JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 8, No. 4, October 2024, pp. 1133-1145

Forecasting Beef Production with Comparison of Linear Regression and DMA Methods Based on n-th Ordo 3

Tundo[1], Mesra Betty Yel[1], Agung Yulianto Nugroho[2]
[1]Department of Computer Science, STIKOM CKI, Indonesia
[2]Department of Informatics, Universitas Cendekia Mitra Indonesia, Indonesia
asna8mujahid@gmail.com

Article History:
Received : 21-06-2024
Revised : 14-08-2024
Accepted : 20-09-2024
Online : 01-10-2024

Keywords:
Forecasting;
Beef;
Production;
Linear regression;
Double Moving Average.

ABSTRACT

Beef is considered a high-value commodity because it is an important food source of protein. Interest in beef is increasing along with increasing people's incomes and awareness of the importance of fulfilling nutrition. Demand for beef is expected to continue to increase. According to the Central Statistics Agency (CSA), beef production in Jakarta shows an increasing trend every year. In the last 10 years, beef production has increased significantly, but in 2020 there was a decrease in production of 7,240.68 tons due to the lockdown due to the corona virus outbreak. After that, in 2021, production reached 16,381.81 tons and will continue to increase in 2022 and 2023. Based on the above phenomenon, the aim of this research is to support the success and sustainability of the beef industry by ensuring that supply matches demand, resources are used optimally, and risks can be managed well. To predict beef production, an accurate method, model or approach is needed. One way to predict beef production in Jakarta is to use the Linear Regression and Double Moving Average (DMA) methodsThe way the Linear Regression and DMA methods work is to forecast based on concepts and properties. The concepts and properties of Linear Regression are models, functions, estimates and forecasting results, while DMA performs time series analysis based on moving averages. After analysis using MAPE, it was found that the algorithm that had the smallest error value was the linear regression algorithm with a percentage for the monthly period of 15% while for the year period it was 17% compared to DMA. So in this case it would be very appropriate to use the Linear Regression method from the error values obtained.

https://doi.org/10.31764/jtam.v8i4.24706

A. INTRODUCTION

One economic sector that plays an important role in people's lives is livestock. Sources of animal protein that are important for health are livestock products. The increasing standard of living of today's society increases awareness of the importance of meeting nutritional needs, including meeting animal protein needs (Astiti et al., 2023). This increases demand for livestock products, one of which is beef. Beef is considered a high-value commodity because it is an example of a food source of protein. Beef is increasingly in demand in line with increasing people's incomes and awareness of the importance of fulfilling nutrition. Demand for beef could be one commodity that will continue to increase (Abdallah et al., 2024).

DKI Jakarta has a total area of 662.33 km^2 consisting of five administrative city areas and one administrative district. The Central Jakarta Administrative City area covers 48.13 km^2, the North Jakarta area covers 146.66 km^2, the West Jakarta area covers 129.54 km^2, the South Jakarta area covers 147.37 km^2, the East Jakarta area covers 188.03 km^2, and the Regency area

Seribu Islands Administration covers an area of 8.70 km^2. Based on the area described above and the large number of residents in Jakarta, it is very natural that beef production in Jakarta is quite large. According to the Central Statistics Agency (CSA), beef production in Jakarta tends to have an increasing trend every year. In the last 10 years the amount of production has increased quite significantly, but in 2020 the amount decreased to 7,240.68 tons, this was due to the lockdown due to the corona virus outbreak. After that, in 2021 total production was 16,381.81 tons and continued to increase in 2022 and 2023 (Agustiar et al., 2021).

Data mining is the analysis of data taken from very large data stores that are widely used as useful information (Tundo & Mahardika, 2023). The term data mining is also called data discovery. One of the techniques created in data mining is a way of tracking existing data to build models (Bhuyan, Kol, A. Adediran, Oluwaseyi Jessy, & Tundo, 2023). One technique in data mining is following existing data to create a model (Sutopo et al., 2023).

To monitor and control beef stocks, information is needed that can be used to predict the movement of beef production in Jakarta. In predicting beef production, a method, model or approach needs to be tested for accuracy. There are many ways to predict beef production in Jakarta, one of which is using the linear regression and Double Moving Average (DMA) methods. The linear regression method is used to calculate the correlation between two or more variables used for prediction via a straight line (Maulud & Abdulazeez, 2020). Variable values are called variables themselves (Sitompul et al., 2023). Linear regression is a form of relationship where the independent variable, namely X, and the dependent variable, namely Y, are factors with a rank of one (Theofani & Sediyono, 2022). The linear regression that is often used is simple linear regression (Wahyudi & Arroufu, 2022). Apart from the linear regression method described above as a comparison, the author also uses the DMA method. DMA is a double moving average where one group average is calculated and then the second group calculates the moving average that was produced by the first group (Khairina et al., 2021). In essence, DMA is a continuation of SMA, where to calculate DMA you need to look for forecasting results from SMA first. A similar thing was stated by Mustapa et al. (2019), that the DMA method is a double-moving average calculation technique that forwards the results of the SMA method for re-forecasting with the aim that the forecasting results will be more optimal (Tobing, 2022). Apart from using linear regression and DMA methods.

