Journal of Animal & Plant Sciences, 35(5): 2025, Page: 1235-1247
ISSN (print): 1018-7081; ISSN (online): 2309-8694
https://doi.org/10.36899/JAPS.2025.5.0105

STATE OF ART OF TOP THREE RICE PRODUCTION COUNTRIES USING GWO AND FIRST ORDER GLDM MODEL

P. Mishra[1*], M. Abotaleb[2], R. Gautam[3], W. Emam[4],Y. Tashkandy[4], A. Matuka[5], S. Ray[6], A.Nayak[7] and S. Yadav[8]

[1] Department of Mathematics and Statistics, College of Agriculture, Rewa, (J.N.K.V.V), India
[2]Department of System Programming, South Ural State University, Chelyabinsk, Russia
[3]Pandit Shambhu Nath.. Shukla, University, Shahdol, India
[4]Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[5]Department of Economics, Faculty of Economics, University of Bologna, Italy.
[6]Department of Agricultural Economics and Statistics, Centurion University of Technology and Management, Odisha, India
[7] College of Agriculture, Jabalpur, (J.N.K.V.V), India
[8]Miranda House, University of Delhi, India
*Corresponding author's E-mail: pradeepjnkvv@gmail.com

ABSTRACT

Stabilising market dynamics, directing agricultural policies, and guaranteeing food security all depend on accurate rice production predictions. The importance of creating accurate prediction models has grown due to the effects of urbanisation, policy changes, and climate variability.This study compares several time-series forecasting algorithms on the rice production data from 1961 to 2023, such as ARIMA, Holt's Linear Trend, and Simple Exponential Smoothing (SES), with improvements made utilising the Grey Wolf Optimiser (GWO) for optimal parameter selection. Furthermore, nonlinear dependencies in production patterns are taken into consideration by using the Generalised Least Deviation Method (GLDM). The findings show that the second-order GLDM model yields the most accurate forecasts for Bangladesh, whereas the GWO-ARIMA model performs best for India. Short-term autoregressive trends for China are well captured by the first-order GLDM model. The study emphasises the advantages of optimization-driven forecasting methodologies over traditional models, providing a strong basis for agricultural planning. These findings help policymakers and stakeholders implement data-driven methods for sustainable rice crop management.

Keywords: Rice production, time-series forecasting, ARIMA, Grey Wolf Optimizer, Generalized Least Deviation Method

INTRODUCTION

Rice (*Oryza sativa*) is one of the world's most significant staple crops, feeding billions of people (Qian *et al.*, 2023; Islam *et al.*, 2024).According to United States Department of Agriculture (USDA), the global production of rice is 521.52 million metric tonnes with china, india and bangladesh as the three leading countries for rice production. The Asia-Pacific area produces and consumes more than 90% of the world's rice (Bandumula *et al.*, 2018**)** , and demand is likely to increase dramatically in the future years.According to the Food and Agriculture Organisation (FAO), rice consumption in Asia has increased by more than 51% since 1995 (FAO, 2024). Furthermore, growing urbanisation and the rise of megacities with populations of 10-15 million would alter demographic patterns, increasing the number of consumers while decreasing the number of producers. These changes demand precise rice production forecasting to assure food security, stabilise market prices, and inform agricultural policy decisions.Considering the impact of climate change, technological improvements, and governmental interventions on rice farming, combining historical data with modern forecasting models is critical for making sound agricultural decisions.Timely and accurate rice yield estimates are critical for understanding food production trends, optimising rice management strategies, and assisting government decision-making (Wang *et al.*, 2021; Tripathy*et al.*, 2022). Effective forecasting models can also assist farmers in developing appropriate management plans by offering reliable insights into future production scenarios (Feng *et al.*, 2020).In theliterature, Many authors developed different time series model as well hybrid approachs to estimate the future prediction of different agricultural commodities.Annamalai and Johnson (2023) estimated the area under cultivated rice using the Autoregressive integrated moving average (ARIMA), Holt's exponential smoothing and Neural netwaork autoregression (NNAR) model to predict the next 5 years. Joseph *et al.* (2023) employed a panel regression model with climate variabilities to estimate the global rice production. ARIMA with Exogenious variable (ARIMAX) model also employed to predict the major food grains *i.e.* rice and wheat production, productivity and per capita net availability (Ray and Bhattacharyya, 2020).Similarly, the hybrid approach with machine learning *i.e.* ARIMA-LSTM with random forest algorithm was developed to predict the volatile agricultural price series (Ray *et al*., 2023; Lama *et al*., 2024). Mustaffa *et al*. (2014)used a relatively new Swarm Intelligence (SI) technique called Grey Wolf Optimizer (GWO) for short-term time series forecasting in order to predict the price of non-renewable commodities like crude oil. The results showed that the GWO generates the lowest Mean Absoluter Percentage Error (MAPE), at 5.4779%, while the Artificial Bee Colony (ABC) achieved a comparable value, at 5.4170%. The Grey Wolf Optimization Technique (GWO) also used to identify the plant disease with deep learning algorithm (Jabbar *et al*., 2023). Abotaleb(2024) investigated the use of the GLDM to ascertain the coefficients of a high-order quasilinear autoregressive model and discovered that a reliable and adaptable technique for modelling non-linear datasets is the GLDM. The GLDM efficiently captures complex data relationships and minimizes the impact of outliers by reducing the absolute differences between actual values and forecasts.

Even with the widespread use of hybrid machine learning techniques and traditional time series models (like ARIMA and Holt's Exponential Smoothing) in agricultural forecasting, there is still a significant gap in the integration of advanced statistical techniques like the First-Order GLDM and metaheuristic optimization techniques like the Grey Wolf Optimizer (GWO) for rice production forecasting. Prior research has mostly concentrated on linear models or stand-alone machine learning approaches, frequently ignoring the dynamic optimization, outlier sensitivity, and nonlinear dependencies necessary for high-accuracy forecasts in unstable agricultural systems. Although GWO has been used in some studies for plant disease detection and commodity price forecasting, its potential for improving production projections for the three primary rice-producing countries (Bangladesh, India, and China) has not yet been investigated. The crop yield modeling has not fully utilized the GLDM's resilience in managing non-Gaussian noise and structural discontinuities in agricultural time series. Our main objective of this study is to build GLDM model with 1st and 2nd order and compared with the ARIMA, Holt's linear and simple exponential smoothing model based on goodness of fit. This study introduces a novel forecasting framework capable of capturing complex, nonlinear trends in rice production while minimizing prediction errors. This approach not only addresses the limitations of conventional models but also provides a scalable solution for long-term agricultural planning under uncertainty.

MATERIALS AND METHODS

Data information: The dataset used in this study comprises annual rice production figures for three major rice-producing countries: Bangladesh, China, and India, collected from the Food and Agriculture Organization (FAO, 2024). The data spans from 1961 to 2023, capturing long-term trends in rice output. It includes production values measured in metric tons, reflecting the agricultural growth and policy shifts in each country. From the data, the information provied that China consistently exhibits the highest production, followed by India and Bangladesh.

