IRJEMS International Research Journal of Economics and Management Studies
Published by Eternal Scientific Publications
ISSN: 2583 – 5238 / Volume 4 Issue 5 May 2025 / Pg. No: 374-381
Paper Id: IRJEMS-V4I5P147, Doi: 10.56472/25835238/IRJEMS-V4I5P147

Original Article

Does Forecasting Using Regression Fare Better Than Other Methods?

[1]Andhalia Liza Marie, [2]Regi Muzio Ponziani*
[1,2]Departement of Accounting, Trisakti School of Management, Jakarta, Indonesia

Received Date: 20 April 2025 Revised Date: 10 May 2025 Accepted Date: 19 May 2025 Published Date: 29 May 2025

Abstract: *This research aims to find the best model to forecast the number of tourists visiting Indonesia among various methods, including multiple regression. The research period was from January 2017 until August 2024. Therefore, the period of the pandemic was covered in this research. The models tested were exponential smoothing, double exponential smoothing, multiple regression, and moving average. The research found that the moving average with the order of 2 is the best-performing model with an error percentage of around 6%. The regression model performs fairly well, with a MAPE of around 16%. The parameters in the regression are all significant, with an adjusted R-squared of 84.49%. This means we can control COVID in the regression model to have a MAPE score below 20%. The second and third best models are exponential smoothing with α of 0.8 and double exponential smoothing with α and β of 0.8. Since the MAPE for exponential and double exponential smoothing does not vary widely, we can infer that the trend component in exponential smoothing does not contribute much to the forecasting accuracy. Future research can include a more established regression, such as ridge regression or lasso regression, for forecasting purposes and compare their performance with machine learning/deep learning models. This will shed light on the efficacy of time series econometrics performance, especially compared to the recent machine learning/deep learning development.*

Keywords: *Double Exponential Smoothing, Exponential Smoothing, Forecasting, Moving Average, Regression.*

I. INTRODUCTION

Tourism has significantly contributed to the Indonesian economy (Azizurrohman, Hartarto, Lin, & Nahar, 2021). Foreign tourists will visit Indonesia and, during their stay, may spend some money that will boost the economic activity of the local area. Therefore, tourism could increase the local own-source revenue of the state government and increase the productivity thereof (Simorangkir, Ramadhan, Sukran, & Manalu, 2024). With every tourist arrival, foreign currency will be flowing into Indonesia, and the residents will provide goods and services to satisfy the needs of the tourists during their visits. Suhel & Bashir (2018) showed that tourists spending during their stay will provide income streams that boost the economy. Therefore, economic growth will be driven by the activity in the tourism sector (Saputra, Sartiyah, & Seftarita, 2022; Hsiao, Chiang, & N, 2024). Tourism has an immediate and significant effect on growth, andong-rung tourism also cointegrates with economic growth (Rasool, Maqbool, & Tarique, 2021). Thus, there will be corrections in the long run whenever economic growth deviates from its fundamental long-term value. The tourism sector provides this correction. Purnomo (2022) stated that, after performing a robustness check, both local and foreign tourists are positively influential toward economic growth since both contribute to the income of the surrounding people. Furthermore, tourism can alleviate unemployment by providing additional income through job creation in the creative and leisure sectors. However, it also brings externalities such as congestion, pollution, or environmental degradation (Naseem, 2021).

Through opening job opportunities, welfare and prosperity can increase, thus boosting the productivity of potential and existing tourist sites (Hutama & Suliswanto, 2023). This, in turn, can even reduce poverty through employment in the tourism sector (Alcala-Ordonez & Segarra, 2023). Nguyen et al. (2024) predicted that low-income economies might develop the tourism sector to increase economic growth and productivity, provided they make atonement to institutional reform. The will to attract more tourists will also prompt the government to increase the quality of the physical infrastructure (Rotar, Gricar, & Bojnec, 2023). Better access, communications, and transportation to tourist attractions may denote the readiness of local preparation to welcome tourists, which signifies tourism's role as a catalyst for further local and regional development (Ernawati, 2019). From the preceding exposition, we may surmise that knowing the information about tourists' arrival will help the government predict the economic growth and the additional income the residents would obtain. This study tries to find the best models to forecast the number of foreign tourists visiting Indonesia. The methodology employed would be exponential smoothing, double exponential smoothing, moving average, and regression that controls the occurrence of COVID-19. This research will shed light on whether controlling for COVID-19 occurrence may yield a powerful model to forecast the number of foreign tourists. The other methods (exponential smoothing, double exponential smoothing, and moving average) will not control the occurrence of