As reinforcement in this research, the author presents references to previous research such as research conducted by Amaliyatul Hasanah entitled "Prediction of Rice Production in Sumenep Regency Using the Single Exponential Smoothing Method". The data analyzed is rice production data in Sumenep Regency from 2005 to 2022. Forecasting results show that rice production in Sumenep Regency in 2023 is estimated to reach 203,132.55 tonnes, with an accuracy level measured using MAPE of 11.243% and a value = 0.3 (Hasanah, 2023). Apart from that, there was also research conducted by M. Rajab Mudatsir, Sudirman Melangi, and Serwin entitled "Prediction of Salted Fish Production Using the Simple Linear Regression Method". Based on research, it is estimated that the amount of salted fish production in May 2022 will reach the value y = 168.53 + 24,175. The calculation accuracy was measured using MAPE and obtained a result of 0.8%. System testing using the white box method shows the same V(G) and CC values, namely 2 (Mudatsir et al., 2022).

Then the final reference is research conducted by Harun Indra Kusuma and Ridwan Saputra entitled "Analysis of Forecasting Demand for Inalcafa Jackets for Men's Products Using the Double Moving Average Method". After carrying out forecasting calculations using the double moving average method, the researcher determined the upper class limit, middle value and lower class limit based on sales data for the previous 24 months. Then, researchers calculated the moving range to compare the actual data values with the forecast values. From the calculation of the double moving average method, an error value of 6.77% was obtained. Estimated production for the next 6 months will average 34.67 units for the variety of men's jackets produced (Kusuma & Saputra, 2024). Of the three studies, none has discussed handling beef production using a comparison of linear regression and SES methods. The reason why we use these two algorithms is to compare which algorithm produces the lowest error rate. Based on the background described above, the identification of the problem is that there is no information that can show predictions of beef production in Jakarta and there is no solution to predict beef production in Jakarta. The aim of this research is to apply data mining using linear regression and DMA algorithms to predict beef production in Jakarta. Apart from that, the aim of this research is also to help CSA in estimating future beef production and assist them in creating a prediction system.

B. METHODS

1. Research Data

In this research, the type of data used is quantitative data, because the data used is in the form of numbers, namely beef production data in Jakarta. This research data was obtained using a secondary approach obtained from the Central Statistics Agency (CSA) website, namely in the form of data on the amount of beef production in Jakarta in 2000-2023. The data used in this research obtained from the BPS website is public. Data is presented in annual and monthly form to compare the results of the linear regression and DMA methods. For the production data obtained, it is more appropriate to use year or month period data.

2. Application of Methodology

This research uses two methods, namely linear regression and DMA on the beef production dataset in Jakarta. This research aims to predict the amount of beef production in the future and can help related agencies in estimating future beef production and assist them in creating a prediction system. This application will be created and arranged in a methodology application flow which is shown in Figure 1.

