



RESEARCH ARTICLE

Development of forecasting models for area and production of pineapple (*Ananas comosus* (L.) Merr.) in Assam

Akankhya Sandilya¹, Borsha Neog^{1*}, Horindra Gogoi², Pallabi Bora³, Utpal Barman³, Mousumi Phukon⁴,
Trishnalee Saikia⁵ & Tonmoy Neog⁶

¹Department of Agricultural Statistics, Assam Agricultural University, Jorhat 785 013, Assam, India

²Department of Agricultural Economics, Assam Agricultural University, Jorhat 785 013, Assam, India

³Department of Extension Education, Assam Agricultural University, Jorhat 785 013, Assam, India

⁴Department of Entomology, Assam Agricultural University, Jorhat 785 013, Assam, India

⁵Department of Agricultural Economics, Krishi Vigyan Kendra, Assam Agricultural University, Jorhat 785 013, Assam, India

⁶Department of Geophysics, Banaras Hindu University, Varanasi 221 005, Uttar Pradesh, India

*Correspondence email - borsha.neog@aau.ac.in

Received: 28 February 2026; Accepted: 16 June 2026; Available online: Version 1.0: 04 August 2026

Cite this article: Akankhya S, Borsha N, Horindra G, Pallabi B, Utpal B, Mousumi P, Trishnalee S, Tonmoy N. Development of forecasting models for area and production of pineapple (*Ananas comosus* (L.) Merr.) in Assam. Plant Science Today. 2026; 13(sp5): 1-7. <https://doi.org/10.14719/pst.14285>

Abstract

Pineapple (*Ananas comosus* (L.) Merr.) is an economically important fruit crop in India with substantial regional production significance. Assam is the leading pineapple producing state of the country. Accurate forecasting of pineapple area and production is critical for agricultural planning and policy formulation. This research included trend pattern analysis and forecasting model development for pineapple area and production data of Assam. Four trendline models were developed for both the area and production series, out of which the logarithmic model was identified as the most suitable for pineapple area and exponential model for pineapple production. Three forecasting approaches were applied: auto-regressive integrated moving average model (ARIMA), artificial neural network (ANN) and non-linear support vector regression (NLSVR). A hybrid modelling framework was employed to integrate linear and non-linear components. Comparative evaluation showed that the ANN model outperformed ARIMA and NLSVR as standalone models, while the ARIMA-ANN hybrid outperformed the ARIMA-NLSVR hybrid for both series. The performances of all models were compared through root mean square error (RMSE) and mean absolute percentage error (MAPE) measures. For pineapple production, ANN model exhibited lowest RMSE and MAPE than among all evaluated models (ARIMA, NLSVR, ARIMA-ANN and ARIMA-NLSVR). On the other hand, hybrid ARIMA-ANN model emerged as best model for pineapple area series than the other four forecasting models. The findings of this study may support agricultural planners, policymakers and stakeholders in developing effective production strategies, resource allocation plans and long-term forecasting frameworks for pineapple cultivation in Assam.

Keywords: ARIMA-ANN model; Assam; crop area forecasting; pineapple; production forecasting; time series analysis

Introduction

Pineapple is a nutritionally valuable fruit crop with recognised health benefits. India is a major global pineapple producer, despite not ranking among the top three pineapple producer of world (1). Although Assam has favourable climatic conditions and leads pineapple production in India, its output remains lower than that of several other major fruit crops. Also, the total area of Assam under pineapple cultivation is also very less and much scope is available for enhancing both the area and production of this crop. Accordingly, the study focuses on trend analysis and forecasting of pineapple area and production in Assam.

Development of forecasting model for crop yield prediction gains much importance in recent two decades. Robust forecasting models improve prediction accuracy and support informed decision-making in agricultural planning. It is helpful for farmers as well as policy makers and traders. Numerous studies have explored forecasting model development. As for example

researchers builds a hybrid time series model for forecasting banana production in Karnataka (2). That study applied auto-regressive integrated moving average model (ARIMA), artificial neural network (ANN) and non-linear support vector regression (NLSVR) models, demonstrating superior performance of the hybrid ARIMA-NLSVR approach.