COVID-19 and, therefore, will act as the benchmarking models for the regression. Research regarding foreign tourists' arrival has been conducted quite extensively in Indonesian and international contexts. Rasul, Dewana, and Saed (2019) compared the forecasting performance of Local Linear Structure, Naïve, Holt, Random Walk, and ARIMA models to forecast tourist visits to the Kurdistan region. They found the Local Linear Structural Model and ARIMA(7,3,0) the best models. Anisa, Irawan, & Widiyanesti (2021) specifically put the Long-Short Term Memory (LSTM) model to the test. They found that this model could consistently be scored above 97% accuracy. In the context of Nepal, Upadhayaya (2021) found that ARIMA(1,1,1) performed best for the forecasting purpose. Nurhasanah, Salsabila, & Kartikasari (2022) employed the SARIMA model and found that ARIMA (1,0,0)(1,1,0) was the best forecasting model. Similarly, Ramadhani, Wahyuningsih, & Siringoringo (2022) also found that the ARIMA model performs best for forecasting tourists. Specifically, the ARIMA model is the ARIMA (0,1,1). Rianda & Usman (2023) compared the ARIMAX model with the intervention model for forecasting foreign tourists visiting Indonesia. They found that the intervention model is much more superior in this case. Kurtulay & Kizilirmak (2024) compared the performance of Naïve III, Moving Average, Double Moving Average, Seasonal Exponential Smoothing, and Artificial Neural Network for forecasting tourists in Turkey extensively. Their research showed that Artificial Neural Network was the best forecasting model. Paudel, Li, & Dhakal (2024) compared Winter Additive, Winter Multiplicative, SARIMA and Exponential Smoothing models in forecasting tourist arrivals in Nepal. They found Winter Multiplicative, the best-performing model due to its high accuracy. Khusna et al. (2024) used multiple input intervention ARIMA models to forecast. They found that SARIMA(1,1,1)(0,1,1) and SARIMA(0,1,1)(1,1,1) were the best models with the highest accuracy. In the Malaysian context, Abu et al. (2025) found that SARIMA(2,1,1)(0,1,1) was the most accurate model.

Since the previous research forecasting using regression is still scant, this research will compare the performance of regression forecasting. However, this research will have a period of the COVID-19 pandemic as a proxy for turbulent times. Therefore, there will be a variable that controls this period. Subsequently, the regression performance will be compared with another method to know its efficacy.

II. METHODOLOGY

This research attempts to forecast foreign tourists visiting Indonesia. The tourist data extends from January 2017 until August 2024. The data will be divided into two categories: training data and test data. Training data range from January 2017 until December 2023. Test data are from January 2024 until August 2024. Training data will be used to generate models that, in turn, predict the total number of foreign tourists visiting Indonesia from January 2024 until August 2024. The prediction yielded will then be compared to determine the models that generate the most accurate predictions. The parameter for accuracy used in this research is MAPE (mean absolute percentage error). This number will be derived from averaging each month's absolute error percentage generated. MAPE is formulated mathematically as follows:

$$\mathbf{MAPE = (\textstyle\sum |A_t - F_t|/A_t)/t}$$

Where $\mathbf{A_t}$ is the actual value of tourists for the month t, and $\mathbf{F_t}$ is the predicted value of the number of tourists. By subtracting F_t from A_t we will get the error of the prediction. Dividing the error by the actual value will get us the error percentage. We will sum the error for January 2024 until August 2024 and divide the result by 8 (the number of months from January 2024 until August 2024) to determine the percentage error. Several forecasting models are used in this research. The first model will be an exponential smoothing model. This model uses the actual and predicted values to forecast the number of tourists for the next period (next month). The model is formalized as follows:

$$\mathbf{F_{t+1} = \alpha A_t + (1 - \alpha)\ Ft}$$

Where α = 0.1, 0.2 ,..., 0.9. The α will vary from 0.1 to 0.9. A_t is the actual value of tourists visiting Indonesia while $\mathbf{F_t}$ is the forecast of the number of tourists. The lower the α, the more the forecast for the next period will depend on the forecast of this period. Therefore, this period's actual value will not contribute much to the forecast for the next period. The higher the α, the more contribution from the actual value of the number of tourists to the forecast for the next period. The second model we use is the double exponential smoothing. This model posits that the forecast for the next period consists of the level and the trend. The formal formulation is as follows:

$$\mathbf{F_{t+1} = a_t + b_t}$$
$$\mathbf{a_t = \alpha\ (A_{t-1} + b_{t-1}) + (1\text{-}\alpha)\ F_{t-1}}$$
$$\mathbf{b_t = \beta\ b_{t-1} + (1 + \beta)\ (a_t - A_{t-1})}$$