Model Development Methodology: The overall forecasting methodology for annual rice production represented in the Figure 1. The process involves data extraction, application of multiple time-series models, evaluation using statistical metrics, and selection of the best-performing model for future forecasting.The forecasting methodology for annual rice production follows a structured multi-step approach to ensure optimal model selection and accurate predictions up to 2035. The workflow consists of the following key steps:

***Database and Rice Production Area Selection*:**The process begins with accessing historical rice production data stored in a structured database. The dataset is filtered based on country-specific rice production areas to ensure that the selected data accurately represents national agricultural trends.

***Time-Series Model Selection*:**A comprehensive set of time-series forecasting models is applied to capture production trends. These models include:

- *ARIMA*: A statistical model that captures autoregressive, differencing, and moving average components for time-series forecasting.
- *Holt's Linear Trend*: An extension of exponential smoothing that accounts for both level and trend components in time-series data.
- *Simple Exponential Smoothing (SES):* A smoothing technique that gives greater weight to recent observations for short-term forecasting.
- *GWO-Augmented Models:* ARIMA, Holt's Linear Trend, and SES models are optimized using the Grey Wolf Optimizer (GWO) to enhance parameter selection and minimize forecasting errors.

- *Generalized Least Deviation Method (GLDM):* Applied in two forms:

*GLDM First Order***:** A model optimized for short-term autoregressive dependencies.

*GLDM Second Order***:** Designed to capture longer-term trends by incorporating nonlinear interactions.

*Model Evaluation Using Statistical Metrics***:**Each model is evaluated on the training dataset using four key statistical performance indicators:

- *Mean Squared Error (MSE):* Measures the average squared difference between observed and predicted values, with lower values indicating better accuracy.
- *Mean Absolute Error (MAE):* Represents the average absolute error in predictions, providing an intuitive measure of deviation from actual values.
- *Mean Absolute Percentage Error (MAPE):* Expresses prediction errors as a percentage of actual values, facilitating interpretability.
- *R-Squared (R^2):* Quantifies the proportion of variance explained by the model, with values closer to 1 indicating better model fit.

*Model Selection and Forecasting Up to 2035:*The model with the lowest MSE is selected as the most reliable predictor for long-term forecasting. This selected model is then used to generate rice production forecasts up to the year 2035, providing insights into future agricultural trends.

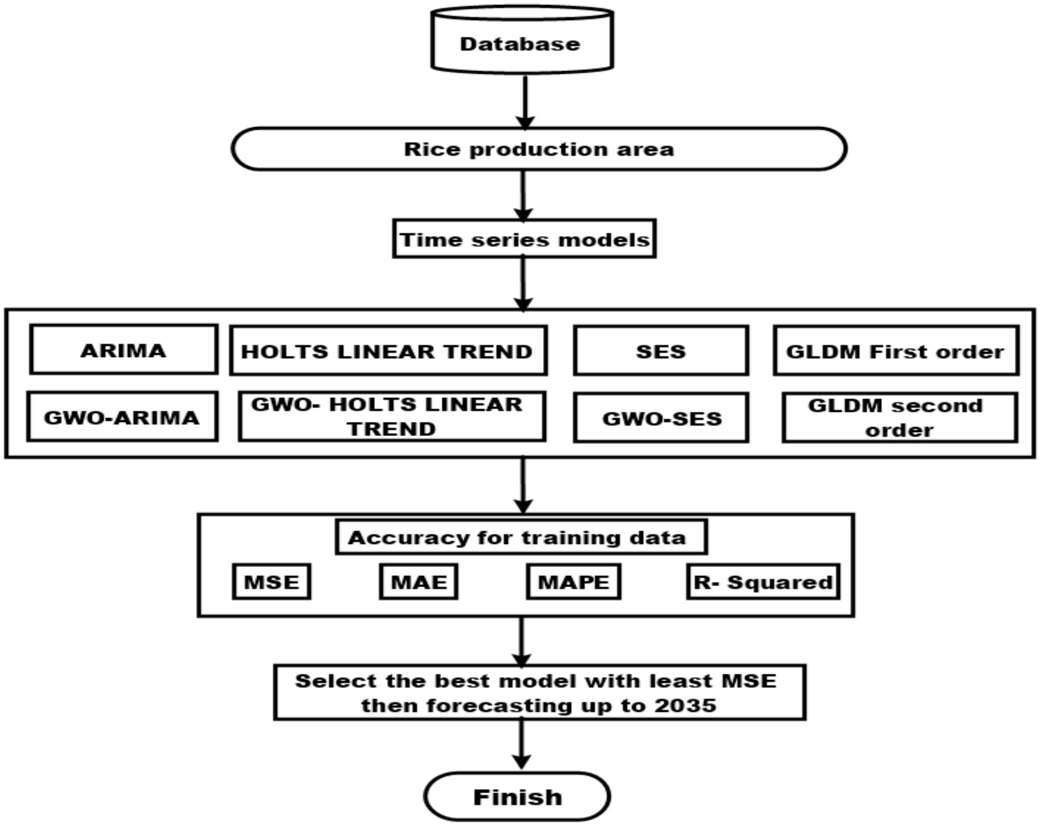
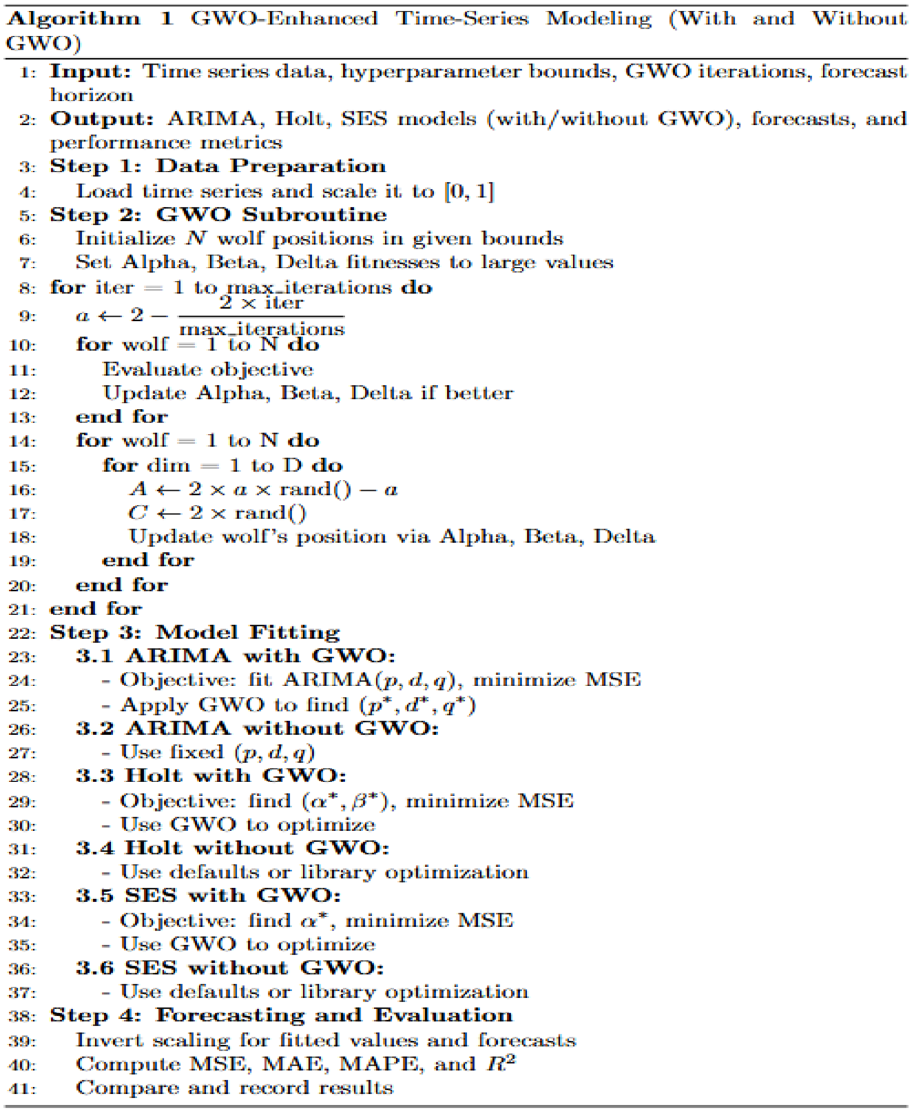


