BIO Web of Conferences **164**, 01003 (2025) https://doi.org/10.1051/bioconf/202516401003
ICAAA 2025

Forecasting Duck Egg Production Using ARIMA Model: A Case Study on Magelang Ducks

Zakia Ulfah[1,3], *Dyah* Maharani[1*], and *Heru* Sasongko[2]

[1]Laboratory of Animal Genetics and Breeding, Department of Breeding and Reproduction, Faculty of Animal Science, Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia
[2]Poultry Science Laboratory, Animal Production Department, Faculty of Animal Science, Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia
[3]Animal Breeding Division, Animal and Fisheries Department, Magelang Regency Government, Central Java, Indonesia

Abstract. This study investigated the application of the ARIMA model to forecast egg production trends in the Magelang ducks. Data were collected from 100 ducks over a 135-day period135 d of monitoring. The mean Daily Duck Production (DDP) was 43.53% with a standard deviation of 23.82%, indicating substantial variability in age-related traits and genetic factors. Following confirmation of data stationarity, the ARIMA (2,1,0) model was identified as the optimal fit. The model effectively captures historical trends and provides accurate forecasts of future production. The mathematic model was $Y_t = 0{,}896 + 0{,}322\ Y_{t-1} + 0{,}315\ Y_{t-2} - 0{,}363\ Y_{t-3} + \varepsilon_t$. These findings provide valuable insights into improving farm management, optimizing resource allocation, and enhancing egg production strategies. Thus, the ARIMA model can be used by researchers or farm managers to help farmers optimize resource management. This model helps farmers optimize resource management during peak and low production periods, making it highly useful even with limited data. Broadly, it enables better decision-making regarding feeding, breeding, and resource allocation for sustainable and profitable farm management.

1 Introduction

1.1 Background

Duck eggs occupy a significant and expanding position in poultry industry. In certain Asian regions, including China, duck eggs constitute a substantial share of total egg consumption, representing between 10% and 30% of the market[1]. Their nutritional benefits, culinary versatility, and increasing market demand underscore their potential as valuable products for producers and consumers. Forecasting egg production is crucial for enhancing poultry operations, facilitating prompt decision-making, and boosting efficiency through precise forecasts derived from past data and model assessments.[2]. This practice ultimately leads to improved profitability and long-term viability in the duck egg industry [3]. Furthermore, understanding future production trends assists farmers in aligning their marketing strategies

* Corresponding author: d.maharani@mail.ugm.ac.id

with consumer demand, ensuring that they can capitalize on peak seasons, even with limited data availability for the Magelang duck. Farmers generally face limitations in data availability owing to insufficient recording. Forecasting helps farmers manage risks by allowing them to anticipate production changes and to take proactive steps to address potential challenges. Effective forecasting enhances productivity and contributes to sustainability and profitability of poultry farming.

The autoregressive integrated moving average (ARIMA) method is a statistical tool for predicting egg production, particularly when dealing with time-series data that exhibit seasonal trends. By leveraging past observations, ARIMA models effectively identify patterns and autocorrelations in time-series data, thereby facilitating accurate prediction of future output levels[4]. The ARIMA model predicts future data points by understanding the autoregressive (p), degree of differencing (d), and moving average (q) of time-series data[5]. ARIMA forecasting of egg production utilizes a time-series analysis to predict future outputs based on historical data. The model enhances its forecasting accuracy by utilizing differencing techniques to handle non-stationary data, which is crucial for accurately predicting fluctuating egg production and similar applications[6]. Combining simplicity with predictive capabilities, ARIMA models offer an efficient approach to forecasting [7]. ARIMA is optimal for predicting Daily Duck Production (DDP) in the absence or minimal presence of seasonal patterns, offering simplicity, flexibility with non-stationary data, and robust short-term forecasts. Daily duck production refers to the number of eggs that a female duck lays per day. Daily duck production was calculated by expressing the average number of eggs laid per day relative to the number of female ducks in a flock. This is referred to as laying percentage or duck day production (DDP). Although SARIMA is suitable for pronounced seasonality, ARIMA's lower complexity renders it ideal for straightforward trend-based predictions.

1.2 Research Objectives

The objective of this study was to investigate the most suitable ARIMA model for forecasting Magelang ducks's egg production trends, considering the time-series characteristics of DDP data on Magelang duck. This study proposed that the ARIMA model is well suited for predicting egg production trends, given the time-series characteristics of Daily Duck Production (DDP) data.