Similarly, in a study, researchers compared the performance of ARIMA, ANN and hybrid ARIMA-ANN for forecasting of rice yield. Results showed that hybrid ARIMA-ANN model explains the rice yield data series better compared to other individual models (3). In one of the studies on forecasting, developed three models for forecasting yield of mango and banana in Tamil Nadu (4). This study also disclosed that hybrid ARIMA-NLSVR model was superior to other developed models that is ARIMA, time delay neural network (TDNN), NLSVR, hybrid ARIMA-TDNN. As hybrid model was seen to give more precise forecasts than the single models in many studies, thus hybrid approach was also adopted in this research.

Time series data on pineapple area and production in Assam (1987–88 to 2023–24) were obtained from indiastatagri.com and the Statistical Handbook of Assam (5). All analyses were conducted using R studio and SPSS software. For development of ANN model, 'Forecast' package of R studio software model was used. On the other hand, 'TSSVM' package was used to develop NLSVR model.

Materials and Methods

Auto-regressive integrated moving average model (ARIMA)

Auto-regressive integrated moving average model (ARIMA) is a classical linear model comprised of three components namely autoregressive (AR), integrated (I) and moving average (MA). Autoregressive (AR) component explains the effect of previous lags upon current observation. The order of AR component is denoted as 'p'. The model is given by,

$$X_t = \alpha + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

where, α = constant

X_{t-i} = lag values

ϕ = coefficients

ε = random error i.i.d, $N(0, \sigma^2)$

The MA component refers to the influence of random shocks of previous lags on current observation. 'q' is the MA parameter. Mathematical model is,

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

where, μ = constant

ε_{t-1} = random shock values

θ = coefficients

The differencing component records the count of differencing process conducted on non-stationarity series to make it a stationary series. These three components were combined to get the typical ARIMA model which is given by,

$$\phi_p(B)(1-B)^d X_t = \alpha + \theta_q(B)\varepsilon_t$$

α = constant

X_t = variable

ϕ = coefficients of AR model

ε_t = random shock values

θ = coefficients of MA model

This model was developed by Gorge Box and Gwilyn Jenkins in 1970 (6), who forwarded four steps of ARIMA modelling; identification stage, estimation stage, diagnostic checking stage and forecasting stage. In the identification step, the orders of the potential ARIMA model were identified. In the estimation phase, coefficients of the identified model were estimated using method of maximum likelihood or iterative least square method. The performance of the selected ARIMA model was compared to some other related models as well in this step. Third stage is diagnostic checking which includes validity evaluation of model. If the residual series follows the characteristic of white noise, then that ARIMA model is regarded as good fit and is promoted to fourth step. However, if the residual series is not a white noise, the process of modelling should be repeated as long as the best model is not identified. In the fourth step, forecasts were generated using the best ARIMA model.

Artificial neural network (ANN)

It is a machine learning (ML) model, suitable for identifying non-linear patterns present in any dataset. It comprises three layers; input layer, hidden layer, output layer. The fundamental unit block of ANN structure is nodes or neurons. Each layer is composed of one or more neurons. There were three important hyperparameters of ANN architecture; number of hidden layers, activation function and training epochs. These three hyperparameters were adjusted repeatedly to minimise the error and get the most suitable ANN model. This model works adopting Marquardt Levenberg backpropagation algorithm. The structure of time delay neural network is defined by I: Hs: Ol where I is the number of nodes in the input layer, s is the sigmoid activation function, O is the number of nodes of the output layer and its activation function is linear. The mathematical form of this model is shown in Fig. 1.

