Prediction of Accident Numbers in Yogyakarta using The Single Moving Average Method

Rizcki Dwi Darmawan[1)], Zakaria Yahya Zudityas[2)], Diaz Yudistira Ardhani[3)], Muhammad Hamid Abdullah[4)]

[1-4)] Departement of Informatics, University of Islamic Majapahit

Correspondence Author: riz.kie720@gmail.com

Article Info :

Article History :

Received :

05 July 2022

Revised :

16 July 2022

Accepted :

28 July 2022

Available Online :

31 August 2022

Keyword :

Prediction, Time Series, Single Moving Average, Root Mean Square Error, Accidents

ABSTRACT

The soaring birth rate among Indonesians has led to an increase in the number of private vehicles which has led to the rise in the number of accidents. The purpose of this article is to predict the number of accidents using the Single Moving Average technique. This article uses data from the Yogyakarta government's My Data web application from 2010 to 2021 in annual time brackets. The prediction model is checked using RMSE, MAD, and MAPE. Using the Single Moving Average technique gets an error value for RMSE of 643.99, MAD of 494.7, and MAPE of 1036.17%.

1. INTRODUCTION

The current data shows that traffic accidents continue to occur frequently, not only in villages and sub-districts but also in big cities. The number of accidents is concerning, with the potential for above 100,000 fatalities. The causes of accidents vary and can be attributed to factors such as driver negligence, overloaded vehicles, or disruptive passengers. It is crucial to prioritize safety when driving to prevent accidents. This includes wearing helmets and jackets while riding a motorbike, using seat belts when driving a car, being attentive to traffic signs, and refraining from running red lights.

To reduce the occurrence of accidents, it's important to stay alert and cautious while travelling on highways. It's estimated that 1.5 million people die from accidents annually, with approximately 50 million sustaining injuries, according to WHO data. Developing countries, such as Indonesia, bear the brunt of these statistics due to the high number of motorized vehicles. Efforts are necessary to implement measures that reduce the number of deaths resulting from accidents.

ISSN Print : 1979-7141
ISSN Online : 2541-1942

Traffic accidents pose a significantly higher risk compared to other types of accidents, especially on roads in Yogyakarta. To address this issue, a traffic accident rate prediction technique has been developed specifically for the city. This technique, based on the Single Moving Average method, aims to identify potential accident-prone locations. Its purpose is to aid in reducing the number of accidents and ultimately minimize fatalities on the city's roads.

2. METHOD

2.1. Method of collecting data.

Our research uses data collection methods and prediction methods. This data collection method is carried out by taking data from the Yogyakarta government's Dataku web application which can be accessed on the page: http://bappeda.jogjaprov.go.id/dataku/.

2.2. Prediction Method.

Time Series is a collection of observational data taken according to time indicators sequentially at fixed time intervals. The time series analysis technique is a statistical strategy used to predict the structure of possible situations that may occur in the future in decision-making (Faishol et al., 2017).

Single Moving Average (Single Moving Average) is a technique for predicting one or several data by finding the average value as a prediction for the future period from a group of observation data (Rachman, 2018). The systematic equation of the Single Moving Average is:

$$Ft = \frac{At + At_{-1} + At_{-2} + \ ... \ + At_{-n}}{n}$$

Information:

Ft = Moving average for the t-th time period

At = Real data in the t-th time period

n = Moving average period limit

The way to calculate using this technique is that every time you get the latest actual data, the old period data will be removed, and the new period data will be entered to get the latest average data. The latest average is used to predict the next period.

2.3. Prediction Model Error.

Root Mean Square Error (RMSE) is the high percentage error of a prediction model, prediction results will be more precise if the RMSE measurement value is smaller (Suprayogi et al., n.d.). This model is used if there is an error in the Data Prediction. The RMSE systematic equation is:

$$RMSE = \sqrt{\frac{\sum(At - Ft)^2}{n}}$$

Information:

At = actual (real) data

Ft = Prediction data

n = Number of data

Mean Absolute Deviation (MAD) is the average value of absolute error over a specific period of time without looking at whether the prediction results are larger or smaller than the actual value, MAD is the average value of the absolute deviation value (Saputro & Purwanggono, n.d.) . The MAD systematic equation is:

ISSN Print : 1979-7141
ISSN Online : 2541-1942

$$MAD = \frac{\sum|At - Ft|}{n}$$

Mean Absolute Percentage Error (MAPE) is the average absolute error over a specific time period and then the result is multiplied by 100% to produce a percentage value (Saputro & Purwanggono, n.d.). The MAPE systematic equation is:

$$MAPE = \sum \left|\frac{At - Ft}{At}\right| \times 100\%$$

3. RESULTS AND DISCUSSION.

3.1. Presentation and Visualization of Actual Data.

After the data is processed it is displayed in the form of a data presentation. The results of the data presentation and visualization are below. Data taken from the dataku web application from the page: http://bappeda.jogjaprov.go.id/dataku/ will be presented in the following table.