Figure 1:Workflow of the forecasting methodology for annual rice production

Algorithm of the proposed model: The optimization-driven framework for forecasting annual rice production in Bangladesh, China, and India using Grey Wolf Optimizer (GWO)-enhanced time-series modelsrepresents in Figure 2 (Algorithm 1). The algorithm integrates the GWO metaheuristic with ARIMA, Holt's Linear Trend, and Simple Exponential Smoothing (SES) models to improve forecasting accuracy by optimizing hyperparameters, minimizing Mean Squared Error (MSE), and enhancing predictive performance. The structured approach consists of four key steps: data preprocessing, GWO optimization, model fitting, and forecasting evaluation.

The first step involves loading and preprocessing historical rice production data for each country. The data is normalized to the range [0,1] to stabilize numerical computations and ensure efficient optimization during training. The second step executes the GWO subroutine, initializing a population of grey wolves with random hyperparameter values. The optimization process iterates through multiple cycles, updating the positions of the Alpha, Beta, and Delta wolves based on their fitness scores. This adaptive mechanism refines the search space by dynamically adjusting the control coefficients A and C, allowing the optimizer to balance exploration and exploitation, leading to globally optimized model parameters.The third step focuses on model fitting. The ARIMA model is trained with and without GWO, where GWO optimizes the autoregressive (p^*), differencing (d^*), and moving average (q^*) parameters to minimize MSE. Similarly, Holt's Linear Trend model and the SES model are optimized using GWO to determine the best smoothing factors (α^*, β^*). The comparative analysis includes a baseline where models are fitted without GWO, relying on default or library-based optimization techniques.The final step involves forecasting and performance evaluation. The predicted rice production values are inverse-scaled to restore their original units. The model's accuracy is assessed using statistical error metrics, including MSE, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). These metrics provide a comparative assessment of the impact of GWO-enhanced optimization on forecasting performance across the three countries.

Applying GWO-enhanced time-series modeling to annual rice production forecasting highlights the effectiveness of evolutionary optimization in improving predictive accuracy. The results indicate that incorporating GWO leads to better parameter selection, reducing forecasting errors compared to models using default configurations. These improvements are particularly relevant for agricultural planning, allowing policymakers and stakeholders to anticipate future production levels, allocate resources efficiently, and develop strategies to ensure food security and sustainable agricultural growth.



Algorithm 1 GWO-Enhanced Time-Series Modeling (With and Without GWO)

```
Input: Time series data, hyperparameter bounds, GWO iterations, forecast horizon
Output: ARIMA, Holt, SES models (with/without GWO), forecasts, and performance metrics
Step 1: Data Preparation
  Load time series and scale it to $[0, 1]$
Step 2: GWO Subroutine
  Initialize $N$ wolf positions in given bounds
  Set Alpha, Beta, Delta fitnesses to large values
for iter = 1 to max_iterations do
  $a \leftarrow 2 - \frac{2 \times \text{iter}}{\text{max\_iterations}}$
  for wolf = 1 to N do
    Evaluate objective
    Update Alpha, Beta, Delta if better
  end for
  for wolf = 1 to N do
    for dim = 1 to D do
      $A \leftarrow 2 \times a \times \text{rand}() - a$
      $C \leftarrow 2 \times \text{rand}()$
      Update wolf's position via Alpha, Beta, Delta
    end for
  end for
end for
Step 3: Model Fitting
  3.1 ARIMA with GWO:
    - Objective: fit ARIMA$(p, d, q)$, minimize MSE
    - Apply GWO to find $(p^*, d^*, q^*)$
  3.2 ARIMA without GWO:
    - Use fixed $(p, d, q)$
  3.3 Holt with GWO:
    - Objective: find $(\alpha^*, \beta^*)$, minimize MSE
    - Use GWO to optimize
  3.4 Holt without GWO:
    - Use defaults or library optimization
  3.5 SES with GWO:
    - Objective: find $\alpha^*$, minimize MSE
    - Use GWO to optimize
  3.6 SES without GWO:
    - Use defaults or library optimization
Step 4: Forecasting and Evaluation
  Invert scaling for fitted values and forecasts
  Compute MSE, MAE, MAPE, and $R^2$
  Compare and record results
```

Figure 2: Algorithm of the proposed model

RESULTS

First the collected data series need to evaluate the visualization through descriptive statistics (Table 1). The mean production values indicated that China leads with an average annual production of 165305.66 metric tons, followed by India at 111205.10 metric tons, and Bangladesh at 31349.28 metric tons. The median values showed a similar trend, with China at 179746.93, India at 111517.41, and Bangladesh at 26784.48, suggested that China's production has been consistently higher over time.The standard deviation values highlighted the degree of variability in rice production across these countries. China exhibited the highest variability at 43412.5, followed by India at 44040.3, and Bangladesh at 14469.5. This indicated that China's and India's rice production fluctuates significantly, whereas Bangladesh maintains more stable production levels. The range values further confirm this pattern, with China displaying the largest production spread (158212.4), followed by India (161070.0), and Bangladesh (45308.9). The skewness and kurtosis revealed the distribution characteristics of production series. Bangladesh and India exhibited slight right-skewness at 0.54 and 0.29, respectively, indicating occasional years of higher-than-average production. In contrast, China's production distribution is negatively skewed at -0.85, reflecting a tendency toward lower values in certain periods. The kurtosis values suggested that Bangladesh and India have relatively flat distributions with fewer extreme values (-1.21 and -0.91, respectively), while China's kurtosis (-0.37) aligns more closely with a normal distribution.These insights are crucial for agricultural policy planning, economic forecasting, and food security strategies. Advanced predictive models, such as the GLDM, can be applied to this dataset to enhance forecasting accuracy and support data-driven decision-making in the agricultural sector.