2 Materials and Method

2.1 Study Design

Thus data collection was performed by completely randomized design. Seventy days aged female Magelang ducks were randomly selected based on hatching records to ensure accurate age documentation. The selection included 100 ducks that were in good health, without physical defects, and reared under the same conditions as individual cages. The number of eggs collected daily was recorded to as a critical metric to determine DDP for assessing the farm productivity.

The first day of recording commenced with discovery of the first egg in the cage. Observations were recorded daily for 135 days after the first egg was laid, number of eggs daily data were recorded by KTT Bina Mandiri, Magelang, and Central Java between April 27, 2024 (157 aged old), and September 8, 2024 (291 aged old).

2.2. Data Analysis

Duck Day Production (DDP) is a quantitative measure of egg-laying efficiency in duck flocks. It is calculated by comparing the number of eggs daily to the number of ducks observed, then multiplying by 100 to yield a percentage. This calculation serves as an effective metric for assessing the productivity of egg-laying ducks in a given population. The DDP data were then analyzed using SPSS 26 to highlight key aspects such as mean, median, standard deviation, frequency distribution, and coefficient variation, which were all examined through descriptive statistics. A DDP data trend analysis was performed to identify the patterns and determine the stationarity of the time series. Finally, an Autoregressive Integrated Moving Average (ARIMA) analysis was conducted. Additionally, using DDP as a time series, this study demonstrates the accuracy of the ARIMA model in forecasting future production, aiding in informed management decisions to optimize efficiency and resource allocation on duck farms.

3 Result and Discussion

3.1 Descriptive Statistic

Table 1. describes the descriptive statistics of the DDP data for Magelang ducks. A 78 percentage-point range indicates substantial variation in egg production percentages, presumably influenced by genetic variation among the ducks and age-related traits since the data were time series.

Table 1. Descriptive statistics of Magelang ducks Daily Duck Production (DDP)

No	Metric	Value
1	Number of duck (N)	100
2	Number of Days	135
3	DDP Range	78%
4	DDP Minimum	1%
5	DDP Maximum	79%
6	DDP Mean	43.53%
7	DDP Standard Deviation	23.82%
8	DDP Coefficient of Variance	54.72%

.

An average DDP of 43.53% reflected moderate productivity, with data collected when the ducks were 291 days old. Duck egg production depends on breed, age, environment, and health, with DDP ranging from 60.57% to 68.86% in various breeds over 12 weeks[8]. Research on Chinese Shan Ma ducks showed peak egg production at 180 days, surpassing 95%, and then declining to 40-50% by 520 days[9]. The high standard deviation of 23.82% suggests considerable daily variability in egg production, potentially owing to age-related trends and genetic variation. The standard deviation shows daily changes in egg production, which makes sense since ducks are aging. As they grew older, they produced more eggs, and genetic variation could also contribute to this variance, which represents a high coefficient of variance value of 54.72%.

3.2 Forecasting egg production by ARIMA model

ARIMA is a forecasting technique for time-series data that can handle trends and nonseasonality. The ARIMA model is consisted of (p, d, q) where, p is the autoregressive parameters, d is the number of differencing passes and q is the moving average parameters. Four critical steps must be considered when applying ARIMA: identification, estimation, diagnostic verification, and forecasting.

3.2.1. *Identification*

The process of achieving stationarity and constructing a suitable model involved an identification phase using autocorrelation functions (ACFs), partial autocorrelation functions (PACFs), and transformations such as differencing and logarithmic conversion [3]. The DDP data qualify as time-series data, making the ARIMA analysis applicable for forecasting. Fig 1 illustrates the Magelang Duck DDP trend over 135 days. The initial DDP blue graph is nonstationary, which necessitates data differencing [10]. After differencing, the red graph shows stationarity. Consequently, the d value in the ARIMA model is 1, as a Single-differencing was performed.