Table 1. Presentation of Actual Data

Year	Accident Rate (Actual)
2010	4,613
2011	4,508
2012	4,457
2013	3,631
2014	3,472
2015	4,313
2016	3,777
2017	4,011
2018	5,061
2019	5,944
2020	4,559
2021	5,350

From these figures, the accident number data (actual) can be visualized using line and bar graphs as below.

ISSN Print : 1979-7141
ISSN Online : 2541-1942

Figure 1. Line graph of actual data on accident numbers

Figure 2. Bar graph of actual data on accident numbers

3.2. Calculation of Predicted Values.

This prediction data calculation uses the Single Moving Average technique and the data used is actual accident data in Yogyakarta, predictions using a 5 year period. The calculation process is as follows.

Predictions for 2014:

$$\mathbf{Ft_{-5}} = \frac{\mathbf{4,613 + 4,508 + 4,457 + 3,631 + 3,472}}{\mathbf{5}}$$

$$\mathbf{Ft_{-5}} = \mathbf{4,136.20}$$

ISSN Print : 1979-7141
ISSN Online : 2541-1942

Predictions for 2015:

$$\mathbf{Ft_{-6}} = \frac{\mathbf{4,508 + 4,457 + 3,631 + 3,472 + 4,313}}{\mathbf{5}}$$

$$\mathbf{Ft_{-6}} = \mathbf{4,076.20}$$

Predictions for 2016:

$$\mathbf{Ft_{-7}} = \frac{\mathbf{4,457 + 3,631 + 3,472 + 4,313 + 3,777}}{\mathbf{5}}$$

$$\mathbf{Ft_{-7}} = \mathbf{3,390}$$

Predictions for 2017:

$$\mathbf{Ft_{-8}} = \frac{\mathbf{3,631 + 3,472 + 4,313 + 3,777 + 4,011}}{\mathbf{5}}$$

$$\mathbf{Ft_{-8}} = \mathbf{3,840.80}$$

Predictions for 2018:

$$\mathbf{Ft_{-9}} = \frac{\mathbf{3,472 + 4,313 + 3,777 + 4,011 + 5,061}}{\mathbf{5}}$$

$$\mathbf{Ft_{-9}} = \mathbf{4,126.20}$$

Predictions for 2019:

$$\mathbf{Ft_{-10}} = \frac{\mathbf{4,313 + 3,777 + 4,011 + 5,061 + 5,944}}{\mathbf{5}}$$

$$\mathbf{Ft_{-10}} = \mathbf{4,621.20}$$

Predictions for 2020:

$$\mathbf{Ft_{-11}} = \frac{\mathbf{3,777 + 4,011 + 5,061 + 5,944 + 4,559}}{\mathbf{5}}$$

$$\mathbf{Ft_{-11}} = \mathbf{4,670.40}$$

Predictions for 2021:

$$\mathbf{Ft_{-12}} = \frac{\mathbf{4,011 + 5,061 + 5,944 + 4,559 + 5,350}}{\mathbf{5}}$$

$$\mathbf{Ft_{-12}} = \mathbf{4,985}$$

From the calculations above, we get a prediction of the number of accidents from 2014 to 2021. They will be represented in the table below.

Table 2. Presentation of Prediction Results Data

Year	Accident Number (Actual)	Accident Number (Predicted)

ISSN Print : 1979-7141
ISSN Online : 2541-1942

2010	4,613.00	
2011	4,508.00	
2012	4,457.00	
2013	3,631.00	
2014	3,472.00	4,136.20
2015	4,313.00	4,076.20
2016	3,777.00	3,930.00
2017	4,011.00	3,840.80
2018	5,061.00	4,126.80
2019	5,944.00	4,621.20
2020	4,559.00	4,670.40
2021	5,350.00	4,985.00

From the accident number data (actual) contained in table 1 and the accident number data (actual) contained in table 2, it will be visualized using line and bar graphs as below.