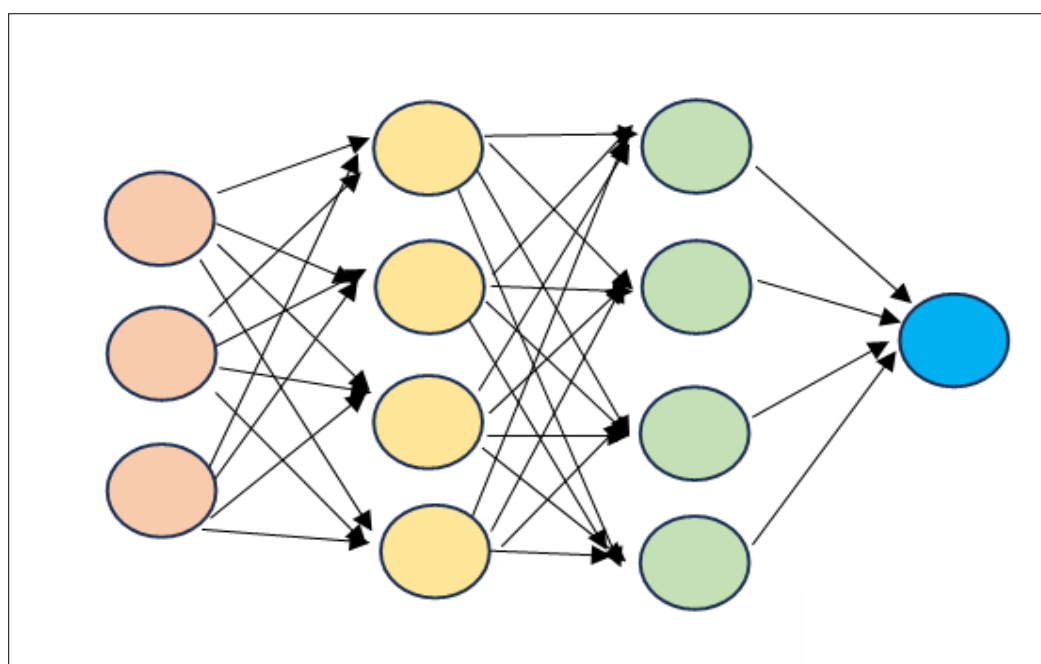


Fig. 1. Artificial neural network (ANN) architecture.

As there is no predefined indicator to identify best ANN model for a series, thus repetitive evaluation was conducted. The dataset was partitioned using a chronological (sequential) 80: 20 split, where the first 80 % observations were used as the training set for model development and the remaining 20 % latest observations were used as the testing set for forecast evaluation. Random splitting was avoided to preserve the temporal dependency structure of the time series data. The same training and testing partitions were consistently used across all forecasting models to ensure unbiased comparative assessment. Different models were developed and their performances were recorded through root mean square error (RMSE) and mean absolute percentage error (MAPE) measures. The model acquiring lowest error was selected as best ANN model for that series.

Non-linear support vector regression (NLSVR)

Non-linear support vector regression (NLSVR) is a model, which utilises ML algorithm with regression concept to identify non-linear patterns. It was developed by Vapnik in 1997 (7). Initially a kernel function is applied to the original data series to transfer it to a high dimensional feature space. Then the model assumes a non-linear regression plane upon the data points in that feature space. This regression plane is called optimal hyperplane and the data points are referred to as support vectors. There is a tolerance area around the plane known as epsilon tube. The gap between epsilon tube boundary to the hyperplane is called margin. The distance of those support vectors that fall outside the epsilon tube from the tube boundary is defined as slack variables. The principle of this model is to determine the coefficients of optimal hyperplane in such a way that it fits most of the support vectors within the epsilon tube by balancing the regularised term (length of margin) and empirical error (prediction error outside the epsilon tube). Regularised term and empirical error term is combined to form the regularised risk function which is minimised with respect to the unknown parameters of optimal hyperplane to get the optimised estimates of the coefficients.

Hybrid model

This is the most popular method in recent time, developed by Zhang in 2003 (8), which withstands the limitations of individual linear and non-linear model. This model assumes that each series is a combination of linear and non-linear component. This approach facilitates combining linear model and non-linear model simultaneously to capture the linear as well as non-linear component of a series. The hybrid model development includes three major steps. In the first step, linear model is applied over a time series Y_t . This model fits the linear component and gives the linear estimated series which is denoted as $\hat{L}_t$. In the second part, the non-linear residuals ($\epsilon_t = Y_t - \hat{L}_t$) obtained from the linear model are modelled using non-linear models. The estimated nonlinear series is denoted as $\hat{N}_t$. Finally, the linear series is added with the non-linear series to make the hybrid model. The mathematical model is given by,

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t$$

$\hat{Y}_t$ is the required hybrid series of the initial series Y_t .

Brock Dechert Scheinkman test (BDS test)

Brock Dechert Scheinkman (BDS) test is a statistical test is used to detect non-linearity in any time series, proposed in the year 1996 (9). Null hypothesis of this test assumes that the given series is identically and independently distributed suggesting absence of non-linear component. Contrarily, the alternative hypothesis assumes that

given series is non-linearly dependent. This test is applied using 'tseries' package of R studio software. If the p -value result of this test is bigger than the specified alpha level, null hypothesis is accepted. This indicates that the series is identically and independently distributed. On the other hand, if p -value of the test is smaller than the level of significance, null hypothesis is rejected. This suggests that series is non-linearly dependent.