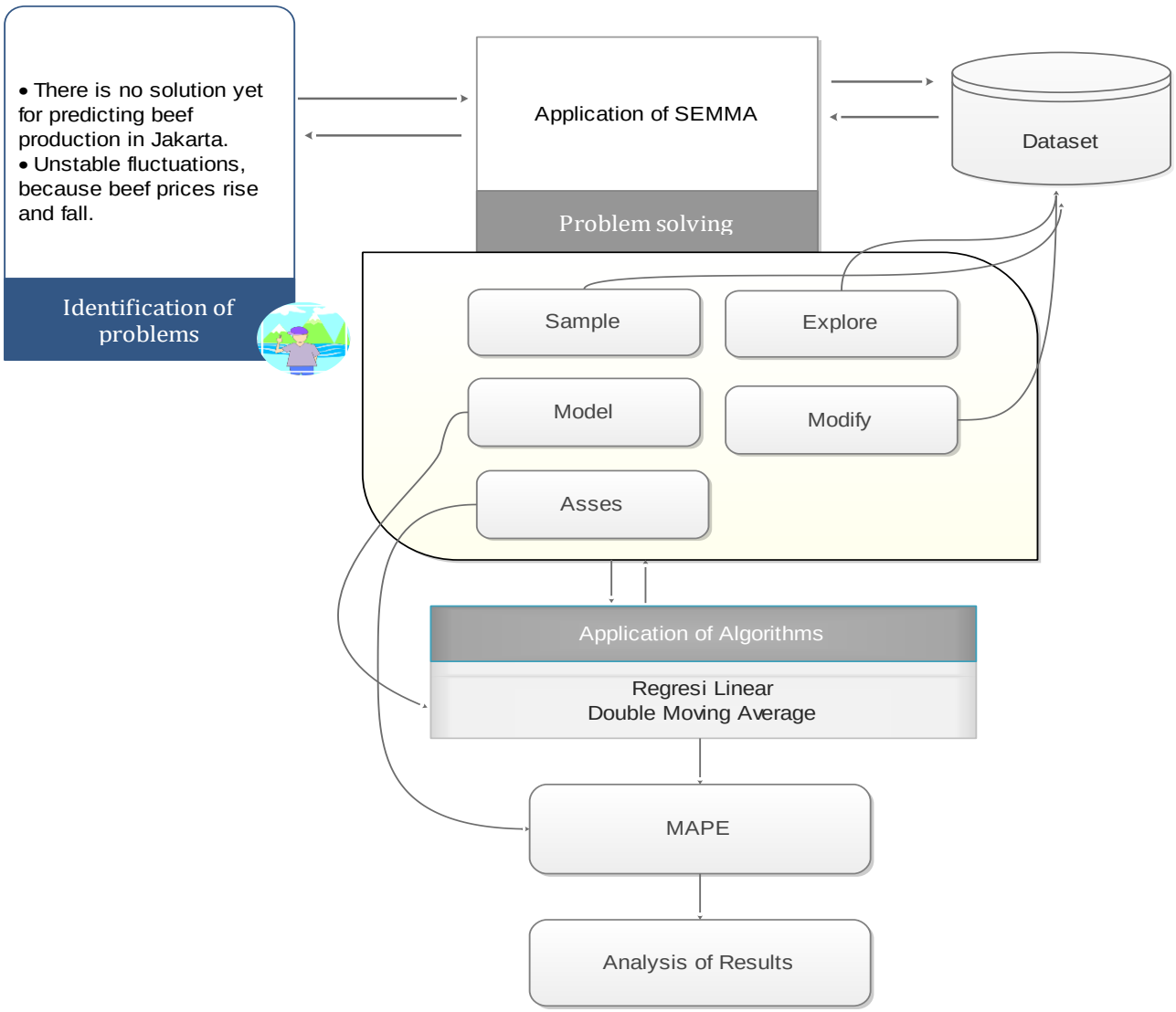


Figure 1. Application of Methodology

3. Test Design

In this research stage, the methodology used is SEMMA. SEMMA is a method used in data mining and data analysis to assist in the process of creating analytical models. SEMMA stands for Sample, Explore, Modify, Model, and Assess. This SEMMA method has five steps in system development (Zawad, 2023). The following are the steps in the SEMMA methodology applied in this research:

a. Sample stage

This stage is data collection to contain the information needed in the research (Cazacu & Titan, 2020). The data collection process was carried out on the Central Statistics Agency (BPS) website regarding beef production in Jakarta in the period 2000 to 2023. This dataset has a special purpose, namely to support predictions of beef production in Jakarta. Apart from searching for datasets, at this stage a literature review is also carried out to look for references to previous research.

b. Explore stage

Explore is a data mining process that can be used to search for data sets and provide information related to unexpected anomalous trends that can be used (Andrade-Arenas

et al., 2024). In the explore stage, use data understanding to examine the data, so you can identify problems in the data. Next, data selection is carried out to select the data that will be used for the mining process while still representing the original data. In the selection data the author determines two attributes, namely the year attribute and the quantity attribute in tons.

c. Modify stage

A data mining process used to modify data and change variables to focus the model selection process (Wiemer, Drowatzky, & Ihlenfeldt, 2019). The modify stage carries out a data cleaning process, missing values can be removed, inconsistent data can be checked, and data can be corrected. At this stage, because the available dataset is only in annual form, the author also presents the data in monthly form. Apart from that, the dataset used in linear regression in the form of univariate data is changed to multivariate, by means of the X value in the next period becoming the Y value in the first period.

d. Model stage

At this stage, a model was created using the linear regression and DMA methods. This stage is also a data mining process that can be used to model data and look for prediction results using the two methods mentioned. After carrying out this stage, it will produce predicted values from each method. The formulation of linear regression is as follows (Syah et al., 2019):

$$y = \alpha + bX \quad (1)$$

$$b = \frac{n(\sum xy)-(\sum x)(\sum y)}{n(\sum x^2)-(\sum x)2} \quad (2)$$

$$a = \frac{\sum y-b*\sum x}{n} \quad (3)$$

Information: Y is Independent Variable (Dependent); X is Dependent Variable (independent); a is constant; and b is regression coefficient (similarity). The steps taken in formulating DMA are as follows (Pranoto, 2022)(Alex & Nur Rahmawati, 2023):