Figure 3. Line graphic visualization of accident number prediction data

ISSN Print : 1979-7141
ISSN Online : 2541-1942

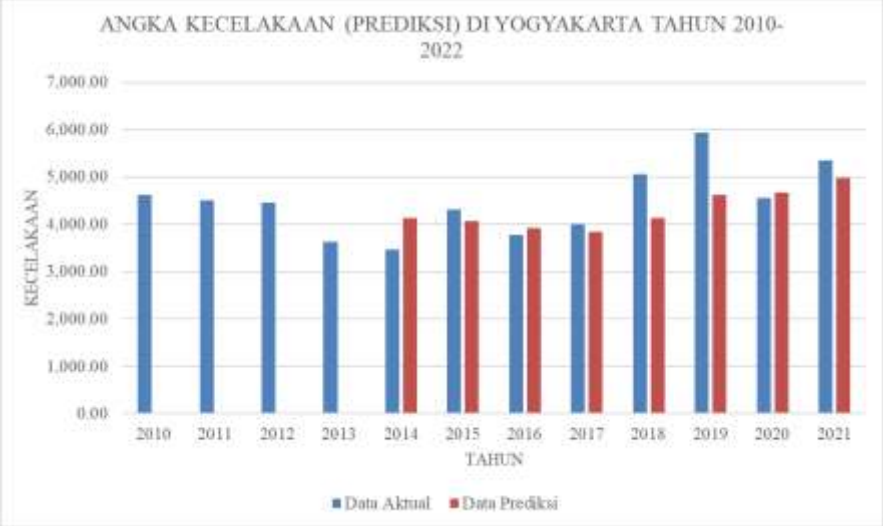

Figure 4. Bar graph visualization of accident number prediction data

3.3. Measurement of Prediction Model Error Values.

The error rate of the prediction model for the Single Moving Average prediction technique is sought using RMSE, MAD, and MAPE. First, look for the error value (Actual Value-Predicted Value) to find RMSE, and the absolute error value (|Actual Value-Predicted Value|) to find MAD and MAPE. The results can be seen below.

Table 3. Error Value Results

Year	Actual (At)	Predicted (Ft)	At-Ft
2010	4,613.00		
2011	4,508.00		
2012	4,457.00		
2013	3,631.00		
2014	3,472.00	4,136.20	-664.20
2015	4,313.00	4,076.20	236.80
2016	3,777.00	3,930.00	-153.00
2017	4,011.00	3,840.80	170.20
2018	5,061.00	4,126.80	934.20
2019	5,944.00	4,621.20	1,322.80
2020	4,559.00	4,670.40	-111.40
2021	5,350.00	4,985.00	365.00
∑			2,100.40

The error value is obtained, so the RMSE can be calculated. The RMSE value can be seen below.

Table 4. RMSE Value Results

Year	Actual (At)	Predicted (Ft)	At-Ft	$\frac{(At-Ft)^2}{n}$

ISSN Print : 1979-7141
ISSN Online : 2541-1942

2010	4,613.00			
2011	4,508.00			
2012	4,457.00			
2013	3,631.00			
2014	3,472.00	4,136.20	-664.20	55,145.21
2015	4,313.00	4,076.20	236.80	7,009.28
2016	3,777.00	3,930.00	-153.00	2,926.13
2017	4,011.00	3,840.80	170.20	3,621.00
2018	5,061.00	4,126.80	934.20	109,091.21
2019	5,944.00	4,621.20	1,322.80	218,724.98
2020	4,559.00	4,670.40	-111.40	1,551.24
2021	5,350.00	4,985.00	365.00	16,653.13
∑			2,100.40	414,722.17
RMSE				**643.99**

So we get an RMSE value of 643.99. Next, the absolute error value will be calculated to find MAD and MAPE. The absolute error value can be seen below.

Table 5. Absolute Error Value

Year	**Actual (At)**	**Predicted (Ft)**	**|At-Ft|**
2010	4,613.00		
2011	4,508.00		
2012	4,457.00		
2013	3,631.00		
2014	3,472.00	4,136.20	664.2
2015	4,313.00	4,076.20	236.8
2016	3,777.00	3,930.00	153
2017	4,011.00	3,840.80	170.2
2018	5,061.00	4,126.80	934.2
2019	5,944.00	4,621.20	1322.8
2020	4,559.00	4,670.40	111.4
2021	5,350.00	4,985.00	365
∑			3957.6

And we get the absolute error value, so we can calculate the MAD and MAPE. MAD and MAPE values can be seen below.