Results and Discussions

Trend analysis

Trend analysis was conducted on two series: pineapple area and pineapple production. To identify the best trendline model, each series was fit with four different mathematical models namely linear, quadratic, logarithmic and exponential. These four curves were selected for evaluation as they represent different types of deterministic growth relationships commonly observed in economic and agricultural time series (10–13). For instance, if a variable increases approximately at a constant rate it may follow a linear trend, whereas if it increases at a moderate but accelerating rate it might follow a quadratic trend (10, 12). Further, a variable following an exponential path generally reflects compound or multiplicative growth, which is typically faster than linear and quadratic growth patterns and frequently observed in agricultural production expansion (12). Conversely, if a series is well explained by a logarithmic model, it indicates an increasing trend with a declining growth rate over time, reflecting diminishing marginal expansion (12).

Model performance was evaluated using the coefficient of determination (R^2) and RMSE, which are widely accepted measures for assessing goodness-of-fit and predictive accuracy in time series modelling (14). The trendline model exhibiting the lowest RMSE and highest R^2 among the competing models was identified as the best trendline model for that series. In the present study, logarithmic and exponential models were found to be the best functional forms for pineapple area and production respectively. Similar modelling strategies and functional form comparisons have been widely adopted in agricultural production studies (14–16). The best trendline model of each variable is presented in Table 1.

Auto regressive integrated moving average (ARIMA)

Auto regressive integrated moving average (ARIMA) methodology was employed to model pineapple area and production series were developed for pineapple area and production series. Both series were modelled using the ARIMA framework following standard steps: identification, estimation, diagnostic checking and forecasting. During the identification stage, stationarity of the series was examined using the Augmented Dickey-Fuller (ADF) test at the 5 % significance level (6). Visual inspection of the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots was also performed (Fig. 2). The original series were found to be non-stationary; however, stationarity was achieved after second-order differencing. Therefore, the differencing order was fixed as $d = 2$. The ADF test results are presented in Table 2.