Like the previous explanation, $\mathbf{A_{t-1}}$ is the actual value of the last period, while $\mathbf{F_{t-1}}$ is the forecast for the last period. A_t is the level term, while $\mathbf{i_t}$ is the trend term. $\mathbf{A_{t-1}}$, $\mathbf{b_{t-1}}$ and $\mathbf{F_{t-1}}$ derive at the level term, $\mathbf{a_t}$. Therefore, double exponential smoothing can identify the trend in the data and use it for prediction. The third model is the regression model. The regression model will have the trend and dummy variable COVID as the dummy variable. This dummy variable will control extraordinary events like the COVID-19 pandemic in the forecasting model. The formal regression model is as follows:

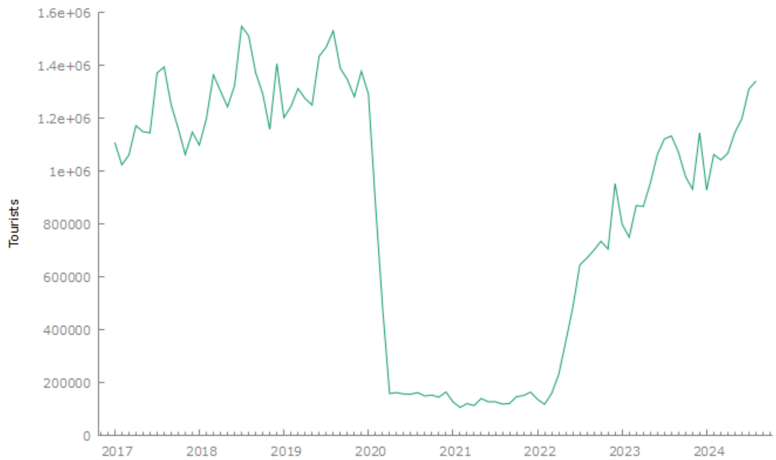
$$\mathbf{TOURISM = \alpha + \beta_1\, TREND + \beta_2\, COVID + \varepsilon}$$

The TREND variable will include 1, 2, and... to indicate the research time. By including the TREND variable, we can capture any pattern emerging from the data. The last model is the moving average model. This model assumes that the forecast for the next period results from the average of some past values. This research will employ the average of 2 past values, 3 past values, 4 past values, and 5 past values. The more past values used for averaging, the smoother the forecast will be. The formula for the moving average is as follows:

$$\mathbf{MA\ (2) = (a_{t-1} + a_{t-2})\ /\ 2}$$
$$\mathbf{MA\ (3) = (a_{t-1} + a_{t-2} + a_{t-3})\ /\ 3}$$
$$\mathbf{MA\ (4) = (a_{t-1} + a_{t-2} + a_{t-3} + a_{t-4})\ /4}$$
$$\mathbf{MA\ (5) = (a_{t-1} + a_{t-2} + a_{t-3} + a_{t-4} + a_{t-5})\ /\ 5}$$

III. RESULTS AND DISCUSSIONS

The time-series plot of the number of tourists from January 2017 until August 2024 is depicted in the figure 1 below:



The trend is clearly visible from the beginning of the research period until early 2020. The seasonality is also obvious. There are times in a year when tourists' visits are at the peak and when it is at the lowest point. The number of tourists, however, was increasing before early 2020. Then, the pandemic struck. Tourists were not allowed to enter Indonesia except for a small amount for special reasons. Therefore, from early 2020 until early 2022, the number of tourists declined abruptly and steeply. Begining of the second quarter of 2022, the government lifted the restriction to enter Indonesia again. The number of tourists began its traction upward. In a short time, 600,000 tourists enter Indonesia around May and June 2022. 2 months later, the number increased to 800,000 tourists. Seasonality also begins to show itself since most tourists choose a holiday for their vacation time. The trend is also increasing over time. This resembles the situation before the pandemic. The upward trend will continue until the end of the research period. The number of tourists has begun to recover and has surpassed the level in 2017. The tables show the forecasts and MAPE using the various methods explained previously