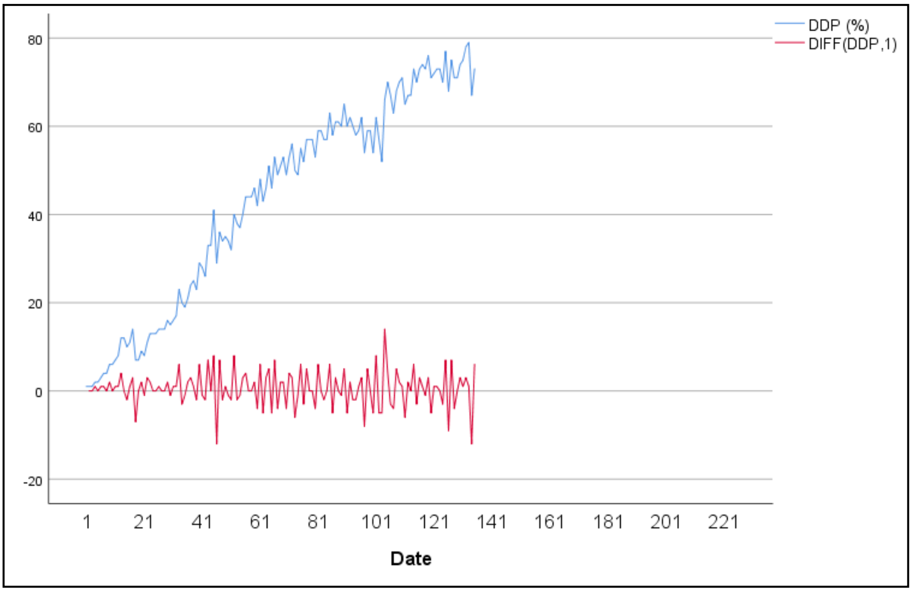

Fig. 1. The Magelang duck production and differencing trend over 135 days (DDP: duck daily production; diff = 1 ; N = 100)

Subsequently, we performed autocorrelation and partial autocorrelation analyses of the differenced DDP data.

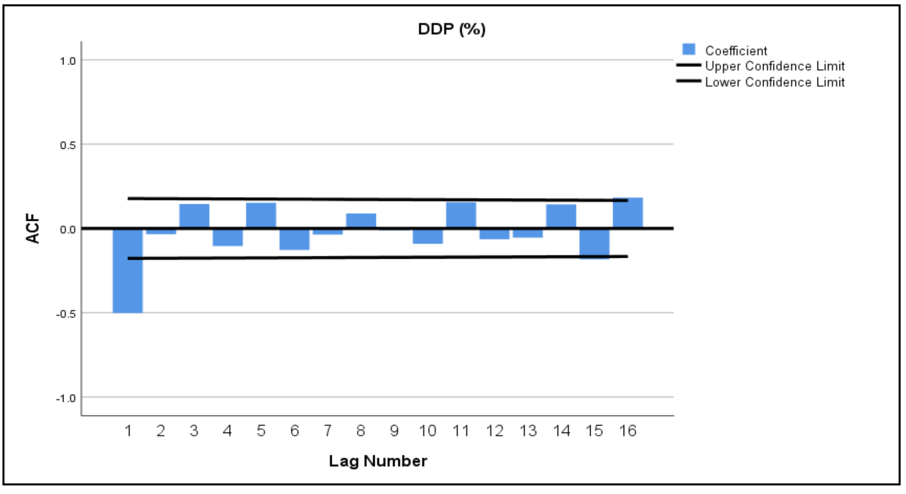

Fig. 2. Autocorrelation (ACF) chart of Magelang duck daily production (DDP)

The autocorrelation function (ACF) plot is shown in Fig. 2. Thus, potential q-value was determined. Because the lag exceeding the upper confidence limit is lag 1, the q value may be 0 or 1.

The p-value can be determined by analysing the partial autocorrelation function (PACF) chart. The lags exceeding the upper confidence limit are Lag 1 and Lag 2; consequently, the p-values applicable to the model are 0, 1, and 2. The autocorrelation function (ACF) and PACF charts indicate that the data are stationary, which is a prerequisite for implementing the ARIMA model. Partial autocorrelation is illustrated in Fig. 3.

