the predicted data, while the Gray square that has a blue centre line is the forecast data for the next quarter. Forecasting data are sufficient to follow the data pattern or not much different from the

actual data pattern. In ***Figure. 3,*** the forecast for the next quarter is shown by the blue line in the middle of the square where the demand for gold jewellery for the next quarter tends to decline. Forecasting using the SES method is usually used for short-term forecasting, which is used for the next 1-3 months.

Table 5. *Forecasting results for the next period.*

Period	Gold Jewelery Request (tons)
4th Quarter 2021	6.0
1st Quarter 2022	6.0

Based on ***Table. 5***. the results of forecasting the amount of demand for gold jewellery for the next quarter or the 4th quarter of 2021, which is 6.0 tons. Forecasting using the SES method tends to have the same value for forecasting results because the SES method is used for short-term forecasting.

4 CONCLUSIONS

The highest demand for gold jewellery in Indonesia was in 2018 with a value of 41.9 tons, an increase of 3.3 tons from the previous year (in 2017). Meanwhile, the lowest gold demand occurred in 2020, which amounted to 20.9 tons. According to World Gold Council data, the decline in gold demand in the jewellery sector experienced a significant decline in the 1st quarter of 2020. The results of forecasting the demand for gold jewellery using the Random Forest method with 50 trees and the Single Exponential Smoothing method using alpha 0.4 both obtained gold demand for the 4th quarter of 2021 to the 4th quarter of 2022 which continued to decline. The results of forecasting the demand for gold jewellery using the Random Forest method with the number of trees, namely 50 and the Single Exponential Smoothing method using alpha 0.4, both show that gold demand for the 4th quarter of 2021 to the 4th quarter of 2022 continues to decline. The results of the analysis show that the Random Forest method produces a smaller forecast error value of 12.87% compared to the Single Exponential Smoothing method based on the MAPE value.

REFERENCES

[1] I. Aksan dan K. Nurfadilah, "Aplikasi Metode Arima Box-Jenkins Untuk Meramalkan Penggunaan Harian Data Seluler," *J. Math. Theory Appl.*, vol. 2, no. 1, hal. 5–10, 2020, [Daring]. Tersedia pada: https://bit.ly/3tKmDgy.

[2] S. I. Wardani, W. P. Nurmayanti, S. A. Wulandya, dan M. Gazali, "Implementation of Exponential Smoothing and Arima on Collection and Distribution Data of Zakat, Infaq and Shadaqah 2015-2020," 2021.

[3] N. L. W. S. R. Ginantra dan I. B. G. Anandita, "Penerapan Metode Single Exponential Smoothing Untuk Analisa Peramalan Penjualan Barang," *J. Sains Komput. Inform.*, vol. 3, no. 2, 2019, doi: http://dx.doi.org/10.30645/j-sakti.v3i2.1622951.

[4] M. J. Wulansari, "Analisis Faktor - faktor Yang Mempengaruhi Seseorang Terkena Penyakit Diabetes Melitus Menggunakan Regresi Random Forest," Universitas Islam Indonesia, 2018.

[5] I. N. S. Aguston, "Pengaruh Harga Emas Terhadap Pencapaian Laba Pada PT. Pegadaian (Persero) Cabang Takalar Sulawesi Selatan," Universitas Negeri Makasar, 2019.

[6] Y. Winarto, "Permintaan Emas Perhiasan Indonesia Capai 389 Ton. Diakses 1 Desember 2021", 2016, https://investasi.kontan.co.id/news/permintaan-emas-perhiasan-indonesia-capai-389-ton

[7] E. Setyowati, "Perbandingan Metode Exponential Smoothing Dan Moving Average Dalam Peramalan Retribusi Pengujian Kendaraan Bermotor Di Dinas Perhubungan Kota Blitar," *J Sains Dasar*, vol. 11, no. 1, hal. 35–38, 2022.

[8] H. Supriyanto, "Perbandingan Metode Supervised Learning Untuk Peramalan Time Series Pada Kunjungan Pasien Rawat jalan," *Junal SimanteC*, vol. 10, no. 2, hal. 67–76, 2022.

[9] M. Rianto dan R. Yunis, "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Random Forest," *Paradig. - J. Komput. dan Inform.*, vol. 23, no. 1, 2021, doi: 10.31294/p.v23i1.9781.

[10] Reinardus A. Haristu dan paulina H. P. Rossa, "Penerapan Metode Random Forest untuk Prediksi Win Ratio Pemain Player Unknown Battleground," *MEANS (Media Inf. Anal. dan Sist.*, vol. 4, no. 2, hal. 120–128, 2019, doi: 10.54367/means.v4i2.545.

[11] R. Dwi Lingga P., C. Fatichah, dan D. Purwitasari, "Deteksi Gempa Berdasarkan Data Twitter Menggunakan Decision Tree, Random Forest, dan SVM," *J. Tek. ITS*, vol. 6, no. 1, hal. 159–162, 2017, doi: 10.12962/j23373539.v6i1.22037.

[12] S. Fachrurrazi, "Peramalan Penjualan Obat Menggunakan Metode Single Exponential Smoothing Pada Toko Obat Bintang Geurugok," *Techsi*, vol. 6, no. 1, hal. 19–30, 2015.

[13] S. Nelfa, "Pendugaan Parameter Model Autoregressive Pada Deret Waktu", (Doctoral dissertation, Universitas Andalas), 2014.

[14] T. D. Andini dan P. Auristandi, "Peramalan Jumlah Stok Alat Tulis Kantor di UD Achmad Jaya Menggunakan Metode Double Exponential Smoothing," *J. Ilm. Teknol. Inf. Asia*, vol. 10, no. 1, hal. 1–10, 2016.