1) Determine the first smoothing value (S'_t)

$$S' = \frac{X_t+X_{t-1}+X_{t-2}+\cdots+X_{t-n+1}}{n} \quad (4)$$

2) Determine the second, third, ...n-th values for smoothing as (S''_t)

$$S'' = \frac{S'_t+S'_{t-1}+S'_{t-2}+\cdots+S'_{t-n+1}}{n} \quad (5)$$

3) Determine the constant value results (a)

$$a_t = 2S'_t - S''_t \quad (6)$$

4) Determine the resulting slope value as the steepness point (b_t)

$$b_t = \frac{2}{n-1}(S'_t - S''_t) \tag{7}$$

Description: S' is the first moving average value in period t; S'' is the value of the second moving average in period t; x_t is the value of the second moving average in period t; and n is the number of limits in the moving average (n-th ordo). Determining future periods will happen when the at and bt values are considered for alpha calculations (Yel et al., 2024). The moving normal is n arranged, and s is the length of the regular cycle. This thinks about assessed distinctive k values to decide their effect on determining exactness. Values of s speaking every day or week by week regularly were chosen to play down the cruel squared blunder on the preparing set (Neisyafitri & Ongkunaruk, 2022). Determine the results of the forecasting value (F_{t+m}) as follows:

$$F_{t+m} = a_t + b_t m \tag{8}$$

Description: F_{t+m} is forecasting results in the next period; X_t is the actual data value in period t; m is the number of subsequent periods to be predicted; a_t is the number of period constant values t; and b_t is the trend value results in the appropriate data.

e. **Assess stage**

At this stage, an evaluation of the modeling that has been created is carried out (Omari Firas, 2023). The evaluation in question is by looking for the Mean Absolute Percentage Error (MAPE) results (Saifullah, Suryotomo, Drężewski, Tanone, & Tundo, 2023). Looking for MAPE results is done to find out how accurate the prediction results that have been made by the three methods are (Saifullah et al., 2024). The smaller the MAPE value, the prediction results can be said to be accurate. The MAPE value can be calculated using the following equation formula:

$$\text{MAPE} = \frac{\sum(\frac{|Actual - forecast|}{Actual})}{n} x\ 100 \tag{9}$$

The MAPE result categories consist of "Very Good", "Good", "Enough", and "Poor" which appear in Table 1.

Table 1. MAPE value level categories

MAPE value	Categories
< 10%	Very Good
10% - 20%	Good
20% - 50%	Enough
> 50%	Poor

C. RESULT AND DISCUSSION

1. Production Prediction for Year Period

Annual data on beef production in Jakarta is presented in Table 2.

Table 2. Beef Production Data (Annual)

Year	Production (Tons)
2000	14282
2001	14888
2002	13719
2003	16610
2004	13045
2005	10060,83
2006	8505,41
2007	7051
2008	8562
2000	14282
...	...
...	...
2023	17664.8

The prediction results for beef production using a linear regression algorithm based on the values α = 5867.6493 and b = 0.530279 are presented in Table 3 below.

Table 3. Prediction using Linear Regression

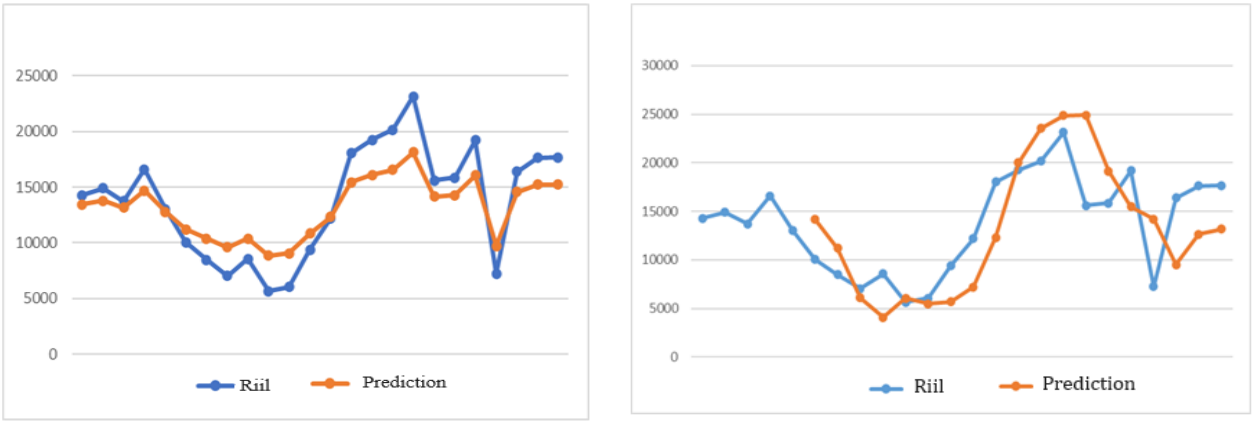
Year	(X)	(Y)	Prediction	MAPE
2000	14282	14888	13441.09	0.06
2001	14888	13719	13762.44	0.08
2002	13719	16610	13142.55	0.04
2003	16610	13045	14675.58	0.12
2004	13045	10060.83	12785.14	0.02
2005	10060.83	8505.41	11202.70	0.11
2006	8505.41	7051	10377.89	0.22
2007	7051	8562	9606.65	0.36
2008	8562	5657	10407.90	0.22
2009	5657	6058	8867.44	0.57
...	...	...	...	...
2023	17664.8	0	15234.92	0.14
2024	15234.92		13946.41	0.08
		MAPE		17%

Meanwhile, the prediction results for beef production using the DMA algorithm based on order 3 values are presented in Table 4 below.