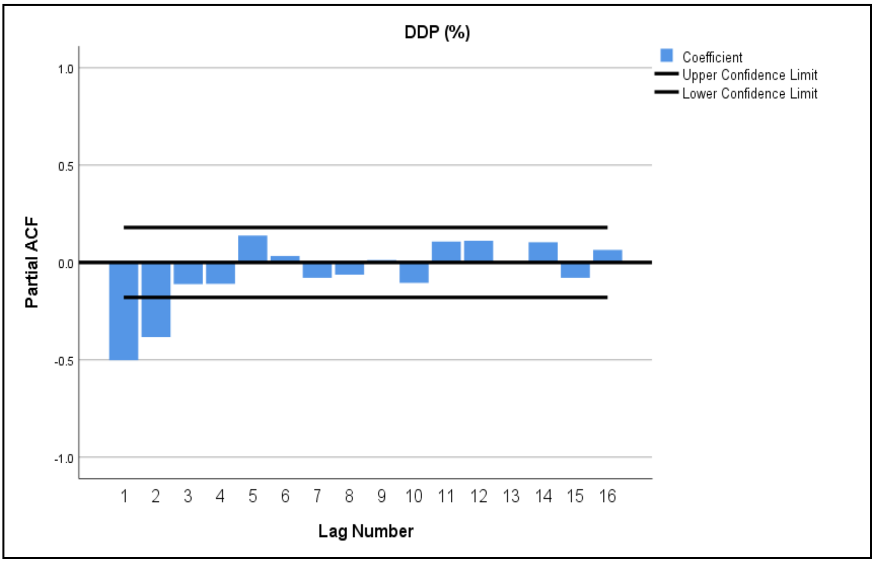

Fig. 3. Partial Autocorrelation (PAFC) chart of Magelang duck daily production (DDP)

Lags exceeding the upper confidence limit were lag 1 and lag 2. Consequently, we defined the potential p-values for the model as 0, 1, and 2. Based on our identification process, we can now determine the potential order of the ARIMA model that we will estimate in the next step. The potential values for p were 0, 1, and 2, d was set to 1, and q was either 0 or 1. The The ARIMA model's ordo suggested for forecasting egg production is presented in Table 2.

Table 2. The ARIMA model's ordo suggested for forecasting egg production

No	p	d	q
1	0	1	0
2	1	1	0
3	2	1	0
4	0	1	1
5	1	1	1
6	2	1	1

3.2.2. *Estimation*

The ARIMA models for predicting egg production were fitted using Statistical Package for Social Sciences (SPSS) software to estimate the model parameters. The estimation or specification step calculates the coefficients based on least-squares and likelihood methods, including the BIC value. The prediction of Magelang duck egg production using ARIMA Model analysis results are presented in Tables 3 and 4.

Table 3. The prediction of Magelang Duck egg production using The ARIMA Model Analysis (R square and error)

No	ARIMA	R square	Error			
			RMSE	MAPE	MAE	BIC
1	ARIMA (0,1,0)	0.97	4.06	10.09	3,02	2.88
2	ARIMA (1,1,0)	0.98	3.52	9.83	2.61	2.63
3	ARIMA (2,1,0)	0.98	3.31	9.75	2.47	2.54
4	ARIMA (0,1,1)	0.98	3.27	9.97	2.49	2.58
5	ARIMA (1,1,1)	0.98	3.27	10.03	2.49	2.52
6	ARIMA (2,1,1)	0.98	3.27	10.03	2.49	2.55

The selection of the optimal ARIMA model involved the evaluation of various configurations, with the ARIMA (2,1,0) model exhibiting superior performance. ARIMA (2,1,0) achieved a high R^2 value of 0.98, indicating that it accounts for 98% of the variation in egg production, thus capturing underlying trends effectively. It also exhibited low error metrics: RMSE of 3.31, MAPE of 9.75, and MAE of 2.47, making it more accurate than the other models. Additionally, it had the lowest BIC score of 2.54, indicating a good balance between model complexity and fit [7]. These low error metrics and high R^2 values collectively demonstrate the model's ability to minimize errors and produce accurate forecasts, reflecting its effectiveness in capturing data variability and generating reliable predictions with minimal inconsistencies [11].

Table 4. The prediction of Magelang duck egg production msing The ARIMA model analysis (variables and white noise assumption)

No	ARIMA	Variables		White Noise Assumption (LJung Box Q test)	
		Significant	Unsignificant	P value	White noise
1	ARIMA (0,1,0)	0	2	0.00	No
2	ARIMA (1,1,0)	2	0	0.06	Yes
3	ARIMA (2,1,0)	3	0	0.70	Yes
4	ARIMA (0,1,1)	2	0	0.85	Yes
5	ARIMA (1,1,1)	2	1	0.84	Yes
6	ARIMA (2,1,1)	2	2	0.85	Yes

As shown in Table 4, the Ljung-Box Q-test for white noise gives a p-value of 0.70, which is relatively high, indicating that the residuals resemble white noise and the model is well-specified. Among the various models, ARIMA (2,1,0) stands out with the greatest number of significant parameters, totalling 3. Thus, ARIMA (2,1,0) offers a more precise and reliable forecast than other ARIMA models, making it the best choice for analysis. Overall, ARIMA (2, 1, 0) is a reliable tool for forecasting egg production trends, providing insights for farm management and resource optimization. This information is vital for researchers and analysts assessing model reliability and effectiveness in predictions.