Table 6. MAD Value Results

Year	**Actual (At)**	**Predicted (Ft)**	**|At-Ft|**
2010	4,613.00		
2011	4,508.00		
2012	4,457.00		

ISSN Print : 1979-7141
ISSN Online : 2541-1942

2013	3,631.00		
2014	3,472.00	4,136.20	664.2
2015	4,313.00	4,076.20	236.8
2016	3,777.00	3,930.00	153
2017	4,011.00	3,840.80	170.2
2018	5,061.00	4,126.80	934.2
2019	5,944.00	4,621.20	1322.8
2020	4,559.00	4,670.40	111.4
2021	5,350.00	4,985.00	365
∑			3957.6
MAD			**494.7**

Table 7. Results of MAPE SCORE

Year	**Actual (At)**	**Predicted (Ft)**	**|At-Ft|**	$\left\|\frac{At - Ft}{At}\right\|$
2010	4,613.00			
2011	4,508.00			
2012	4,457.00			
2013	3,631.00			
2014	3,472.00	4,136.20	664.2	0.19
2015	4,313.00	4,076.20	236.8	0.05
2016	3,777.00	3,930.00	153	0.04
2017	4,011.00	3,840.80	170.2	0.04
2018	5,061.00	4,126.80	934.2	0.18
2019	5,944.00	4,621.20	1322.8	0.22
2020	4,559.00	4,670.40	111.4	0.02
2021	5,350.00	4,985.00	365	0.07
∑			3957.6	0.83
MAPE				**1036.17%**

Finally, we got a MAD value of 494.7 and a MAPE value of 1036.17%. And from the three measurements of error in the prediction value, a comparison can be made as follows.

Table 8. Comparison of RMSE, MAD and MAPE values

Year	**Prediction (MA5)**
MAD	**494.7**
RMSE	**643.99**
MAPE	**1036.17%**

4. CONCLUSION.

According to the problem model and the results of the predicted values in the articles that have been carried out, it can be concluded:

ISSN Print : 1979-7141
ISSN Online : 2541-1942

1. Predictions using the Single Moving Average technique have a fairly high error rate.
2. Based on the prediction results for the last three years (2019, 2020, 2021), the number of accidents has not increased drastically. The increase was only around 49 accidents in 2019-2020, but there was a drastic increase of around 315 accidents in 2020-2021.
3. In contrast to the predicted results, the actual (real) results of accidents in the last year (2021) had several accidents of 5,350, while the predicted data had several accidents of 4,985. So there is a difference of 365 accidents.
4. Testing the error level in the prediction method has an RMSE value of 643.99, MAD of 494.7, and MAPE of 1036.17%. This is quite a large amount of measurement error for a prediction model.
5. Looking at the 3 points above (points 2, 3, and 4), it is necessary to review the factors used in this article, whether due to human error during the data entry process or the inaccurate prediction method used.

5. DECLARATION OF COMPETING INTEREST.

We declare that we have no conflict of interest.

6. REFERENCES.

Faishol, Y. M., Purnamasari, I., & Goejantoro, R. (2017). Peramalan Regarima Pada Data Time Series (Studi Kasus: Penjualan Tiket Pesawat PT. Kumala Wisata Tenggarong). *Jurnal EKSPONENSIAL*, *8*(1).

Nyoman, N., Pinata, P., Sukarsa, M., Kadek, N., & Rusjayanthi, D. (2020). Prediksi Kecelakaan Lalu Lintas di Bali dengan XGBoost pada Python. *JURNAL ILMIAH MERPATI*, *8*(3).

Prana, S., Isesa, W., Prayogi, A. S., & Fahrudin, T. M. (2018). PEMODELAN DAN EVALUASI TREND FORECASTING PADA KONDISI KORBAN KECELAKAAN LALU LINTAS M ENGGUNAKAN TREND M OMENT DAN LEAST SQUARE. *Jurnal Sistem Cerdas*, *01*(02), 56–66.

Rachman, R. (2018). Penerapan Metode Moving Average dan Exponential Smoothing pada Peramalan Produksi Industri Garment. *JURNAL INFORMATIKA*, *5*(1), 211–220.

Rakhmat, L. A., Kusumawati, A., Bona Frazila, R., & Hendarto, S. (2012). Pengembangan Model Prediksi Kecelakaan Lalu Lintas pada Jalan Tol Purbaleunyi. *Jurnal Teknik Sipil*, *19*(3).

Saputro, A., & Purwanggono, B. (n.d.). *PERAMALAN PERENCANAAN PRODUKSI SEMEN DENGAN METODE EXPONENTIAL SMOOTHING PADA PT. SEMEN INDONESIA*.

Suprayogi, I., Trimaijon, & Mahyudin. (n.d.). *MODEL PREDIKSI LIKU KALIBRASI MENGGUNAKAN PENDEKATAN JARINGAN SARAF TIRUAN (JST) (Studi Kasus: Sub DAS Siak Hulu)*. http://ce.unri.ac.id

ISSN Print : 1979-7141
ISSN Online : 2541-1942