Table 1. Forecasting Results using Exponential smoothing

Time	Real Data	A								
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
January 2024	927,746.00	858,814.54	1,000,468.26	1,038,770.27	1,052,795.85	1,062,527.13	1,073,649.01	1,087,685.24	1,104,595.29	1,123,781.20
February 2024	1,062,149.00	865,707.69	985,923.81	1,005,462.99	1,002,775.91	995,136.57	986,107.21	975,727.77	963,115.86	947,349.52
March 2024	1,041,861.00	885,351.82	1,001,168.85	1,022,468.79	1,026,525.15	1,028,642.78	1,031,732.28	1,036,222.63	1,042,342.37	1,050,669.05
April 2024	1,066,958.00	901,002.74	1,009,307.28	1,028,286.46	1,032,659.49	1,035,251.89	1,037,809.51	1,040,169.49	1,041,957.27	1,042,741.81
May 2024	1,145,499.00	917,598.26	1,020,837.42	1,039,887.92	1,046,378.89	1,051,104.95	1,055,298.61	1,058,921.45	1,061,957.85	1,064,536.38
June 2024	1,197,941.00	940,388.34	1,045,769.74	1,071,571.24	1,086,026.94	1,098,301.97	1,109,418.84	1,119,525.73	1,128,790.77	1,137,402.74
July 2024	1,310,756.00	966,143.60	1,076,203.99	1,109,482.17	1,130,792.56	1,148,121.49	1,162,532.14	1,174,416.42	1,184,110.95	1,191,887.17
August 2024	1,339,946.00	1,000,604.84	1,123,114.39	1,169,864.32	1,202,777.94	1,229,438.74	1,251,466.45	1,269,854.13	1,285,426.99	1,298,869.12

Table. 2 MAPE (Mean Absolute Percentage Error) of Exponential smoothing

Time	MAPE								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
January 2024	0.0743	0.0784	0.1197	0.1348	0.1453	0.1573	0.1724	**0.1906**	0.2113
February 2024	0.1849	0.0718	0.0534	0.0559	0.0631	0.0716	0.0814	**0.0932**	0.1081
March 2024	0.1502	0.0391	0.0186	0.0147	0.0127	0.0097	0.0054	**0.0005**	0.0085
April 2024	0.1555	0.0540	0.0362	0.0321	0.0297	0.0273	0.0251	**0.0234**	0.0227
May 2024	0.1990	0.1088	0.0922	0.0865	0.0824	0.0787	0.0756	**0.0729**	0.0707
June 2024	0.2150	0.1270	0.1055	0.0934	0.0832	0.0739	0.0655	**0.0577**	0.0505
July 2024	0.2629	0.1789	0.1536	0.1373	0.1241	0.1131	0.1040	**0.0966**	0.0907
August 2024	0.2532	0.1618	0.1269	0.1024	0.0825	0.0660	0.0523	**0.0407**	0.0307
Average	**18.69%**	**10.25%**	**8.83%**	**8.21%**	**7.79%**	**7.47%**	**7.27%**	**7.20%**	**7.41%**

Table 3

Forecasting Result Using Double Exponential Smoothing

Time	Real Data	forecast (a = 0.1, b=0.1)	forecast (a = 0.2, b=0.2)	forecast (a = 0.3, b=0.3)	forecast (a = 0.4, b=0.4)	forecast (a = 0.5, b=0.5)	forecast (a = 0.6, b=0.6)	Forecast (a = 0.7, b=0.7)	forecast (a = 0.8, b=0.8)	forecast (a = 0.9, b=0.9)
January 2024	927,746	847,803	1,000,425	1,038,770	1,052,796	1,062,527	1,073,649	1,087,685	1,104,595	1,123,781
February 2024	1,062,149	855,696	985,888	1,005,463	1,002,776	995,137	986,107	975,728	963,116	947,350
March 2024	1,041,861	875,430	1,001,134	1,022,469	1,026,525	1,028,643	1,031,732	1,036,223	1,042,342	1,050,669

April 2024	1,066,958	891,982	1,009,278	1,028,286	1,032,659	1,035,252	1,037,810	1,040,169	1,041,957	1,042,742
May 2024	1,145,499	908,659	1,020,810	1,039,888	1,046,379	1,051,105	1,055,299	1,058,921	1,061,958	1,064,536
June 2024	1,197,941	932,261	1,045,747	1,071,571	1,086,027	1,098,302	1,109,419	1,119,526	1,128,791	1,137,403
July 2024	1,310,756	958,091	1,076,182	1,109,482	1,130,793	1,148,121	1,162,532	1,174,416	1,184,111	1,191,887
August 2024	1,339,946	993,284	1,123,096	1,169,864	1,202,778	1,229,439	1,251,466	1,269,854	1,285,427	1,298,869