3.2.3. *Diagnostic*

The diagnostic process involves analysing parameter significance through visual representations and statistical measures to assess model fit [11]. Table 4 demonstrates that the ARIMA (2,1,0) model is optimal due to its incorporation of the highest number of statistically significant parameters compared to alternative models. Non-significant parameters were excluded to enhance the model's reliability and ensure that only the variables meaningfully contributing to the model's predictive capability were retained.

The model visual representation shown in Fig. 4 provides a reasonably good fit to the observed data, and the forecast shows a steady trend with some uncertainty, reflected in the widening confidence intervals. This is typical for ARIMA or similar time-series models, where future predictions are based on past trends and uncertainty grows as we move further from the last observed point. The area between the upper Control Limit (UCL) and lower Control Limit (LCL) reflects the range of uncertainty or variability in the forecast. The forecasted data stays within the UCL and LCL, it suggests that the model is performing as expected. Forecasting process with the model (2,1,0) for forecasting of future egg production indicated a good fitting of the model for prediction.

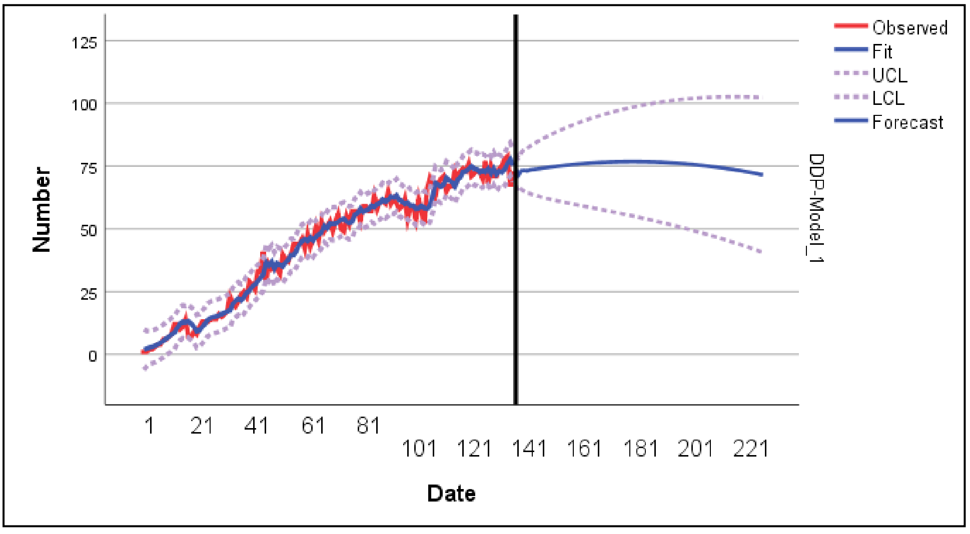

Fig 4. Actual and estimated egg production of Mageland duck based on ARIMA (2,1,0)

3.2.4. *Forecasting*

Based on analysing parameter significance through visual representations and statistical measures, we determined that the ARIMA (2,1,0) mathematical model is the most suitable for forecasting egg production in Magelang ducks. The significant AR terms (lag 1 and lag 2), as shown in Table 5, indicate that the past values of the series strongly influence future values in a negative direction. The predictive capability of the model is significantly bolstered by the presence of the constant term. These parameters confirm that the ARIMA (2,1,0) model is well-fitted and reliable for forecasting, as all components are statistically significant at a high confidence level.