Table 1 : Descriptive Statistics of Annual Rice Production (1961–2023) for Bangladesh, China, and India

	Bangladesh	**China**	**India**
Mean	31349.28	165305.66	111205.10
Standard Error	1822.98	5469.45	5548.56
Median	26784.48	179746.93	111517.41
Standard Deviation	14469.49	43412.45	44040.30
Sample Variance	209366222.10	1884640557.00	1939548117.00
Kurtosis	-1.21	-0.37	-0.91
Skewness	0.54	-0.85	0.29
Range	45308.94	158212.35	161069.99
Minimum	13304.52	56217.60	45657.01
Maximum	58613.46	214429.95	206727.00
Sum	1975004.92	10414256.28	7005921.31
Count	63.00	63.00	63.00

After summarize the series, next we estimatedthe forecasting results for annual rice production in Bangladesh, China, and India using both statistical and optimization-driven time-series models. The comparative evaluation included the GLDM with first- and second-order formulations, as well as ARIMA, Holt's Linear Trend, and SES models, optimized with and without the GWO. The primary objective is to assess the effectiveness of these models in capturing historical trends and projecting future production levels while evaluating the impact of GWO in improving forecasting accuracy through optimal hyperparameter selection.The GLDM model, applied in both first- and second-order forms, was designed to account for nonlinear dependencies in rice production data. The second-order GLDM model was used for Bangladesh to capture long-term production dynamics, while the first-order GLDM model was applied to China, emphasizing short-term autoregressive properties. For India, the ARIMA model was augmented with GWO to fine-tune its autoregressive (p^*), differencing (d^*), and moving average (q^*) components, enhancing its predictive reliability.

To quantitatively evaluate model performance, key statistical metrics were computed, including Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2). These metrics facilitate a comparative analysis between models with and without optimization, highlighting the advantages of integrating GWO for parameter selection. Additionally, graphical representations of time series data, fitted models, and projected forecasts provided a visual interpretation of the observed and predicted values. The findings offer critical insights into the production trends of each country, supporting data-driven agricultural planning and policy development.

The compartive analysis of various forecasting models of the series was presented in Table 2. The evaluation metrics provided a quantitative assessment of forecasting accuracy, highlighting the strengths and limitations of each model in predicting future rice production trends.For Bangladesh, the forecasting results indicated that the GLDM second-order model outperforms the other models, achieving the lowest MSE (1973896.50), MAE (1032.06), and MAPE (3.94%). These values demonstrated its superior predictive accuracy in capturing the underlying patterns in Bangladesh's rice production. The R^2 value of 0.99 suggested a strong correlation between the predicted and actual production values, confirming the model's effectiveness. Traditional models such as ARIMA and exponential smoothing methods exhibit slightly higher error values, indicated that these approaches may be less effective in capturing complex variations in rice production trends over time.For China, the results suggested that the GLDM first-order model delivers the most accurate forecasts, with an MAE of 4652.53 and a MAPE of 3.08%. Despite having a higher MSE compared to the second-order GLDM model, the first-order variant demonstrated a better balance between error reduction and computational efficiency. The R^2 value of 0.98 remains consistently high, indicated strong model reliability. Alternative approaches, including ARIMA and hybrid models such as GWO-Holt's Linear Trend, shown slightly higher error margins, suggested that they may not fully capture the complexities of China's rice production fluctuations. The larger production scale of China contributed to relatively higher MSE values across all models, but the GLDM-based methods provided enhance forecasting stability by effectively modeling nonlinear trends.For India, the GWO-ARIMA model performed as the most effective forecasting approach, achieving the lowest MSE (58169833) and the best MAPE (6.90%). The MAE value of 6161.9 indicated that this model minimizes absolute deviations more effectively than other approaches. The R^2 values for India's rice production models are slightly lower than those observed for Bangladesh and China, ranging from 0.95 to 0.97, suggesting that the variability in India's rice production may be influenced by additional external factors. The GLDM second-order model also performed competitively, with relatively low error values compared to other traditional statistical models. However, the GWO-ARIMA model demonstrated superior robustness, likely due to its hybrid optimization technique that integrates evolutionary computation with time series analysis.

The performance evaluation underscored the effectiveness of the GLDM models in producing accurate forecasts, particularly for Bangladesh and China, where they consistently achieve lower error values and high predictive reliability. The findings suggested that the GLDM framework, which incorporates nonlinear recurrence equations, provided a more refine forecasting approach than conventional time series models. For India, the GWO-ARIMA model proved to be highly effective, indicated that different forecasting methodologies may be optimal depending on the specific characteristics of each country's rice production trends.

Table 2: Performance evaluation criteria of Forecasting Models.

	MSE	MAE	MAPE	R-Squared
	Bangladesh			
GWO-ARIMA	2178310.02	1123.10	4.47	0.99
ARIMA	2332541.45	1156.59	4.53	0.99
GWO- HOLTS LINEAR TREND	2305951.38	1135.62	4.36	0.99
HOLTS LINEAR TREND	2165258.17	1134.77	4.51	0.99
GWO-SES	2701031.31	1230.70	4.60	0.99
SES	2694499.06	1220.64	4.53	0.99
GLDM First order	2184030.63	1158.96	4.56	0.99
GLDM second order	**1973896.50**	**1032.06**	**3.94**	**0.99**
	China			
GWO-ARIMA	38240139.74	4347.98	2.94	0.98
ARIMA	41735773.08	4843.44	3.26	0.98
GWO- HOLTS LINEAR TREND	41671319.34	4893.68	3.33	0.98
HOLTS LINEAR TREND	40714868.72	4767.40	3.12	0.98
GWO-SES	46641346.74	5281.39	3.64	0.98
SES	46566691.10	5247.34	3.58	0.98
GLDM First order	**6357520.10**	**4652.53**	**3.08**	**0.98**
GLDM second order	37336584.45	4683.32	2.94	0.98
	India			
GWO-ARIMA	**58169833.22**	**6161.86**	**6.90**	**0.97**

ARIMA	78103731.29	7482.57	7.85	0.96
GWO- HOLTS LINEAR TREND	60976160.03	6268.31	6.72	0.97
HOLTS LINEAR TREND	60847181.82	6289.54	6.75	0.97
GWO-SES	80642640.57	7517.03	7.74	0.96
SES	80377763.11	7466.08	7.65	0.96
GLDM First order	89556533.50	6849.91	7.46	0.95
GLDM second order	70495861.52	6119.94	6.74	0.96

Note- MSE: Mean square error; MAE: Mean absolute error; MAPE: Mean absolute percentage error

After evaluating the comparision matrix, the GLDM was seleted for rice production forecasting in Bangladesh and China series (Table 3). For Bangladesh series, the model follows a second-order GLDM structure, incorporated multiple past production values along with nonlinear terms, while for the China series, the model utilized a first-order GLDM structure with quadratic dependence. The second-order GLDM model for Bangladesh is given by the equation:

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + a_3 Y_{t-1}^2 + a_4 Y_{t-2}^2 + a_5 Y_{t-1} Y_{t-2} \qquad (1)$$

where substituting the estimated coefficients results in:

$$Y_t = 0.4534 Y_{t-1} + 0.5892 Y_{t-2} + 0.0001 Y_{t-1}^2 + 0.0001 Y_{t-2}^2 - 0.0003 Y_{t-1} Y_{t-2} \ (2)$$

The coefficient $a_1 = 0.4534$ indicated that the previous year's production (Y_{t-1}) has a moderate influence on the forecasted production, contributing approximately 45.34% of its value. The coefficient $a_2 = 0.5892$ suggested that production from two years prior (Y_{t-2}) has an even greater impact on the forecast, contributing 58.92% of its value. The nonlinear coefficient $a_3 = 0.0001$ represent the squared term of Y_{t-1}, while $a_4 = 0.0001$ represent the squared term of Y_{t-2}, both of which indicated a very small but existing quadratic effect. The interaction term $a_5 = -0.0003$ represents a minor negative correlation between Y_{t-1} and Y_{t-2}, implying that the interaction between these past values slightly reduces forecasted production. The presence of second-order terms suggested that Bangladesh's rice production was influenced by longer-term dependencies, requiring the inclusion of multiple past values to improve forecasting accuracy.