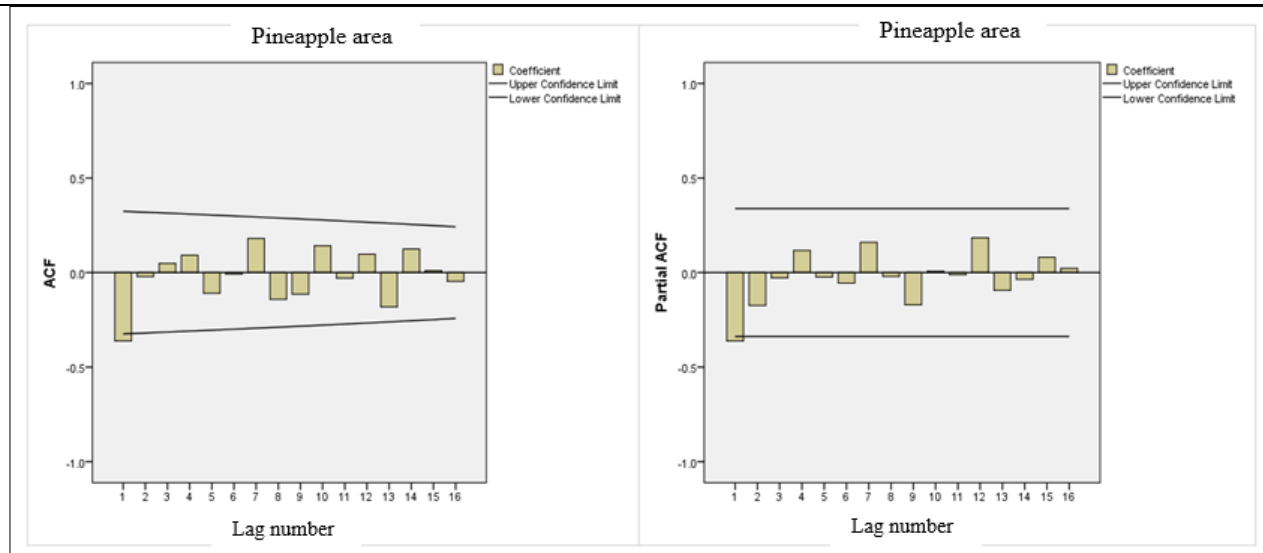
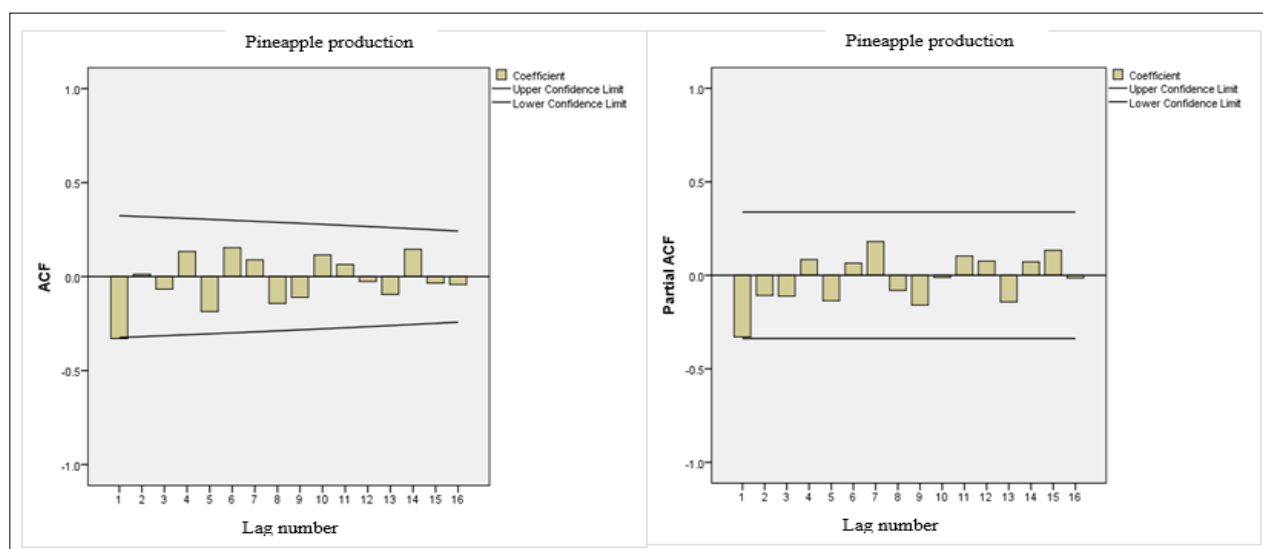
The autoregressive order (p) and moving average order (q) were identified from the PACF-ACF plots respectively, based on the number of spikes observed in the plot (Fig. 2 and 3) was selected as the tentative model ARIMA (0, 2, 1) for both pineapple area and pineapple production series.

Table 1. Description of best trendline model of the two series

Pineapple area	$y = 2.1884\ln(x) + 7.9563$	0.67	1.350
Pineapple production	$y = 137.63e^{0.0233x}$	0.7555	27.858

Table 2. Augmented Dickey-Fuller (ADF) test results of pineapple area and pineapple production

Variables	Series	Dicky Fuller statistic	lag	p-value
Pineapple area	1 st differenced series	-3.0393	3	0.17
	2 nd differenced series	-3.7712	3	0.03
Pineapple production	1 st differenced series	-2.9784	3	0.19
	2 nd differenced series	-4.5265	3	0.01

**Fig. 2.** Autocorrelation function (ACF) and partial autocorrelation function (PACF) plot of 2nd differenced series of area of pineapple.**Fig. 3.** Autocorrelation function (ACF) and partial autocorrelation function (PACF) plot of 2nd differenced series of production of pineapple.

Model parameters were estimated using the maximum likelihood estimation approach available in the 'forecast' package of R, to validate model selection, the proposed ARIMA (0, 2, 1) model was compared with competing ARIMA specifications using goodness of fit and forecast error measures; including akaike information criterion (AIC), RMSE and MAPE (10). The model with lowest error values selected as a best model for diagnostic checking. For residual analysis was done through Ljung box test (6) and residual ACF-PACF

plot. The results of Ljung box test of all the models were given in Tables 3 and 4. The Ljung-Box test yielded *p* values greater than 0.05, indicating no significant autocorrelation in the residuals. It depicted that the residual series of the two models is white noise, thus the models were considered as good fit. Hence, ARIMA (0, 2, 1) was considered an adequate linear model for forecasting pineapple area and production. Five-year forecasts were subsequently generated using the finalised ARIMA models.