Table. 4 MAPE (Mean Absolute Percentage Error) Of Double Exponential Smoothing

Time	MAPE								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
January 2024	0.086169	0.078339	0.119671	0.134789	0.145278	0.157266	0.172396	**0.190623**	0.2113
February 2024	0.194373	0.071799	0.053369	0.055899	0.063091	0.071592	0.081365	**0.093238**	0.1081
March 2024	0.159744	0.039090	0.018613	0.014720	0.012687	0.009722	0.005412	**0.000462**	0.0085
April 2024	0.163995	0.054060	0.036245	0.032146	0.029716	0.027319	0.025107	**0.023432**	0.0227
May 2024	0.206757	0.108851	0.092197	0.086530	0.082404	0.078743	0.075581	**0.072930**	0.0707
June 2024	0.221780	0.127047	0.105489	0.093422	0.083175	0.073895	0.065458	**0.057724**	0.0505
July 2024	0.269055	0.178961	0.153556	0.137297	0.124077	0.113083	0.104016	**0.096620**	0.0907
August 2024	0.258714	0.161835	0.126932	0.102368	0.082471	0.066032	0.052309	**0.040687**	0.0307
Average	**19.5073%**	**10.2498%**	**8.8259%**	**8.2146%**	**7.7863%**	**7.4707%**	**7.2705%**	**7.1965%**	**7.4137%**

Table 5 Forecast Results and MAPE of Regression

Time	Real Data	Forecast	MAPE
January 2024	962,305	927,746	0.04
February 2024	958,327	1,062,149	0.10
March 2024	954,349	1,041,861	0.08
April 2024	950,371	1,066,958	0.11
May 2024	946,393	1,145,499	0.17
June 2024	942,415	1,197,941	0.21
July 2024	938,437	1,310,756	0.28
August 2024	934,459	1,339,946	0.30
MAPE			16%

Table 6 Forecast Results of Moving Average

Time	Data Real	Moving Average Order			
		MA(2)	**MA(3)**	**MA(4)**	**MA(5)**
January 2024	927,746.00	1037884	1018096	1031133	1051434
February 2024	1,062,149.00	1036144	1001172	995508.5	1010456
March 2024	1,041,861.00	994947	1044812	1016416	1008837
April 2024	1,066,958.00	1052005	1010585	1044075	1021505
May 2024	1,145,499.00	1054409	1056989	1024679	1048651
June 2024	1,197,941.00	1106228	1084773	1079117	1048843
July 2024	1,310,756.00	1171720	1136799	1113065	1102882
August 2024	1,339,946.00	1254348	1218065	1180289	1152603

Table 7 MAPE of Moving Average

Time	MAPE			
	MA (2)	**MA (3)**	**MA (4)**	**MA (5)**
January 2024	**0.118716**	0.097387	0.111439	0.133321
February 2024	**0.024483**	0.057409	0.062741	0.048669
March 2024	**0.045029**	0.002833	0.024423	0.031698
April 2024	**0.014015**	0.052835	0.021447	0.042601
May 2024	**0.079519**	0.077267	0.105474	0.084546
June 2024	**0.076558**	0.094469	0.09919	0.124462
July 2024	**0.106073**	0.132715	0.150822	0.158591
August 2024	**0.063881**	0.090959	0.119152	0.139814
Average	**6.6034%**	7.5734%	8.6836%	9.5463%

Table 1 shows the forecast results using exponential smoothing. This method does not recognize any trend or seasonality. The feedback to the model consists of the actual value of the previous period and the forecast. There will be weight for each lagged actual value and forecast. This weight is denoted by α. This research varies the value of α from 0.1 to 0.9. The higher the value of the α, the more the contribution of the lagged actual value to the current period forecast (and the less contribution by the lagged forecast value). For January 2024, the higher the value of α, the higher the forecast. In fact, only at α = 0.1 is the forecast value less than 1 million.

The other values of α yield forecasts of more than 1 million. In February 2024, the α of 0.3 and 0.4 generated the most forecasts. They are indeed the closest to the actual value for February 2024, which is 1.06 million. The forecast generated by α 0.1 is the least accurate for the month. The exponential model using α of 0.1 consistently undershoots the actual data. Apart from January 2024, it always produces forecasts below the actual value. This is an indication of the low accuracy of the model. Hence, a small component of the actual lag value cannot contribute much to the forecast accuracy. Later, in Table 2, we will see that the MAPE of α 0.1 is the highest. This denotes the magnitude of deviation of the forecast value from the actual value. We can infer that the actual value of the number of tourists depends significantly on the lagged value, a property of time series data. Whenever the tourist's arrival in a certain month is high, we can expect the arrival for the next month to be more or less high and vice versa. In Table 1, however, all the exponential smoothing models always underpredict the number of tourists. Therefore, it is wiser to adjust the forecasts of tourists whenever we use the exponential smoothing model. Qualitatively, we may add up some adjusting numbers to cope with the underprediction issue.