Table 5. The Prediction of Magelang Duck Egg Production Using The ARIMA (2,1,0) Model Parameters

ARIMA (2,1,0) Model Parameters								
					Estimate	SE	t	Sig.
DDP (%)-Model (2,1,0)	DDP (%)	No Transformation	Constant		.869	.288	3.014	.003
			AR	Lag 1	-.678	.082	-8.263	.000
				Lag 2	-.363	.085	-4.272	.000
			Difference		1			

In general, the mathematical model based on ARIMA (2,1,0) is as follows::

$$Y_t = \underbrace{\Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} \ldots + \Phi_n Y_{t-n}}_{AR} + \underbrace{\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots \theta_q\varepsilon_{t-q}}_{MA} \qquad (1)$$

The parameter analysis results obtained using SPSS 26, presented in Table 43, were incorporated into the formula. Given that the model is differenced once, as indicated by the d value of 1, it is necessary to reformulate it in terms of the original *Yt* values. The P value of the constant was < 0.05 (0.003) and this indicated that the model fitted with the constant The mathematical model of ARIMA (2,1,0) is as follows:

$$\Delta Y_t = 0{,}896 - 0{,}678(Y_{t-1} - Y_{t-2}) - 0{,}363\,(Y_{t-2} - Y_{t-3}) + \varepsilon_t \qquad (2)$$

$$Y_t - Y_{t-1} = 0{,}896 - 0{,}678(Y_{t-1} - Y_{t-2}) - 0{,}363\,(Y_{t-2} - Y_{t-3}) + \varepsilon_t \qquad (3)$$

$$Y_t = 0{,}896 + Y_{t-1} - 0{,}678\,Y_{t-1} + 0{,}678\,Y_{t-2} - 0{,}363 Y_{t-2} + 0{,}363\,Y_{t-3} + \varepsilon_t \qquad (4)$$

So the final mathematic model was :

$$Y_t = 0{,}896 + (1 - 0{,}678)\,Y_{t-1} + (0{,}678 - 0{,}363)\,Y_{t-2} - 0{,}363 Y_{t-3} + \varepsilon_t \qquad (5)$$

$$Y_t = 0{,}896 + 0{,}322\,Y_{t-1} + 0{,}315\,Y_{t-2} - 0{,}363\,Y_{t-3} + \varepsilon_t \qquad (6)$$

Where:

Y_t : The value of Daily Duck Production (DDP) at time t
Y_{t-1} : This is the DDP at time t-1 (the previous day)
Y_{t-2} : This is the DDP at time t-2 (two days ago)
Y_{t-3} : This is the DDP at time t-3 (three days ago)
ϵ_t : The error term at time t.

Based on the mathematical model, it appears that peak egg production was likely reached around day 161 after the first egg laying recorded refers to week 23, with a value of 77%. The previous research indicated that the peak of egg production of Magelang ducks was 60,57% lower than that of Mojosari ducks were 68.86% [8]. This means that the peak production was higher in this study, which may have been caused by intensive rearing using individual cages. A previous study also [12] revealed that egg production in Alabio ducks occurs at week 24. This means that Magelang ducks are predicted to reach egg production a week earlier than Alabio ducks. The value at peak production is 77%, which is lower than that in the Shan Ma duck, which exceeds 90% and reaches 180 days old duck [9]. These results suggest that farmers must enhance the nutrition and management of their farms to optimize egg production. Duck egg production can be enhanced through various strategies, including dietary modifications[13], genetic improvements[14], and management practices[15, 16]. The ARIMA (2, 1, 0) model predicts future values based on past observations by applying first differencing (d=1) for stationarity and using two lagged observations (p=2) without a moving average component (q=0). While suitable for forecasting egg production, advanced methods, such as ETS or SARIMA, are preferred when seasonal patterns or external factors are significant for the next study. ARIMA's simplicity and reliability of ARIMA make it appropriate for datasets lacking strong seasonal variations. Egg production can be forecasted using limited data, such as a single time series, helping duck farmers manage real-world data limitations caused by poor recording. This tool helps farmers optimize resource management during peak and low production, proving useful even with limited data. Broadly speaking, it improves decisions on feeding, breeding, and resource allocation for sustainable and profitable farming.

4 Conclusion

The ARIMA (2,1,0) model was identified as the optimal fit, effectively capturing historical trends and providing accurate forecasts of future egg ducks production. These findings provide valuable insights into improving farm management, optimizing resource allocation, and enhancing egg production strategies. Thus ARIMA model can conducted by researcher or farm manager to help farmers optimize resource management during peak and low production periods, making it highly useful, even with limited data. Predictive models, such as ARIMA, enhance resource efficiency and sustainability, allowing farmers to adjust operations according to predicted trends, thereby reducing environmental impacts and increasing profitability. The ARIMA model provides valuable insights for decision-making regarding future egg production of Magelang ducks. Broadly, it enables better decision-making regarding feeding, breeding, and resource allocation for sustainable and profitable farm management. This forecasting technique is not only beneficial for this particular farm but can also be effectively utilized by ducks egg producers worldwide to forecast future performance.