Table 1. Prediction using DMA

Year	Production (Tons)	SMA (S')	DMA(S'')	Prediction (F_{t+m})	MAPE
2000	14282	-	-	-	-
2001	14888	-	-	-	-
2002	13719	14296.33	-	-	-
2003	16610	15072.33	-	-	-
2004	13045	14458	14608.89	-	-
2005	10060.83	13238.61	14256.31	14156.22	41%
2006	8505.41	10537.08	12744.56	11203.20	32%
2007	7051	8539.08	10771.59	6122.11	13%
2008	8562	8039.47	9038.543	4074.06	52%
2009	5657	7090	7889.517	6041.32	7%
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...
2023	17664.8	17221.41	15080.15	13160.22	26%
2024				21503.92	
			MAPE		31%

The prediction results for January 2024 using the linear regression algorithm are 13946.41 tons with a MAPE value of 17%. Meanwhile, using the DMA algorithm, it was 21503.92 tons with a MAPE value of 31%. In terms of trends, the graph of linear regression production predictions and DMA with actual production is shown in Figure 2.



a. Linear Regression b. DMA

Figure 2. Trend graph of production predictions with actual production (Annual)

2. Monthly Period Production Prediction

Apart from annual data, monthly data is also presented in Table 5.

Table 5. Beef Production Data (Monthly)

Year	Month	Production (Tons)
2000	January	1245
	February	987
	March	445
	...	...
2001	January	1238
	February	1367
	March	1148
	...	...
2002	January	1211
	February	1146
	March	1134
	...	...
...	...	...
	...	...
	...	...
2024	January	1763
	February	1518
	March	1644
	...	...

The prediction results of beef production for the month period using a linear regression algorithm based on the values α = 385.87894 and b = 0.660526, are presented in Table 6 below.

Table 6. Prediction using Linear Regression

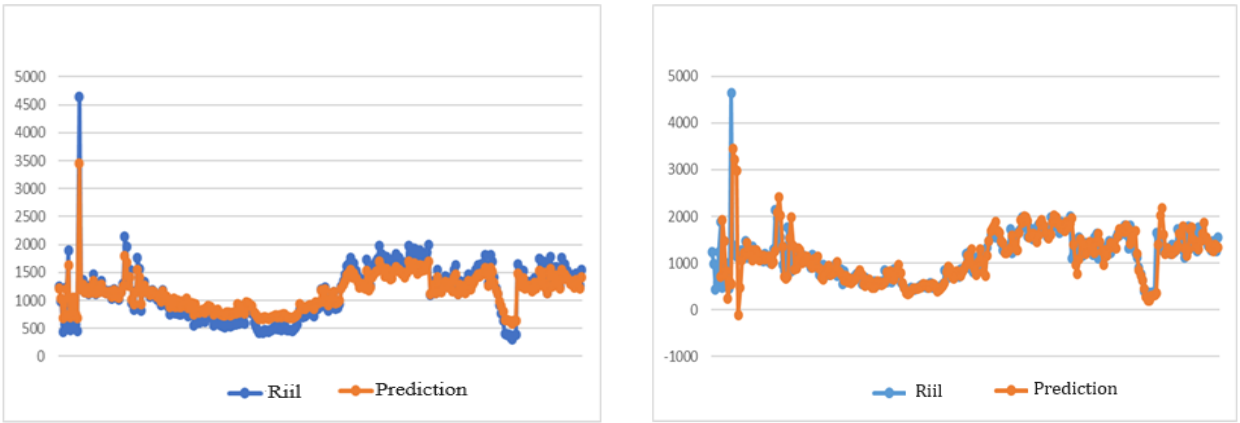
Year	Month	(X)	(Y)	Prediction	MAPE
2000	January	1245	987	1208.23	3%
	February	987	445	1037.81	5%
	March	445	567	679.81	53%
	...	...	...	...	...
2001	January	1238	1367	1203.610128	3%
	February	1367	1148	1288.817982	6%
	March	1148	1234	1144.162788	0%
	...	...	...	...	...
2002	January	1211	1146	1185.775926	2%
	February	1146	1134	1142.841736	0%
	March	1134	1166	1134.915424	0%
	...	...	...	...	...
...	...	...	...	...	...
	...	...	...	...	...
	...	...	...	...	...
2023	October	1356	1259	1281.552196	5%
	November	1259	1545	1217.481174	3%
	December	1545	1634	1406.39161	9%
2024	January	1634		1465.178424	10%
			MAPE		15%

Based on table 6, it can be seen that the resulting MAPE is 15%. The results of monthly beef production predictions produce production predictions for January 2024 for beef production in Jakarta. Meanwhile, the prediction results for beef production using the DMA algorithm based on order 3 values are presented in Table 7 below.