Table 2 shows that the MAPE will decrease when we increase the value of α. However, the value of the MAPE will rebound after α equals 0.8. Hence, α of 0.8 has the lowest MAPE among all the models and, therefore, is the most accurate method. Large α indicates the pivotal role of the actual number from the previous period. The second-best model is the one whose α is 0.7, followed by 0.9. The error rate for exponential smoothing with α 0.8 is 7.2 %, while the one for α 0.7 is 7.27%. These two models do not differ very much in terms of accuracy. The model with α 0.9 scores the MAPE of 7.41%. The worst-performing model for exponential smoothing is the α of 0.1. The MAPE is 18. 69%. Therefore, it is recommended to use an exponential smoothing model that relies more on the lagged actual value.

Tables 3 and 4 report the results and MAPE of double exponential smoothing. Double exponential smoothing considers the data's trend beside the data's level. Therefore, it is posited that this model will yield better results than just exponential smoothing. There are two parameters in double exponential smoothing, namely α and β. This research does not see many combinations of α and β. The α and β used are 0.1, 0.2, and 0.9 for both parameters. The forecast results are compared with the actual data to derive their accuracies. When we consider the trend factor in the exponential smoothing, the results are pretty similar. The worst performer in exponential smoothing is still the worst performer in double exponential smoothing. Again, the best performer in exponential smoothing is the best performer in double exponential smoothing. The lower the value of α and β, the higher the MAPE will be. For January, higher α and β will generate higher results. Therefore, the forecast using α 0.1 and β 0.1 is the lowest, while α 0.9 and β 0.9 is the highest. For February, the pattern of forecasts resembles that of exponential smoothing. For α 0.3 and β 0.3 and α 0.4 and β 0.4, the forecasts are the highest among all. These forecasts are the closest to the actual value. All the other models yield forecasts of below than 1 million tourists. These two models generate forecasts of 1.005 million and 1.002 million tourists. The results are very similar to those of exponential smoothing. The value of α and β of 0.8 scores the lowest MAPE. Therefore, it is the most accurate. The second best is also α and β of 0.7. This resembles the results of exponential smoothing in which α 0.7 is the second best. The first model, α and β of 0.1 performs worst. The MAPE is around 19%. This is even bigger than the exponential smoothing in which the α is 0.1. Overall, the result resembles and is comparable to the exponential smoothing model. Therefore, we posit that exponential smoothing is enough instead of double exponential smoothing. Research concerning the number of tourists arriving visiting Indonesia could employ exponential smoothing since the MAPE is very good that is around 7%.

Table 5 presents the forecast results of regression. The regression performs very well in January and March. In January, the forecast error is only 4%, while in March is 8%. Beginning in April, the performance deteriorates until it reaches its worst performance in August, with an error of 30%. In February, the forecast error was 10%, which is still relatively good. These forecast results are derived from a regression model that involves trend and COVID, a dummy variable to control the incidence of COVID (1 if COVID still plagues the period, 0 if not). The values of regression parameters are as follows:

TOURISTS =	**1,300,436 -**	**3,978 TREND*** -**	**938,269 COVID*****
		(760.47)	**(44,852.12)**

The above regression coefficients are all significant at 1%. Overall, the trend is negative. It looks like the recovery after the pandemic is not enough to turn back tourist arrivals to the level before the pandemic. However, the decreasing trend is not worrying because it tends to almost reach a flat trend. The decreasing trend is around 3,978 number of tourists decreases every month. As a percentage of around 1 million tourists visit per month, the proportion is relatively small. The pandemic also contributes significantly to the decrease in tourist numbers. The total lockdown during the pandemic will decrease the number of tourists by 938,269. This are practically almost zero tourists visiting Indonesia. The regression model is fit (the p-value for the F test is 0,000), and the adjusted R-square is 84.49%. This means 84.49% of variances in the number of tourists can be explained by the TREND and COVID variables. The model constructed is very good for forecasting. However, the MAPE for the regression model is 16%. The regression model does not perform better than exponential smoothing and exponential smoothing for forecasting purposes. Nonetheless, we have shown and proven that TREND and COVID variables are enough to construct a regression model for forecasting tourist visits to Indonesia. Table 6 presents the forecast results of the moving average model. For January 2024, all models overshoot the actual data. The number of tourists is 927,746, while the forecasts generated are above 1 million. Interestingly, for February 2024, all the models undershoot the actual number of tourists with the least accurate forecast generated by MA(2). For March 2024, only MA(3) overshoots the actual data, while all the others undershoot. Regarding the last month of prediction, i.e. August 2024, only MA(2) yield the forecast closest to the actual data, while MA(5) is the least accurate.

