The 11th Asian-Australasian Conference on Precision Agriculture (ACPA 11)
October 14-16, 2025, Chiayi, Taiwan

Machine Learning Prediction Models for Dual-Horizon Egg Production Forecasting

Tzu-Ling Wang, Shih-Fang Chen*

Department of Biomechatronics Engineering, National Taiwan University, Taiwan

*Corresponding Author: sfchen@ntu.edu.tw

ABSTRACT

Egg production forecasting presents significant challenges in agricultural supply chain management due to complex seasonal patterns, disease outbreaks, and market volatility. Although various forecasting models have been developed for agricultural production, limited research has systematically compared model performance across different temporal horizons or developed integrated frameworks optimized for diverse planning needs. This study develops a comparative dual-horizon machine learning framework using monthly egg production data from 2012 to 2024 provided by the Taiwan Poultry Association, with specialized models optimized for one-month operational adjustments and one-year strategic planning. For one-month forecasting, three models were tested and compared: autoregressive integrated moving average with exogenous variables (ARIMAX), seasonal autoregressive integrated moving average with exogenous variables (SARIMAX), and the extreme gradient boosting (XGBoost) algorithm. SARIMAX achieved superior performance, with a mean absolute percentage error (MAPE) of 3.04%. For one-year forecasting, a direct Long Short-Term Memory (LSTM) neural network architecture enhanced with feature engineering was employed to capture complex temporal dependencies while mitigating overfitting. The LSTM model achieved a MAPE of 3.61% and demonstrated robust trend consistency with observed production patterns (Kendall's tau, τ = 0.667). The proposed framework offers a flexible solution for egg production forecasting. The one-month model is suitable for monthly supply chain adjustments and market response, such as feed distribution and price stabilization, while the one-year model supports policy planning and capacity management. Future improvements will incorporate additional production variables such as layer age and feed conversion ratio to enhance prediction accuracy and broaden practical applications.

Keywords: Egg production, Time series analysis, Trend forecasting, SARIMAX, LSTM

INTRODUCTION

Egg production forecasting faces inherent complexity due to seasonal climate variation, major diseases (e.g., avian influenza), changes in flock size, and shifts in market demand, which can create imbalances between supply and demand and lead to price volatility. Most existing methodologies focus on a single forecasting horizon, resulting in a gap between immediate operational adjustments and long-term strategic planning. To address this, this study proposes a dual-model machine learning framework that incorporates both a one-month prediction model for monthly operational planning and a one-year prediction model for annual strategic planning. The one-month model employs algorithms such as ARIMAX, SARIMAX, and XGBoost to capture recent market dynamics and short-term fluctuations. In contrast, the one-year model utilizes LSTM networks to identify cyclical patterns in production and support long-term planning and decisions. By differentiating forecasting strategies according to planning requirements and avoiding reliance on future exogenous variables, the framework provides targeted solutions across multiple horizons, filling a critical methodological gap in agricultural forecasting.

MATERIALS AND METHODS

DATA AND EXPERIMENTAL DESIGN

This study utilized monthly egg production data from 2012 to 2024 provided by the Taiwan Poultry Association. The dataset was partitioned into training (2012–2021), validation (2022–2023), and testing (2024) periods, encompassing a total of 144 monthly observations. This comprehensive coverage ensures that seasonal variations and market cycles throughout the study period are adequately represented. Model performance was evaluated using rolling window cross-validation with multiple metrics, including MAPE for relative prediction accuracy, RMSE and MAE for absolute error quantification, and Kendall rank correlation coefficient (Kendall's tau, τ) to assess trend consistency.

ONE-MONTH FORECASTING MODELS

Three models were developed for one-month predictions. The ARIMAX model combines autoregressive and moving average components with exogenous inputs (Gonçalves et al., 2021). SARIMAX extends ARIMAX by incorporating seasonal components to model recurring production cycles (Alharbi & Csala, 2022). XGBoost, a gradient boosting algorithm (Chen & Guestrin, 2016), was applied to capture nonlinear relationships using multiple lagged features as input. External variables included seasonal indicators and market trend factors to enhance short-term prediction accuracy.

ONE-YEAR FORECASTING MODEL

For one-year forecasting, an LSTM architecture was developed to avoid dependence on future exogenous variables (Hochreiter & Schmidhuber, 1997). The model incorporated time-based features such as cyclical month encoding and annual trend indicators to capture seasonal patterns and long-term dynamics. A sliding window approach was applied, where 6 months of past data were used to directly predict the next 12 months of production. The LSTM model incorporated dropout regularization and dense layers to enhance robustness and predictive performance. In addition, feature engineering incorporated cyclical encoding of monthly patterns using trigonometric transformations to capture seasonal periodicities inherent in egg production cycles.

RESULTS & DISCUSSION

SHORT-TERM FORECASTING PERFORMANCE

Among one-month forecasting, SARIMAX demonstrated superior performance, achieving MAPE of 3.04%, outperforming ARIMAX (3.36%) and XGBoost (4.33%) (Table 1). This demonstrates that incorporating seasonal components can significantly enhance prediction accuracy. Traditional statistical models maintained advantages in scenarios where short-term fluctuations are relatively stable. The performance improvement of SARIMAX over ARIMAX confirms the importance of seasonal modeling in egg production forecasting, where biological cycles and market patterns exhibit seasonal dependencies.