For China, the first-order GLDM model is given by:

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-1}^2 \qquad (3)$$

where substituting the estimated coefficients results in:

$$Y_t = 1.1181 Y_{t-1} - 0.0000 Y_{t-1}^2 \qquad (4)$$

The coefficient $a_1 = 1.1181$ indicated that the previous year's production (Y_{t-1}) has a dominant influence, contributing 111.81% of its value, suggesting a slight growth tendency in China's rice production. The quadratic coefficient $a_2 \approx -0.0000$ implies that nonlinearity does not play a significant role in China's production trends, meaning that a linear dependence on Y_{t-1} is sufficient for forecasting purposes. The near-complete dependence on Y_{t-1} with a coefficient slightly greater than one, suggested that China's rice production follows a stable and self-sustaining trend, exhibiting minor variations over time. Unlike Bangladesh, where multiple lag terms and interactions improve forecasting accuracy, China's production can be effectively predicted using a simpler first-order structure.

Table 3: GLDM Model Coefficients for Bangladesh (Second-Order) and China (First-Order)

Coefficient	Bangladesh (Second-Order)	China (First-Order)
a_1	0.4534	1.1181
a_2	0.5892	≈ 0.0000
a_3	0.0001	–
a_4	0.0001	–
a_5	-0.0003	–

The optimized parameters and performance metrics for the Grey Wolf Optimizer-Augmented Autoregressive Integrated Moving Average (GWO-ARIMA) model applied to rice production forecasting in India are presented in Table 4. The ARIMA model was parameterized by three components: the autoregressive order (p), the degree of differencing (d), and the moving average order (q). The optimization process determined the best configuration as $p = 5$, $d = 1$, and $q = 4$, indicating that five past observations contribute to the prediction, one differencing operation was required to ensure stationarity, and four lagged forecast errors were included for improved accuracy.

The ARIMA(p, d, q) model follows the mathematical representation:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \qquad (5)$$

where Y_t represents the predicted rice production at time t, ϕ_i are the autoregressive coefficients, θ_j are the moving average coefficients, and ϵ_t is the error term. The differencing parameter $d = 1$ implies that the model was applied to the first-differenced series, meaning that Y_t is replaced by $\Delta Y_t = Y_t - Y_{t-1}$ to eliminate trends.

The optimization of the ARIMA model parameters was performed using the Grey Wolf Optimizer (GWO), a nature-inspired metaheuristic algorithm. The objective function minimized in this process is the mean squared error (MSE):

$$\min_{\{p,d,q\}} MSE = \frac{1}{n}\sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2 \qquad (6)$$

where $\hat{Y}_t$ represents the predicted values. The search agents in the GWO algorithm iteratively refine the ARIMA hyperparameters by adjusting p, d, and q within predefined bounds ($p \in [0,5]$, $d \in [0,2]$, $q \in [0,5]$). The optimization resulted in a final scaled MSE of 0.002242, indicating a high degree of accuracy in capturing India's rice production trends. The **convergence time** for the optimization process was **15.38seconds**, reflecting the computational efficiency of the GWO-based tuning approach.The obtained parameter values suggest that India's rice production time series exhibits significant short-term dependencies, as indicated by $p = 5$, which implies that five past observations are required to model the data effectively. The differencing term $d = 1$ confirmed the presence of a trend in the original series, required transformation to achieve stationarity. The moving average order $q = 4$ highlighted the necessity of incorporating past forecast errors to refine predictions. The low MSE value and the rapid convergence of the GWO-ARIMA model confirmed its suitability for agricultural forecasting applications, providing policymakers and researchers with a robust tool for data-driven decision-making.

Table 4: GWO-ARIMA Model coefficient for India

	P	d	Q	GWO MSE (scaled)	Convergence Time
(GWO-ARIMA) India	5	1	4	0.002242	15.38 sec

Figure 3 illustrated the time series representation of annual rice production in Bangladesh, displaying the original recorded data, the fitted values from the GLDM second-order model, and the forecasted production values. The blue line represents the actual historical data, reflecting the natural fluctuations in rice production over the years. The black line corresponds to the fitted values generated by the second-order GLDM model, demonstrating its ability to closely capture the underlying trend and variability of rice production. The red line extended beyond the historical data, depicting the future projections based on the GLDM model.The second-order GLDM model effectively accounts for both linear and nonlinear dependencies by incorporating multiple past observations and interaction terms, ensuring that the forecast aligns with the long-term production trends. The fitted values closely follow the original data, highlighting the model's capability in reducing residual errors. The forecasted trend indicates continued growth in rice production, suggesting an upward trajectory driven by advancements in agricultural techniques, improved irrigation infrastructure, and policy interventions aimed at enhancing yield efficiency.The smooth transition from fitted values to forecasted trends reflects the stability of the GLDM second-order model in capturing the historical patterns and extending them into future estimates. This projection was provided valuable insights for policymakers, researchers, and stakeholders in agricultural planning, enabling strategic decision-making to ensure food security and resource allocation. The visualization in Figure 3 was underscored the effectiveness of the GLDM framework in modeling complex agricultural time series data and generating reliable future forecasts.

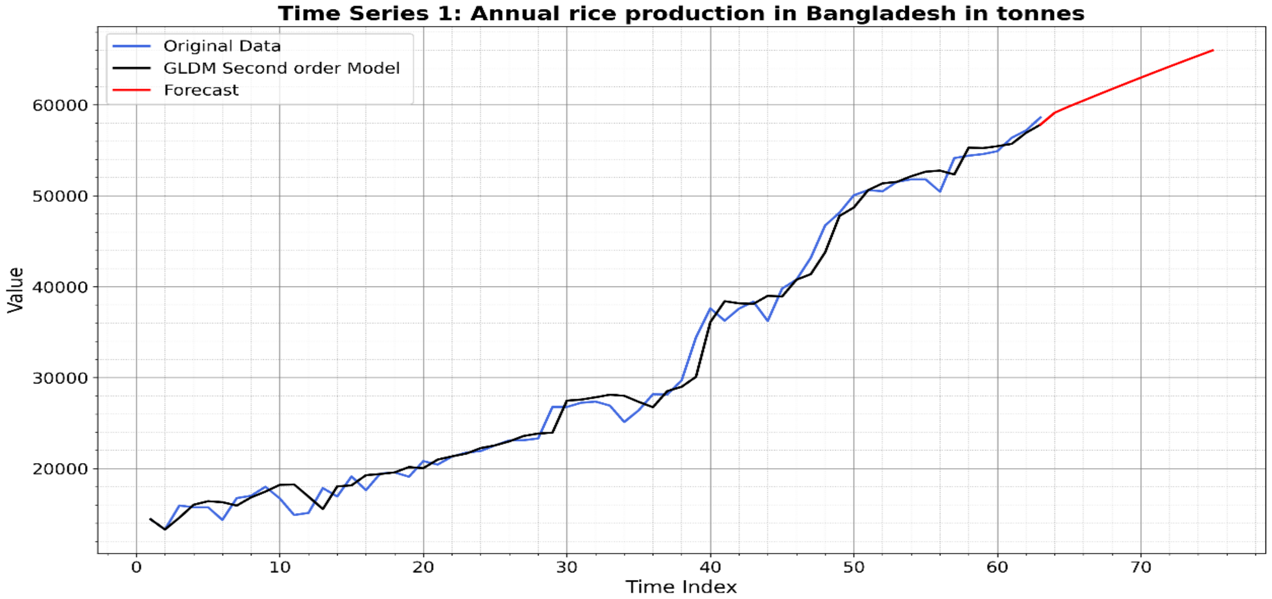

Time Series 1: Annual rice production in Bangladesh in tonnes

Figure 3: Time Series Plot of Annual Rice Production in Bangladesh: Original Data, GLDM Second-Order Model, and Forecast