Table 2. Prediction using DMA

Year	Month	Production (Tons)	SMA (S')	DMA(S'')	Prediction (F_{t+m})	MAPE
2000	Jan	1245	-	-	-	-
	Feb	987	-	-	-	-
	Mar	445	892,33	-	-	-
	Apr	567	666,33	-	-	-
	Mei	1234	748,66	769,11	-	-
	Jun	1890	1230,33	881,77	707,77	63%
	Jul	467	1197	1058,66	1927,44	313%
	...	...	...	...	...	...
2001	Jan	1238	2114,33	1559,77	3441,55	178%
	Feb	1367	2417,66	2140,88	3223,44	136%
	Mar	1148	1251	1927,66	2971,22	159%
	...	...	...	...	...	...
2023	...	...	...	...	...	...
	Okt	1356	1369	1376	1393,77	3%
	Nov	1259	1365,66	1376,88	1355	8%
	Dec	1545	1386,66	1373,77	1343,22	13%
2024	Jan				1412,444	14%
			MAPE			19%

The prediction results for January 2024 using a linear regression algorithm for the monthly period are 1465.1784 tons with a resulting MAPE value of 15%. Meanwhile, using the DMA algorithm, it was 1412.444 tons with a MAPE value of 19%. In terms of trends, the graph of linear regression production predictions and DMA with actual production is shown in Figure 3.

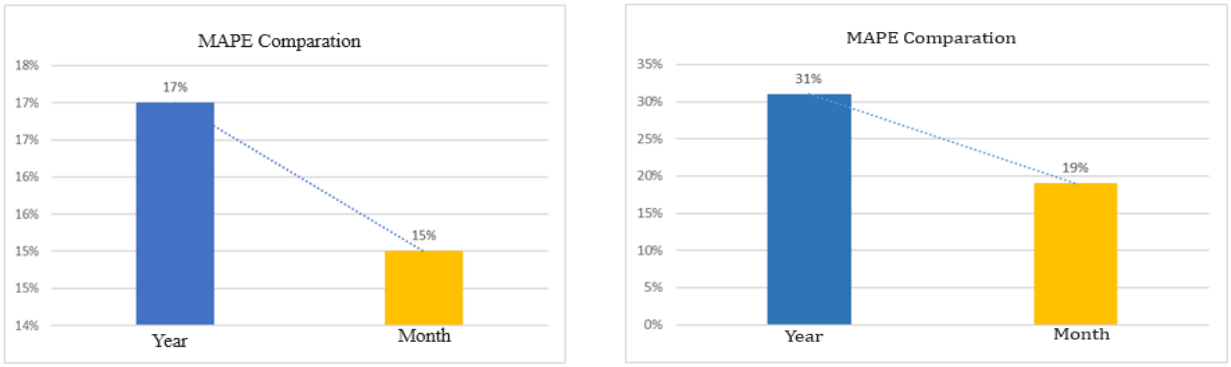


a. Linear Regression b. DMA

Figure 3. Trend graph of production predictions with actual production (Monthly)

3. Comparative Analysis of Algorithms

In finding the most suitable algorithm in the case of beef production in Jakarta based on the three algorithms used, a bar chart was created from the MAPE results to simplify the analysis. The initial concept, combining prediction results over monthly and yearly periods for each algorithm. Based on Figure 4 it can be seen that the MAPE value for the monthly period is smaller than the annual period based on the three algorithms.



a. Linear Regression (Year and Month) b. DMA (Year and Month)

Figure 4. Comparison of MAPE Linear Regression and DMA for Year and Month periods

The next concept combines all the prediction results from both algorithms for both monthly and annual periods. Based on Figure 5, the smallest MAPE value obtained is the linear regression algorithm for either a month or year period, where the MAPE results are 15%, 17%. Thus, the linear regression algorithm is more suitable for solving beef production problems in Jakarta.