LONG-TERM FORECASTING PERFORMANCE

For one-year forecasting, the LSTM model achieved strong one-year prediction accuracy with MAPE of 3.61% and Kendall's tau(τ) of 0.667, indicating good trend consistency with observed production patterns (Table 2). The model successfully captured the production decline during June and July 2024 and the subsequent rebound from August, demonstrating its capability for early warning (Fig. 1). The LSTM model's ability to maintain trend consistency while achieving competitive accuracy demonstrates its effectiveness for strategic planning applications, where understanding long-term patterns is more critical than precise point predictions.

FRAMEWORK INTEGRATION AND APPLICATIONS

The dual-horizon framework addresses a methodological gap in agricultural time series prediction by providing specialized solutions for different planning horizons. The design avoiding future exogenous variable dependence enhances practical deployment feasibility compared to conventional econometric models. In particular, this approach enables the model to be robustly applied even when future information is unavailable or difficult to obtain, which is common in real-world agricultural scenarios. By integrating feature engineering and tailored modeling strategies for both one-month and one-year models, the dual-horizon framework allow stakeholders in the poultry industry to adjust production plans more flexibly and respond to changing conditions more effectively.

Table 1 Evaluation metrics of one-month forecasting models.

Model	RMSE (boxes)	MAE (boxes)	MAPE
ARIMAX	5221.99	3917.23	**3.36%**
SARIMAX	4761.84	3526.98	**3.04%**
XGBoost	7446.85	5295.72	**4.33%**

Table 2 Evaluation metrics of one-year forecasting model.

Model	RMSE (boxes)	MAE (boxes)	MAPE	Kendall’s tau (τ)
LSTM	4972.32	4516.43	**3.61%**	**0.67**

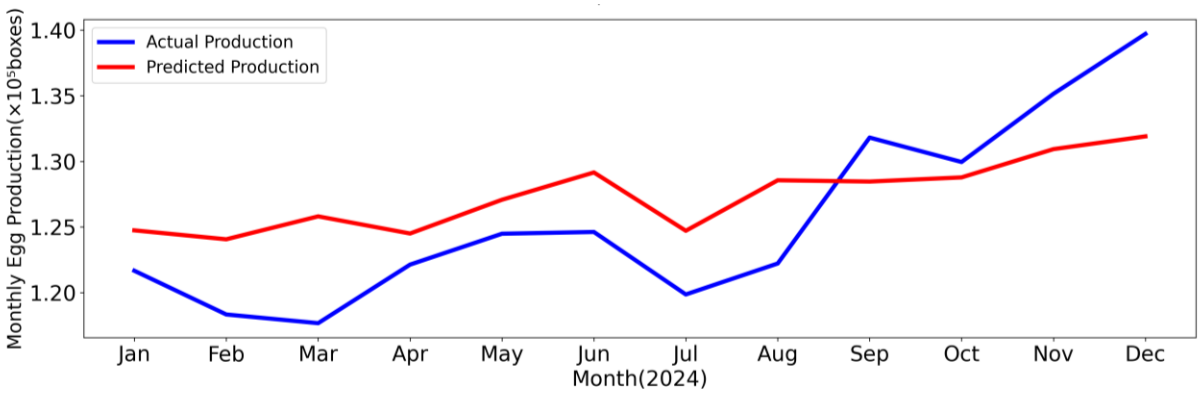

Fig.1 Demonstration of one-year prediction curve for egg production.

CONCLUSIONS

This study developed a dual-horizon machine learning framework for egg production forecasting, systematically comparing statistical and machine learning approaches across different temporal scales. The one-month SARIMAX model achieved 3.04% MAPE, while the one-year LSTM model delivered 3.61% MAPE with early warning capabilities. The framework provides a technological foundation for poultry industry decision-making, enabling data-driven approaches to both operational and strategic planning challenges. The specialized model architectures demonstrate that temporal scale optimization significantly enhances forecasting performance compared to generic time series approaches. Future research will integrate additional external factors to enhance anomaly prediction capabilities and extend applications to broader livestock industries.

ACKNOWLEDGEMENTS

This study was funded by the Ministry of Agriculture, Executive Yuan, Taiwan (114AS-2.3.2-AD-01).

REFERENCES

Alharbi, F. R., & Csala, D. (2022). A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach. *Inventions*, *7*(4), 94. https://doi.org/10.3390/inventions7040094

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). https://doi.org/10.1145/2939672.2939785

Gonçalves, J. N., Cortez, P., Carvalho, M. S., & Frazao, N. M. (2021). A multivariate approach for multi-step demand forecasting in assembly industries: Empirical evidence from an automotive supply chain. *Decision Support Systems*, *142*, 113452. https://doi.org/10.1016/j.dss.2020.113452

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735