The time series analysis of annual rice production in China, illustrating the historical data, the fitted values from the GLDM first-order model, and the projected forecast are given in Fig. 4. The original recorded data were shown in blue, while the black line represented the fitted values generated by the first-order GLDM model. The red line depicted the forecasted rice production beyond the historical data. The first-order GLDM model effectively captured the underlying trend in China's rice production, relying primarily on the immediate past values for forecasting. The fitted values aligned closely with the historical data, highlighting the model's ability to replicate the observed production dynamics. Unlike Bangladesh, which exhibited a strong growth pattern, China's rice production appeared to have reached a saturation phase, with the forecast suggesting a slight stabilization or marginal decline in output over the coming years. The relatively stable forecasted trend indicated that China's rice production system was highly autoregressive, with production levels heavily influenced by the previous year's values. The model's first-order structure implied minimal dependency on long-term historical fluctuations, making it suitable for short-term forecasting but less adaptable to major external shocks. The slight downward trend in the projections may have been attributed to factors such as land-use policies, climate-related constraints, or a shift in agricultural priorities toward alternative crops. The visualization in Figure 4 underscored the practical application of the GLDM first-order model in forecasting rice production trends in China. The results suggested that while China had maintained a high level of rice output, future production might stabilize or decline slightly, necessitating adaptive agricultural policies and sustainable resource management strategies to maintain food security. The findings provided valuable insights for policymakers and agricultural researchers, ensuring data-driven decision-making in planning future rice production strategies.

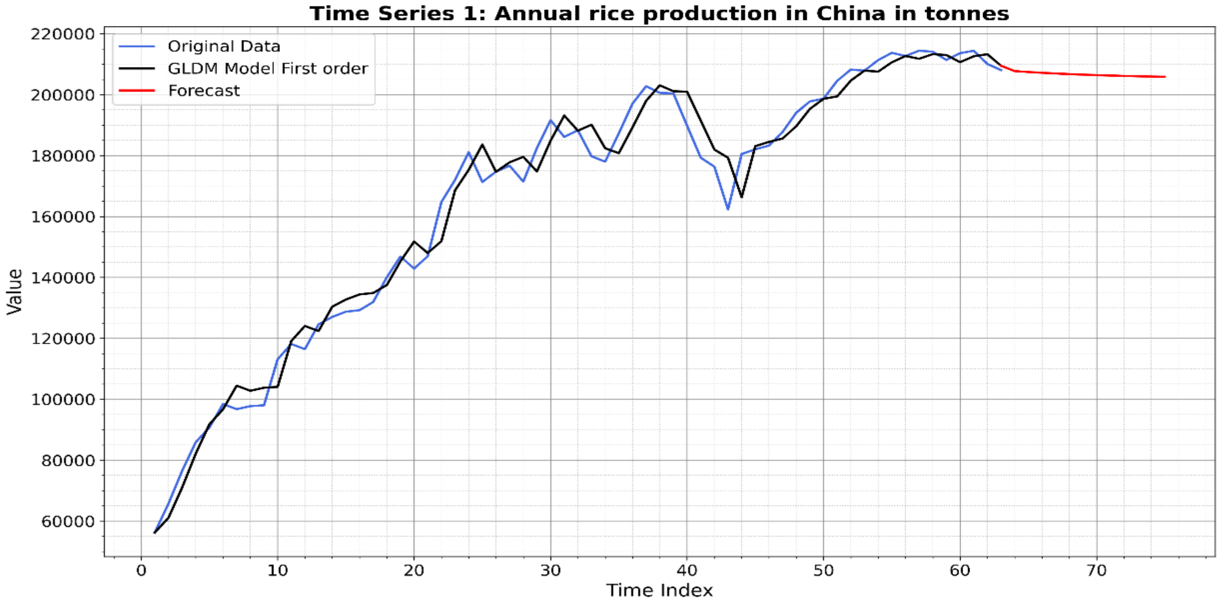


Figure 4: Time Series Plot of Annual Rice Production in China: Original Data, GLDM First-Order Model, and Forecast

The time series representation of annual rice production in India, illustrating the historical data, the fitted values from the ARIMA+GWO model, and the projected forecast are given in Fig. 5. The original recorded data were displayed in blue, while the red line represented the fitted values derived from the ARIMA+GWO model. The forecasted rice production was shown in orange, extending the trend beyond the observed dataset. The ARIMA+GWO model demonstrated a strong capability in capturing the underlying patterns and variability in India's rice production. The fitted values aligned closely with the actual data, confirming the model's ability to reduce forecasting errors. The GWO enhanced ARIMA's parameter selection by minimizing the MSE, leading to an improved balance between overfitting and underfitting. The high accuracy of the fitted values suggested that the model effectively integrated autoregressive (AR) and moving average (MA) components to account for both short-term dependencies and long-term trends. The forecasted trajectory indicated a sustained increase in rice production, reaching progressively higher levels over the coming years. This upward trend reflected the impact of technological advancements, improved irrigation infrastructure, and policy-driven initiatives aimed at enhancing agricultural productivity. The smooth transition from fitted values to forecasted projections highlighted the model's robustness in handling dynamic agricultural time series data. Figure 5 provided valuable insights into India's future rice production trends, assisting policymakers and researchers in making informed decisions. The ARIMA+GWO framework proved to be a reliable tool for long-term agricultural forecasting, allowing stakeholders to anticipate future production levels and implement strategies for sustainable food security and resource management.