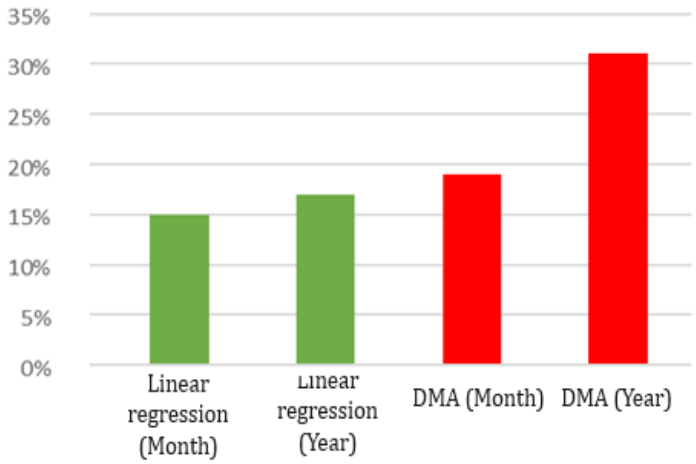


Figure 5. Comparison of linear regression MAPE and DMA (Year and Month)

D. CONCLUSION AND SUGGESTIONS

The conclusion that can be obtained from research on beef production predictions in Jakarta is that the Linear regression and DMA methods can be applied to predict beef production in Jakarta with results obtained in January 2024, using the linear regression algorithm it produces a prediction of 13946.41 in the year period and for the month period it

produces a prediction of 1465.1784, and the prediction from the results of the DMA algorithm produces 21503.92 in the year period and for the month period it produces predictions of 1412.444. After analysis using MAPE, it was found that the algorithm that had the smallest error value was the linear regression algorithm with a percentage for the monthly period of 15% while for the year period it was 17% compared to other algorithms. This forecast is quite accurate because the MAPE value for each algorithm used has an error of less than 31%. Suggestions for further developers are to use various variants of the α and b values in order to reduce the MAPE value to a smaller one or to compare the results of this research with the results of variants of the α and b values. One way to explore the α and b values is by comparing the minimum and maximum α values between the range (0.1 – 0.9) while the b value will follow the magnitude of the α value, in accordance with the concept of the equation method.

REFERENCES

Abdallah, B. N., Wardani, F. S., Prabandewa, C. D., & Hertadi. (2024). Determination of Supply Strategy for Beef Price Stability in Balikpapan: Game Theory Approach. *G-Tech : Jurnal Teknologi Terapan*, *8*(1), 26–35. Retrieved from https://ejournal.uniramalang.ac.id/index.php/g-tech/article/view/1823/1229

Agustiar, R., Triatmojo, A., & Guntoro, B. (2021). The study of local beef market structure in Jakarta, Indonesia. In *IOP Conference Series: Earth and Environmental Science* (Vol. 888, pp. 1–5). https://doi.org/10.1088/1755-1315/888/1/012082

Alex, M. A. H., & Nur Rahmawati. (2023). Application of the Single Moving Average, Weighted Moving Average and Exponential Smoothing Methods For Forecasting Demand At Boy Delivery. *Tibuana*, *6*(1), 32–37. https://doi.org/10.36456/tibuana.6.1.6442.32-37

Andrade-Arenas, L., Rubio-Paucar, I., & Yactayo-Arias, C. (2024). Data mining for predictive analysis in gynecology: a focus on cervical health. *International Journal of Electrical and Computer Engineering*, *14*(3), 2822–2833. https://doi.org/10.11591/ijece.v14i3.pp2822-2833

Astiti, N. M. A. G. R., Wedaningsih, K. N., & Parwata, I. K. W. (2023). Potential demand and supply of beef cattle in Indonesia. *Eximia*, *11*, 24–32. https://doi.org/10.47577/eximia.v11i1.274

Bhuyan, H., Kol, M., A. Adediran, D., Oluwaseyi Jessy, B., & Tundo, T. (2023). Predicting Uterine Fibroids with Multiple Classifiers: An Analysis. *SciWaveBulletin*, *01*(02), 18–26. https://doi.org/10.61925/swb.2023.1203

Cazacu, M., & Titan, E. (2020). Adapting CRISP-DM for Social Sciences. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, *11*(2sup1), 99–106. https://doi.org/10.18662/brain/11.2sup1/97

Hasanah, A. (2023). Prediksi Produksi Padi di Kabupaten Sumenep Menggunakan Metode Single Exponential Smoothing. *Jurnal Arjuna: Publikasi Ilmu Pendidikan, Bahasa Dan Matematika*, *1*(4), 264–272. Retrieved from https://doi.org/10.61132/arjuna.v1i4.136

Khairina, D. M., Khairunnisa, R., Hatta, H. R., & Maharani, S. (2021). Comparison of the trend moment and double moving average methods for forecasting the number of dengue hemorrhagic fever patients. *Bulletin of Electrical Engineering and Informatics*, *10*(2), 978–987. https://doi.org/10.11591/eei.v10i2.2711

Kusuma, H. I., & Saputra, R. (2024). Analisis Peramalan Permintaan Jaket Inalcafa pada Produk Pria dengan Metode Double Moving Average. *G-Tech: Jurnal Teknologi Terapan*, *8*(2), 1213–1219. https://doi.org/10.33379/gtech.v8i2.4222

Maulud, D., & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, *1*(2), 140–147. https://doi.org/10.38094/jastt1457

Mudatsir, R. M., Melangi, S., & Serwin. (2022). Prediksi Jumlah Produksi Ikan Asin Menggunakan Metode Regresi Linear Sederhana. *JURNAL BALOK - Banthayo Lo Komputer*, *1*(2827–9425), 118–124.