Table 7 presents the accuracy measures of each MA model. We can see that MA(2) has the best accuracy. The MAPE is 6.6%. This is even better than exponential and double exponential smoothing, whose best-performing models hover around 7%. The higher the order of the moving average, the higher the MAPE. This means the forecast of tourists' arrival need not be smoothed out. The less smoothed, the more accurate the forecast. This explains why MA(2) performs very well. MA(2) is the last smoothed since it yields a forecast from the average of two previous data. Although the forecast is more volatile, the probability of it being accurate is high. In fact, for the April forecast, the error rate is only around 1%. No model has achieved this before.

IV. CONCLUSION

This research aims to find the best model to forecast the number of tourists visiting Indonesia. The research period was from January 2017 until August 2024. Therefore, the period of the pandemic was covered in this research. This will test the ability of each model to cope with high uncertainty and turbulence in forecasting periods. The models tested were exponential smoothing, double exponential smoothing, multiple regression, and moving average. The research found that the moving average with the order of 2 is the best-performing model. This indicates that a smoothed-out model is unsuitable for forecasting the number of tourists. Moving average with the order of 2 is not a very smooth model. Hence, it allows for some volatilities in the forecast that can capture the volatility in the actual data.

The second and third best models are exponential smoothing with α of 0.8 and double exponential smoothing with α and β of 0.8. Since the MAPE for exponential and double exponential smoothing does not vary widely, we can infer that the trend component in exponential smoothing does not contribute much to the forecasting accuracy. The high value of α also shows the very characteristic of time-series data. The value for the next period does not differ very much from the current value. Therefore, we cannot expect a sudden increase or decrease while forecasting the number of tourists. Therefore, using high α is recommended to make future forecasts that do not diverge drastically from the current data. The last method is the multiple regression method. This method yields a MAPE of 16%. Although it is not the best performer, its accuracy is still very good because the MAPE is below 20%. This research found that the independent variables in the regression must include a trend element and a variable to control the pandemic. Therefore, two variables are enough. The adjusted R square of the regression model is very high, and the model is also fit. This means the independent variable chosen contributes so much to the value of the dependent variable. Future research can include a more established regression, such as ridge or lasso regression.

V. REFERENCES

[1] Abu, N., Taib, S. A., Zainal, N. A., Ramli, N. A., & Go, C. K. (2025). Time Series Forecasting for the Tourism Industry in Malaysia. *Advances and Applications in Statistics, 92*(1), 77-87. doi:https://doi.org/10.17654/0972361725004

[2] Alcala-Ordonez, A., & Segarra, V. (2023). Tourism and Economic Development: A Literature Review to Highlight Main Empirical Findings. *Tourism Economics, 31*(1), 76-103. doi:https://doi.org/10.1177/13548166231219638

[3] Anisa, M. P., Irawan, H., & Widiyanesti, S. (2021). Forecasting Demand Factors of Tourist Arrivals in Indonesia's Tourism Industry Using Recurrent Neural Network. *IOC Conf. Series: Materials Science Engineering, 1077*. doi:10.1088/1757-899X/1077/1/012035

[4] Azizurrohman, M., Hartarto, R. B., Lin, Y. M., & Nahar, F. H. (2021). The Role of Foreign Tourists in Economic Growth: Evidence from Indonesia. *Jurnal Ekonomi & Studi Pembangunan, 22*(2), 313-322. doi:https://journal.umy.ac.id/index.php/esp/article/view/11591

[5] Ernawati, D. P. (2019). Development of the Tourism Industry as the Motor of Economic Growth in Indonesia. *International Journal of Science and Society (IJSOC), 1*(4), 145-153. doi:https://doi.org/10.54783/ijsoc.v1i4.30

[6] Hsiao, W. C., Chiang, H. C., & N, A. (2024). The Study of Economic Growth and Tourism Industry in Thailand. *European Journal of Development Studies, 4*(4), 62-68. doi:https://doi.org/10.24018/ejdevelop.2024.4.4.370

[7] Hutama, D. A., & Suliswanto, M. S. (2023). The Influence of the Tourism Sector on Economic Growth and Labor Absorption in the Province of Bali. *Journal of Indonesian Tourism, Hospitality and Recreation, 6*(1), 15-26. doi:https://doi.org/10.17509/jithor.v6i1.54589