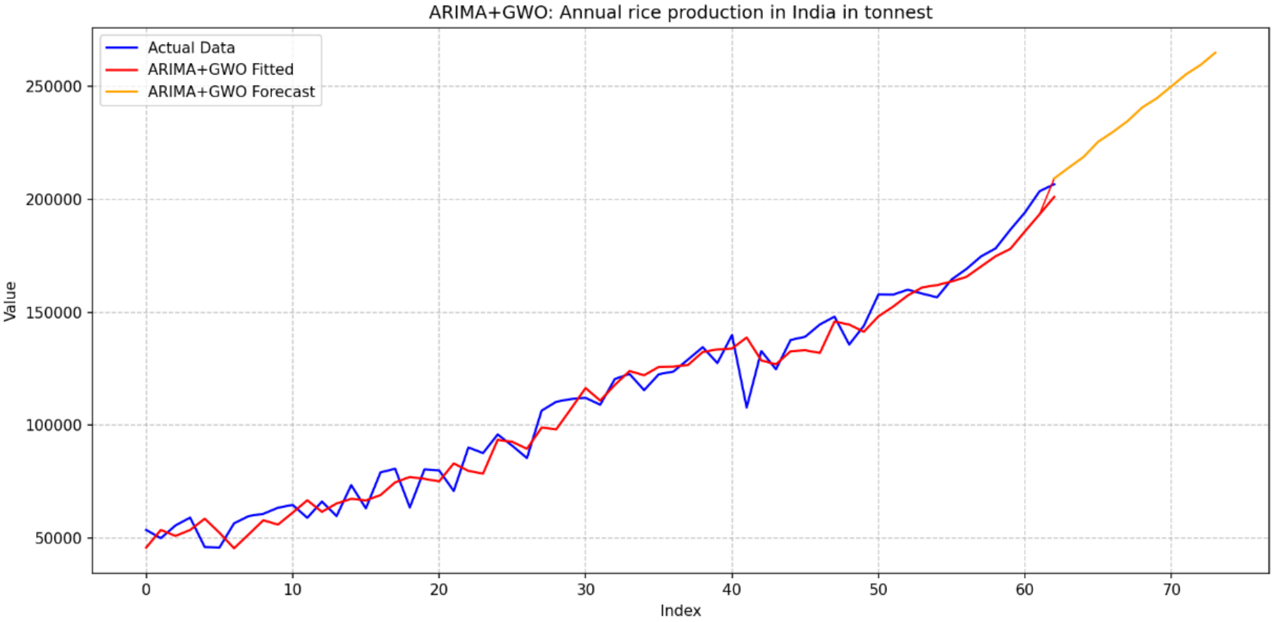


Figure 5: Time Series Plot of Annual Rice Production in India: Original Data, ARIMA+GWO Fitted Model, and Forecast

The projected rice production for Bangladesh, China, and India from 2024 to 2035, utilizing the GLDM and the GWO-ARIMA model is given in Table 5. The forecast for Bangladesh was derived using a second-order GLDM model, while China's predictions followed a first-order GLDM structure. India's future rice production was forecasted using the GWO-ARIMA model, which optimally selected ARIMA parameters through metaheuristic optimization.

Table 5: Future Rice Production (*metric tons*) Forecast for Bangladesh, China, and India

Year	Bangladesh **GLDM second order**	China **GLDM First order**	India **GWO-ARIMA**
2024	59140.5968	207727.2109	209301.7217
2025	59832.6509	207431.0356	214126.3508
2026	60468.3592	207170.6231	218798.5506
2027	61112.1591	206941.5719	225557.9995
2028	61745.6375	206740.0405	229853.5767
2029	62373.7438	206562.6727	234666.7625
2030	62994.7810	206406.5326	240710.6605
2031	63608.9619	206269.0496	244811.6033
2032	64215.9685	206147.9712	250106.2179
2033	64815.6509	206041.3220	255445.1573
2034	65407.8268	205947.3684	259619.7162
2035	65992.3392	205864.5884	265026.2622

For Bangladesh, the second-order GLDM model forecasted an increasing trend in rice production, starting at 59,141 thousand tons in 2024 and reaching 65,992thousand tons in 2035. The model captured long-term dependencies in historical production values, incorporating quadratic and interaction terms that accounted for nonlinearity. The forecasted values exhibited a steady growth pattern, reflecting the country's ongoing advancements in agricultural productivity, improved irrigation systems, and policy-driven efforts to enhance yield efficiency. For China, the first-order GLDM model predicted a relatively stable production trend, with rice output decreasing slightly from 207,727 thousand tons in 2024 to 205,865 thousand tons in 2035. The forecasted values suggested a marginal decline in production over the years, which may have indicated the saturation of arable land, shifting agricultural policies, or the adoption of alternative crops due to economic and environmental considerations. The first-order structure of the GLDM model implied that China's rice production followed a highly autoregressive pattern, where current production was largely dictated by the previous year's output, with minimal influence from higher-order dependencies. For India, the GWO-ARIMA model forecasted a consistent increase in rice production, beginning at 209,302 thousand tons in 2024 and rising to 265,026 thousand tons in 2035. The GWO-ARIMA model optimally selected the ARIMA parameters to minimize forecast error, capturing both autoregressive and moving average components. The increasing trend in India's forecasted production highlighted the country's expanding agricultural infrastructure, advancements in farming technologies, and government initiatives to ensure food security. The integration of GWO allowed the ARIMA model to achieve a high degree of accuracy while reducing computational overhead, making it an effective tool for long-term forecasting. The comparative analysis of these forecasts revealed distinct trends across the three countries. Bangladesh and India exhibited a positive growth trajectory, whereas China's production remained stable with a slight downward tendency. The second-order GLDM model for Bangladesh captured multi-period dependencies, making it suitable for dynamic production patterns. The first-order GLDM model for China suggested that rice production was largely self-sustaining with minor fluctuations. Meanwhile, the GWO-ARIMA model for India effectively captured both short-term fluctuations and long-term growth trends, demonstrating its flexibility in forecasting applications.

DISCUSSION

The forecasting results obtained from the GLDM and the ARIMA+GWO models provided significant insights into the future trends of rice production in Bangladesh, China, and India. Each model was selected based on the historical behavior of rice production in the respective countries, ensuring optimal accuracy and predictive reliability. For Bangladesh, the second-order GLDM model effectively captured long-term dependencies and nonlinear relationships within the historical data. The model's structure incorporated interaction terms, allowing it to adapt to fluctuations in production trends. The forecasted values indicated a steady increase in rice production, suggesting continued agricultural growth supported by technological advancements and improved farming techniques. The upward trend underscored the impact of policy-driven efforts and investments in irrigation infrastructure, ensuring sustained productivity in the coming years. In China, the first-order GLDM model was employed, as it efficiently captured (Abotaleb *et al*., 2024; Ma *et al*., 2024) short-term dependencies while maintaining a stable forecast. The model's reliance on the most recent observations highlighted the autoregressive nature of China's rice production system. The forecasted values suggested a slight decline in production over the next decade, which could be attributed to factors such as shifting agricultural priorities, land-use policies, or climate-related constraints. The stabilization trend observed in the projections indicated that China's rice production had reached a saturation point, necessitating adaptive strategies to maintain or enhance output levels. For India, the ARIMA+GWO model provided an optimal forecasting framework by leveraging both statistical modeling and metaheuristic optimization. The GWO improved ARIMA's parameter selection, minimizing forecasting errors and enhancing model performance. The results indicated a continuous upward trajectory in rice production, driven by advancements in agricultural technologies, government initiatives, and increasing efficiency in resource utilization. The strong alignment between the fitted and actual values confirmed the model's robustness in capturing seasonal variations and long-term growth patterns. The increasing trend in India's forecasted production highlighted the country's expanding agricultural infrastructure and its potential to meet future food security demands.