Mustapa, R., Latief, M., & Rohandi, M. (2019). Double moving average method for predicting the number of patients with dengue fever in Gorontalo City. *Sciences and Technology (GCSST)*, *2*, 332–337.

Retrieved from https://series.gci.or.id

Neisyafitri, R. J., & Ongkunaruk, P. (2022). The Use of Intervention Approach in Individual and Aggregate Forecasting Methods for Burger Patties: A Case in Indonesia. *Agraris*, *8*(1), 20–33. https://doi.org/10.18196/agraris.v8i1.12842

Omari Firas. (2023). A combination of SEMMA & CRISP-DM models for effectively handling big data using formal concept analysis based knowledge discovery: A data mining approach. *World Journal of Advanced Engineering Technology and Sciences*, *8*(1), 009–014. https://doi.org/10.30574/wjaets.2023.8.1.0147

Pranoto, G. T. (2022). Forecasting With Weighted Moving Average Method for Product Procurement Stock. *Jurnal Sistem Informasi Dan Sains Teknologi*, *4*(2). https://doi.org/10.31326/sistek.v4i2.1268

Saifullah, S., Dreżewski, R., Dwiyanto, F. A., Aribowo, A. S., Fauziah, Y., & Cahyana, N. H. (2024). Automated Text Annotation Using a Semi-Supervised Approach with Meta Vectorizer and Machine Learning Algorithms for Hate Speech Detection. *Applied Sciences*, *14*(3), 1078. https://doi.org/10.3390/app14031078

Saifullah, S., Suryotomo, A. P., Dreżewski, R., Tanone, R., & Tundo. (2023). *Optimizing Brain Tumor Segmentation Through CNN U-Net with CLAHE-HE Image Enhancement. 1st International Conference on Advanced Informatics and Intelligent Information Systems (ICAI3S 2023)*. Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-366-5

Sitompul, M., Hasan, M. A., & Devega, M. (2023). Forecasting Simcard Demand Using Linear Regression Method. *IT Journal Research and Development*, *8*(1), 48–60. https://doi.org/10.25299/itjrd.2023.12202

Sutopo, J., Khanapi, M., Ghani, A., Burhanuddin, M. A., Septianti, A. N., & Tundo, T. (2023). Dance Gesture Recognition Using Laban Movement Analysis with J48 Classification. *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, *11*(2), 528–536. https://doi.org/10.52549/ijeei.v11i2.4314

Syah, L. Y., Nafsiah, S. N., & Saddhono, K. (2019). Linear regression statistic from accounting information system application for Employee integrity. *Journal of Physics: Conference Series*, *1339*(1). https://doi.org/10.1088/1742-6596/1339/1/012131

Theofani, G., & Sediyono, E. (2022). Multiple Linear Regression Analysis on Factors that Influence Employees Work Motivation. *SinkrOn*, *7*(3), 791–798. https://doi.org/10.33395/sinkron.v7i3.11453

Tobing, D. N. L. (2022). Indihome Product Sales Forecasting with the Double Moving Average and Double Exponential Smoothing Methods on PT. Telkom Witel Sumut Pematang Siantar. *Formosa Journal of Science and Technology*, *1*(8), 1201–1222. https://doi.org/10.55927/fjst.v1i8.2281

Tundo, T., & Mahardika, F. (2023). Fuzzy Inference System Tsukamoto – Decision Tree C 4 . 5 in Predicting the Amount of Roof Tile Production in Kebumen. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, *7*(2), 533–544.

Wahyudi, T., & Arroufu, D. S. (2022). Implementation of Data Mining Prediction Delivery Time Using Linear Regression Algorithm. *Journal of Applied Engineering and Technological Science*, *4*(1), 84–92. https://doi.org/10.37385/jaets.v4i1.918

Wiemer, H., Drowatzky, L., & Ihlenfeldt, S. (2019). Data mining methodology for engineering applications (DMME)-A holistic extension to the CRISP-DM model. *Applied Sciences (Switzerland)*, *9*(12). https://doi.org/10.3390/app9122407

Yel, M. B., Tundo, T., & Arinal, V. (2024). Forecasting Roof Tiles Production with Comparison of SMA and DMA Methods Based on n-th Ordo 2 and 4. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, *8*(3), 667–679.

Zawad, N. M. (2023). Application of Data Mining in Healthcare of Bangladesh. *IJISCS (International Journal of Information System and Computer Science)*, *7*(2), 89. https://doi.org/10.56327/ijiscs.v7i2.1433