Table 3. Ljung box test result of average model-artificial neural network (ARIMA) models of pineapple area and pineapple production

Variables	Best models	Ljung box	lag	p-value
Pineapple area	ARIMA (0,2,1)	11.561	17	0.83
Pineapple production	ARIMA (0,2,1)	12.967	17	0.74

Table 4. Description of best average model-artificial neural network (ARIMA) models of pineapple area and pineapple production

Variables	Best models	AR coefficient	p-value	MA coefficient	p-value
Pineapple area	ARIMA (0,2,1)	Constant = -0.299	0.4	1.000	0.9
Pineapple production	ARIMA (0,2,1)	-	-	0.999	0.9

Following ARIMA modelling, the residual series were further examined for the presence of non-linear dependence using the BDS test (9) implemented through the t-series package in R (12). The test was conducted for multiple embedding dimensions and epsilon values at the 5 % significance level. The resulting *p* value was less than 0.05 leading to rejection of the null hypothesis of independent and identically distributed residuals. This indicated the existence of non-linear structures not captured by the linear ARIMA framework (Fig. 4, 5).

The presence of significant nonlinear dependence in the residuals justified the application of non-linear models in the subsequent stage of hybrid modelling. Accordingly, non-linear techniques were employed to model the residual component remaining after ARIMA estimation, thereby improving forecasting performance by capturing both linear and non-linear patterns in the series.

Artificial neural network (ANN)

For non-linear modelling initially ANN model was applied. Multilayer feed-forward neural networks are capable of approximating complex non-linear relationships (15). The dataset was divided into training (80 %) and testing (20 %) subsets to enable model validation and reduce overfitting (16). Different ANN models were developed considering different combinations of hidden nodes and input nodes. Total 30 models were developed considering 1–6 numbers of input node and 2–10 number of hidden nodes at an increment of 2 and their performances were recorded through RMSE and MAPE values. The model showcasing lowest error both in training as well as testing set is considered as the best ANN model for that series. The modelling strategy adopted in this study is consistent with earlier applications of ANN and hybrid ARIMA-ANN models in economic and agricultural time series forecasting, where experimentation with network structure and error-based comparison is recommended for identifying robust predictive models (16, 17). Best ANN models for the two series were given in Tables 5–7.

Non-linear support vector regression (NLSVR)

Non-linear support vector regression (NLSVR) modelling was performed using the 'TSSVM' package in R Studio (18). The theoretical framework of support vector regression (SVR) was originally developed by Vapnik and Vladimir N under statistical learning theory, which provides strong generalisation capability through structural risk minimisation (19, 20). Adaptive recursive support vector machine (ARSVM) function was used to build the NLSVR model over pineapple area and production series (18). The model was developed over the training data set and forecasts were

generated for testing set. The automated tuning mechanism within the TSSVM framework optimises kernel and regularisation parameters using the training data, ensuring robust non-linear pattern extraction (20). TSSVM is an automated package that used the training data to develop the model. Model performance was evaluated by comparing forecasts with observed values using error metrics. At last forecasts were generated for each variable using their respective best NLSVR model.

Hybrid model comparison and interpretation

Hybrid models were developed to combine linear and non-linear components of the time series. The rationale behind hybridisation lies in the assumption that real-world time series often contain both linear and non-linear components, which cannot be adequately captured by a single model (21). In the hybrid ARIMA-ANN modelling framework, ARIMA was first applied to model the linear component of the original time series. The residuals obtained from the fitted ARIMA model were subsequently used as input for ANN modelling to capture the remaining non-linear structure. The forecasts generated from the ANN residual model were then added to the ARIMA forecasts to obtain the final hybrid forecasts (15, 19). A similar additive hybridisation procedure was followed for the ARIMA-NLSVR model.