A comparative analysis of the forecasting models revealed distinct trends across the three countries. While Bangladesh and India exhibited a clear growth pattern in rice production, China's projections suggested stabilization or a slight decline. These variations highlighted the importance of country-specific agricultural policies and resource management strategies. The models used in this study provided a valuable framework for decision-makers, enabling proactive planning to mitigate risks associated with climate change, population growth, and global market fluctuations. The results also emphasized the effectiveness of integrating advanced forecasting techniques with optimization algorithms. The GLDM models for Bangladesh and China successfully captured the underlying dynamics of their respective production systems, while the ARIMA+GWO model for India demonstrated superior adaptability and precision (Bagalkot & Naik, 2024). Future research could focus on incorporating additional external variables, such as climate conditions and economic factors, to further enhance forecasting accuracy. Additionally, exploring hybrid modeling approaches that combine statistical and machine learning techniques might provide even more reliable and interpretable predictions. The findings presented in this study had significant implications for policymakers, researchers, and agricultural stakeholders. Understanding future production trends was crucial for ensuring food security, optimizing resource allocation, and developing sustainable agricultural policies. By utilizing advanced forecasting methodologies, decision-makers could implement data-driven strategies to enhance productivity, mitigate risks, and support long-term agricultural development.

Conclusion: In this research, the GLDM and an improved ARIMA model combined with the GWO (ARIMA+GWO) are two sophisticated forecasting techniques used to examine rice production trends in Bangladesh, China, and India. With Bangladesh and India displaying upward trends, the results demonstrate clear country-specific patterns that point to continuous advancements in agricultural infrastructure, policy, and technology. The ARIMA+GWO model's optimization greatly improved India's estimates, whereas a second-order GLDM model best represented Bangladesh's growth. The first-order GLDM model, on the other hand, accurately depicted China's moderately falling or comparatively constant trend, which suggested a potential output plateau brought on by policy or environmental restrictions. The study emphasizes the importance of combining conventional forecasting models with metaheuristic optimization in order to improve accuracy. Further improvement can be made by utilizing the machine learning algorithm with these technique for forecasting different time series patterns. These insights offer strategic value for policymakers to develop targeted agricultural policies that promote sustainability and food security. The research also highlights the potential of incorporating external variables and hybrid modeling techniques in future studies to further improve predictive performance.

DECLARATIONS

Data Availability; On reasonable request, the corresponding author will provide data supporting the study’s results. The raw data cannot be made public for reasons of confidentiality and privacy. However, researchers who satisfy the requirements for access to confidential data can be given access to aggregated and anonymized data as well as the statistical analysis codes. To request access to the data, interested researchers can get in touch with the corresponding author at WhatsApp+919,560,073,489 or pradeepjnkvv@gmail.com.

Competing interests: The authors declare no competing interests.

Acknowledgements: The study was fundedby Ongoing Research Funding Program (ORF-2025-749), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest: All authors declare no competing interests.

REFERENCES

Abotaleb, M.O. (2024). Solving the optimizing parameters problem for non-linear datasets using the high-order general least deviations method (GLDM) algorithm. Comp. M. Differ.Equ.1-30.https://doi.org/10.22034/cmde.2024.62441.2751

Annamalai, N. and A. Johnson (2023). Analysis and forecasting of area under cultivation of rice in India: Univariate time series approach. Sn.Comput. Sci. 4(193): 1-15. https://doi.org/10.1007/s42979-022-01604-0

Bandumula, N. (2018). Rice production in Asia: Key to global food security. Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci. 88: 1323-1328.https://doi.org/10.1007/s40011-017-0867-7.

Bagalkot, S.S. and N. Naik (2024). Novel grey wolf optimizer-based parameters selection for GARCH and ARIMA models for stock price prediction. PeerJComput. Sci. 10: e1735.https://doi.org/10.7717/peerj-cs.1735

Booranawong, T. and A. Booranawong (2017). Simple and double exponential smoothing methods with designed input data for forecasting a seasonal time series: in an application for lime prices in Thailand. Suranaree J. Sci. Tech. 24(3): 301-310.

FAO- FAOSTAT (2024). https://www.fao.org/faostat/en/#home [Accessed on 12/04/24]

Feng, P., B. Wang, D. L. Liu, C. Waters, D. Xiao, L. Shi and Q. Yu (2020). Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. Agril. For. Meteor. 285: 107922.https://doi.org/10.1016/j.agrformet.2020.107922

Islam, M.M., G.M.S. Himel, M.S. Uddinand and M.G. Moazzam (2024). A visual dataset for recognition of rice varieties. Data in Brief, 54: 110442 https://doi.org/10.1016/j.dib.2024.110442

Jabbar, A.N. and H. Koyuncu (2023). Deep learning and grey wolf optimization technique for plant disease detection: a novel methodology for improved agricultural health. Traitement du Sig. 40(5): 1961-1972.https://doi.org/10.18280/ts.400515

Joseph, M., S. Moonsammy, H. Davis, D. Warner, A. Adams and T.D.T. Oyedotun (2023). Modelling climate variabilities and global rice production: A panel regression and time series analysis. Heliyon. 9(4): e15480https://doi.org/10.1016/j.heliyon.2023.e15480

Lama, A., S. Ray, T. Biswas, L.Narsimhaiah, Y.S. Raghav, P. Kapoor, K.N. Singh, P. Mishraand and B. Gurung (2024). Python code for modeling ARIMA-LSTM architecture with random forest algorithm. Soft. Impacts. 20: 100650.https://doi.org/10.1016/j.simpa.2024.100650

Ma, C., K. Gu, Y. Zhao and T. Wang (2024). Research on highway traffic flow prediction based on a hybrid model of ARIMA–GWO–LSTM. Neural Comput. App. 1-20.https://doi.org/10.1007/s00521-024-10550-2

Mustaffa, Z., Y. Yusof and S.S. Kamaruddin (2014). Application of grey wolf optimizer for time series forecasting. The Asian Conf. on Socie. Edu. Tech. 2014: Offic. Conf. Proc.

OECD/FAO (2024), OECD-FAO Agricultural Outlook 2024-2033, Paris and Rome, https://doi.org/10.1787/4c5d2cfb-en

Qian, H., X. Zhu, S. Huang, B. Linquist, Y. Kuzyakov, R.Wassmann, K.Minamikawa, M. Martinez-Eixarch, X. Yan, F. Zhou, B.O. Sander, W. Zhang, Z. Shang, J. Zou, X. Zheng, G. Li, Z. Liu, S. Wang, Y. Ding, K.J. van Groenigen and Y. Jiang (2023). Greenhouse gas emissions and mitigation in rice agriculture. Nature Rev. Earth. Environ. 4(10): 716–732.https://doi.org/10.1038/s43017-023-00482-1

Ray, S. and B. Bhattacharyya (2020). Statistical modeling and forecasting of ARIMA and ARIMAX models for food grains production and net availability of India. J. Expet. Bio. Agril. Sci. 8(3): 296-309. http://dx.doi.org/10.18006/2020.8(3).296.309

Ray, S., A. Lama, P. Mishra, T. Biswas, S.S. Dasand B. Gurung (2023). An ARIMA-LSTM model for predicting volatile agricultural price series with random forest technique. Appl. Sof. Comp. 149: 110939.https://doi.org/10.1016/j.asoc.2023.110939

Tripathy, R., K. Chaudhari, G. Bairagi, O. Pal, R. Das and B.K. Bhattacharya (2022). Towards fine-scale yield prediction of three major crops of India using data from multiple satellites. J. Ind. Socie. Remo. Sen. 50: 271–284.https://doi.org/10.1007/s12524-021-01361-2

Wang, F., X. Yao, L. Xie, J. Zheng and T. Xu (2021). Rice yield estimation based on vegetation index and florescence spectral information from UAV Hyperspectral Remote Sensing. Remo. Sen. 13(17): 3390. https://doi.org/10.3390/rs13173390.