This research was funded by the LPDP Scholarship, Ministry of Finance of the Republic of Indonesia, Faculty of Animal Science at UGM, and the Animal and Fisheries Department of Magelang Regency. All errors and omissions were responsible for this study. This study was approved by the Ethics Committee of the Faculty of Veterinary Medicine, UGM (ethical clearance no. 144/EC-FKH/eks./2024, issued on January 8, 2024). Funding was provided by the LPDP (Scholarship No. 202212210312291), with no involvement in the study design, data collection, analysis, or manuscript preparation. We sincerely thank Mr. Bowo, Chairman of KTT Bina Mandiri, Kalangan Ambartawang, Mungkid, Magelang, for his invaluable support in providing the research site and assisting with data collection. The data supporting this study are available from the corresponding author upon request.

Author contributions: Z. Ulfah: Study conceptualization, data collection, analysis, interpretation, and manuscript writing. D. Maharani: intellectual support and manuscript revision. H. Sasongko: Experimental design and manuscript revision.

References

1. J. Sumarmono, Duck production for food security, The 1st Animal Science and Food Technology Conference (AnSTC) in IOP Conference Series. J. Earth and Env. Sci., **372**, 012070. (2019)

2. N. Bumanis, et al., Hen Egg production forecasting: capabilities of machine learning models in scenarios with limited datasets. J. App. Sci. **13**, 7607 (2023).
3. A. Wijaya, et al. Predictions of duck egg production in 34 Indonesian provinces using backpropagation algorithm. International Conference on Informatics, Multimedia, Cyber and Informations System (ICIMCIS), Jakarta, Indonesia, 7-8 November 2023, Desember, 14 (2023), 27-32.
4. E.B. Sinu, M.A. Kleden, and A. Atti, Application of ARIMA model for forecasting national economic growth: a focus on gross domestic product data, Berekeng, J. Math. Sci. App., **18,** 1261-1272 (2024).
5. C. Deemee, et al. Statistical comparison ARIMA order performance in stock market. in 2023 7th International Conference on Information Technology (InCIT). Chiang rai, Thailand, 16-17 November 2023, January, 30 (2024), 81-85.
6. D.K Yadav, K. Soumya, and L. Goswami, Autoregressive integrated moving average model for time series analysis. International Conference on Optimization Computing and Wireless Communication (ICOCWC), Debre Tabor, Ethiopia, 29-30 January 2024, March, 21 (2024), 1-6.
7. Y. Zhang and G. Meng. Simulation of an adaptive model based on AIC and BIC ARIMA predictions. J. Phy, **2449**, 1-7 (2023).
8. H. Henrik and M. Marhayani, Egg production and quality of Magelang duck, Mojosari duck, and their reciprocal crosses. Jurnal Ilmu-Ilmu Peternakan, **30,** 180-83, (2020)
9. Z. Zhu, et al., Ovarian transcriptomic analysis of Shan Ma ducks at peak and late stages of egg production, Asian-Australian J. Anim.Sci., **30**, 1215-1224 (2017).
10. R.J. Hyndman and G. Athanasopoulos, 8.1. Stationarity and differencing. Forecasting: Principles and practices, (O texts, Melbourne, Australia, 2013).
11. F.D.M Abdallah, Role of time series analysis in forecasting egg production depending on ARIMA model. J.App. Math., **9**, 1-5 (2019).
12. T. Susanti and L. Prasetyo. Pendugaan parameter genetik sifat-sifat produksi telur itik alabio. J. Tech. Anim. Vet. dan Veteriner, (2008)
13. J. Loing, et al., Increasing duck egg production using grow enhancer treatment. International J.l of Bio. Sci., **6,** 31-35 (2015).
14. M. Oyebanjo, O. Osaiyuwu, and A. Salako, Prolactin, genetics, and bioinformatics: The trinity for improving duck egg production in Nigeria–A review. J. Anim. Sci. and Vet. Med., **8**(5), 235-246 (2023).
15. H. Sasongko, Feeding management evaluation of duck farmer groups in Brebes. in International Seminar on Tropical Animal Production (ISTAP). 2010. Yogyakarta.
16. I. Arhipova, et al. Smart platform designed to improve poultry productivity and reduce greenhouse gas emissions. in Proceedings of Sixth International Congress on Information and Communication Technology: ICICT 2021, Singapore, September, 24, (2021), 35-46.