Finally, values of hybrid series were compared with original values through RMSE measure. Performance results for both series are presented in Tables 8 and 9. From the results ARIMA-ANN performed better than other selected models like ARIMA-NLSVR, ANN, NLSVR, ARIMA in case of pineapple area and in pineapple production ANN model performed good as compared with ARIMA-ANN, ARIMA-NLSVR, NLSVR and ARIMA. Previous studies have reported strong forecasting performance of hybrid ARIMA-ANN models in agricultural and economic time series forecasting (14, 21). However, in the present study, ANN provided better forecasts for pineapple production, whereas the hybrid ARIMA-ANN model performed better for pineapple area.

Conclusion

Trend analysis revealed that pineapple production in Assam followed an exponential growth pattern, whereas pineapple area exhibited a logarithmic trend. Among the forecasting approaches, the hybrid auto-regressive integrated moving average model-artificial neural network (ARIMA-ANN) model demonstrated superior performance for forecasting pineapple area, while the ANN model produced the most accurate forecasts for pineapple production

Table 5. Brock Dechert Scheinkman (BDS) test result of pineapple area

Dimension	Epsilon	Statistic	<i>p</i> -value
2	0.507	-5.008	0.00
2	1.521	2.421	0.02
3	0.507	-4.955	0.00
3	1.521	8.578	0.00

Table 6. Brock Dechert Scheinkman (BDS) test result of pineapple production

Dimension	Epsilon	Statistic	<i>p</i> -value
2	1.262	33.285	0.00
2	3.785	8.650	0.00
3	1.262	-4.165	0.00
3	3.785	4.638	0.00

Table 7. Best artificial neural network (ANN) model description of pineapple area and pineapple production series