[8] Khusna, H., Mashuri, M., Ahsan, M., Wibawati, W., Aksioma, D. F., & Suhermi, N. (2024). Forecasting Number of International Tourist Arrivals Using Multi-Input Intervention ARIMA Model. *BAREKENG: Journal of Mathematics and Its Application, 18*(3), 1539-1549. doi:https://doi.org/10.30598/barekengvol18iss3pp1539-1548

[9] Kurtulay, Z., & Kizilirmak, I. (2024). Forecasting Turkiye's International Tourism Demand. *Journal of Global Tourism and Technology Research, 5*(1), 12-26. doi:10.54493/jgttr.1408566

[10] Naseem, S. (2021). The Role of Tourism in Economic Growth: Empirical Evidence from Saudi Arabia. *Economies, 9*(3). doi:https://doi.org/10.3390/economies9030117

[11] Nguyen, P. C., Schinckus, C., Chong, F. H., Nguyen, B. Q., & Tran, D. L. (2025). Tourism and Contribution to Employment: Global Evidence. *Journal of Economics and Development, 27*(1), 22-37. doi:https://doi.org/10.1108/JED-07-2024-0269

[12] Nurhasanah, D., Salsabila, A. M., & Kartikasari, M. D. (2022). Forecasting International Tourist Arrivals in Indonesia Using the SARIMA Model. *International Journal of Statistics and Data Science, 2*(1), 19-25. doi:https://doi.org/10.20885/enthusiastic.vol2.iss1.art3

[13] Paudel, T., Li, W., & Dhakal, T. (2024). Forecasting Tourist Arrivals in Nepal: A Comparative Analysis of Seasonal Models and Implications. *Journal of Statistical Theory and Applications, 23*, 206-223. doi:https://doi.org/10.1007/s44199-024-00079-7

[14] Purnomo, S. D. (2022). The Effect of Tourism on Economic Growth: Empirical Study in Eastern Indonesia. *E-Jurnal Ekonomi dan Bisnis Universitas Udayana, 11*(8), 959-968.

[15] Ramadhani, A., Wahyuningsih, S., & Siringoringo, M. (2022). Forecasting the Number of Foreign Tourist Visits to Indonesia Using Intervention Analysis with Step Function. *Jurnal Matematika, Statistika dan Komputasi, 19*(1), 146-162. doi:https://doi.org/10.20956/j.v19i1.21607

[16] Rasool, H., Maqbool, S., & Tarique, M. (2021). The Relationship Between Tourism and Economic Growth Among BRICS Countries: A Panel Cointegration Analysis. *Future Business Journal, 7*.

[17] Rianda, F., & Usman, H. (2023). Forecasting Tourism Demand DUring the COVID Pandemic: ARIMAX and Intervention Modelling Approaches. *BAREKENG: Journal of Mathematics and Its Applications, 17*(1), 285-294. doi:https://doi.org/10.30598/barekengvol17iss1pp0285-0294

[18] Rotar, L. J., Gricar, S., & Bojnec, S. (2023). The Relationship Between Tourism and Employment: Evidence from the Alps-Adriatic Country. *Economic Research-Ekonomska Istrazivanja, 36*. doi:https://doi.org/10.1080/1331677X.2022.2080737

[19] Saputra, Z., Sartiyah, S., & Seftarita, C. (2022). The Impact of Tourism on Economic Growth: Evidence from Aceh Province, Indonesia. *International Journal of Global Optimization and Its Application, 1*(3), 223-228. doi:https://doi.org/10.56225/ijgoia.v1i3.74

[20] Simorangkir, C. O., Ramadhan, G., Sukran, M. A., & Manalu, T. (2024). Tourism Development Impact on Economic Growth and Poverty Alleviation in West Java. *Jurnal Kepariwisataan Indonesia, 18*(2), 175-196. doi:https://doi.org/10.47608/jki.v18i22024.175–196

[21] Suhel, S., & Bashir, A. (2018). The Role of Tourism Toward Economic Growth in the Local Economy. *Economic Journal of Emerging Markets, 10*(1), 32-39. doi:10.20885/ejem.vol10.iss1.art4

[22] Upadhayaya, R. P. (2021). Forecasting International Tourists Arrival to Nepal Using Autoregressive Integrated Moving Average (ARIMA). *Janapriya Journal of Interdisciplinary Studies, 10*(1), 107-117. doi:https://doi.org/10.3126/jjis.v10i01.42614