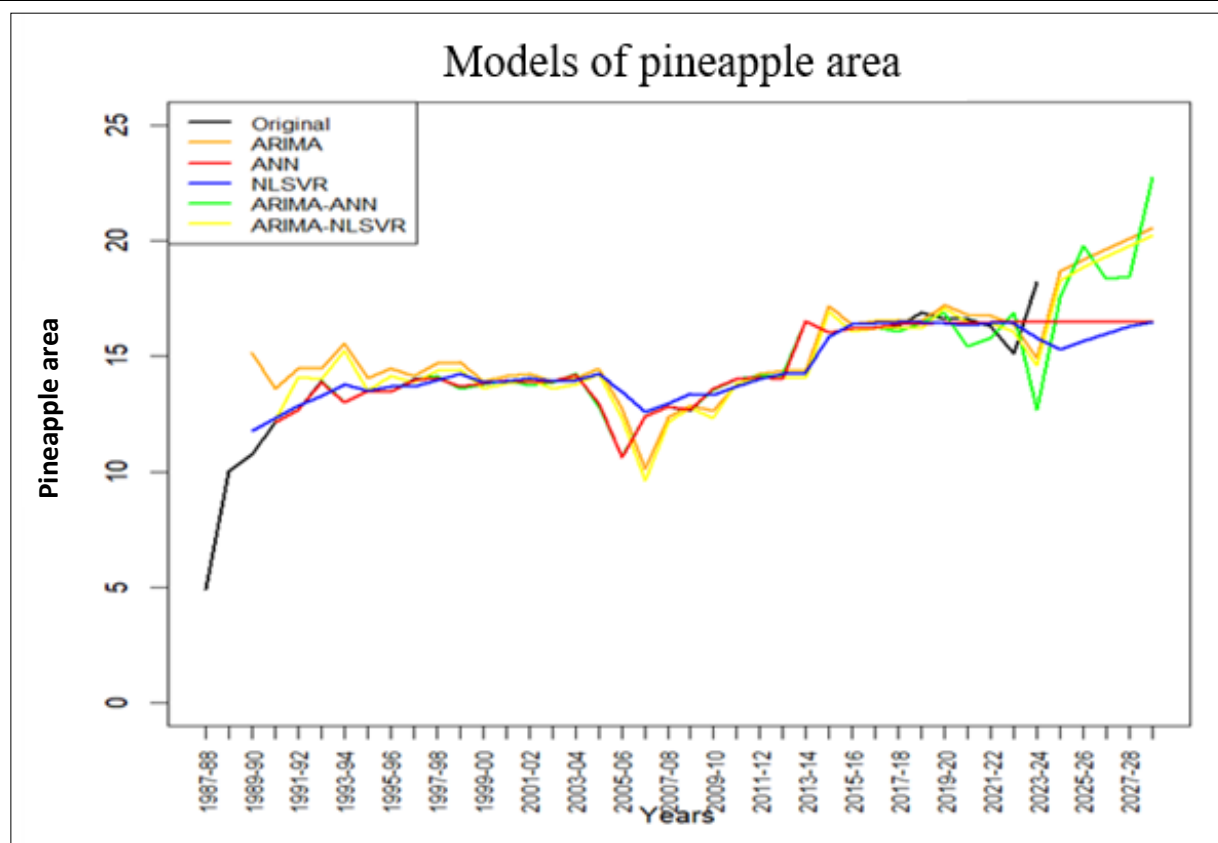
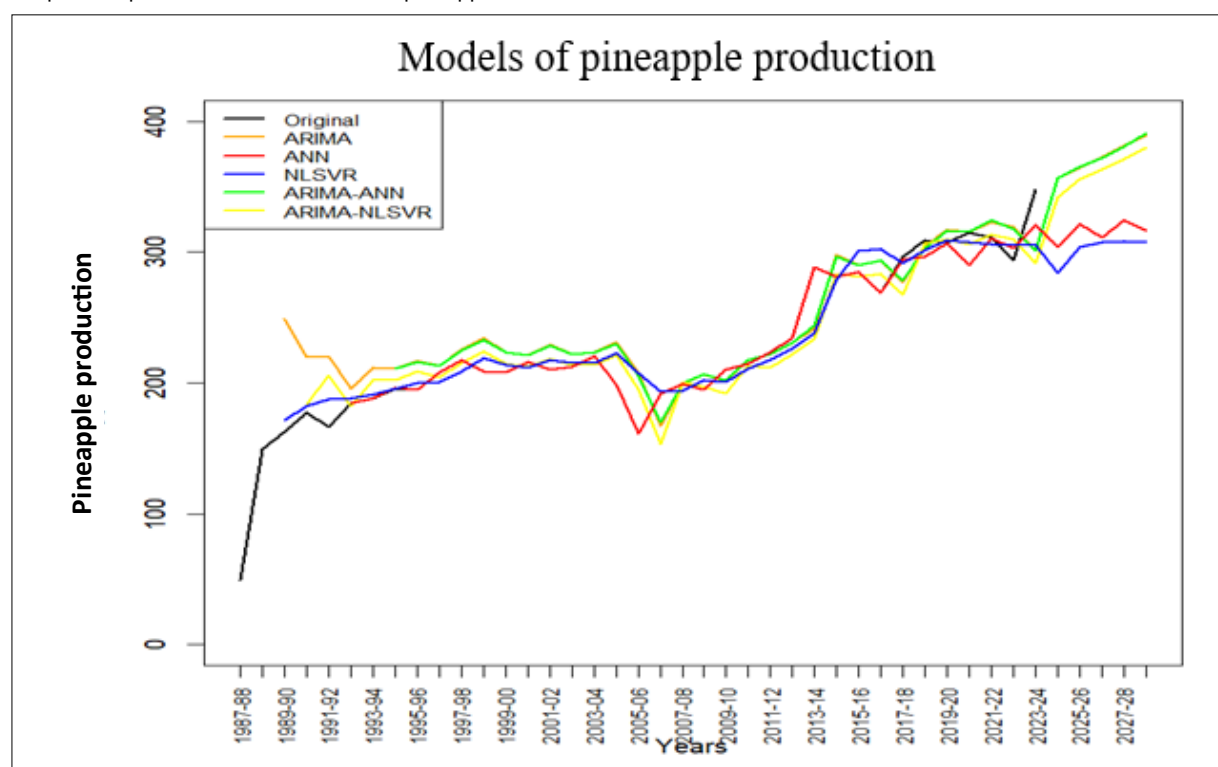
Variables	Best ANN models	No. of weights	No. of input node	No. of hidden node
Pineapple area	3:8s:1l	61	3	8
Pineapple production	5:10s:1l	71	5	10

Table 8. Best non-linear support vector regression (NLSVR) description of all pineapple area and production series

Variables	Measures	ARIMA	ANN	NLSVR	ARIMA-ANN	ARIMA-NLSVR
Pineapple area	Training	1.423	0.061	0.829	0.008	1.084
	Testing	1.294	0.807	1.005	2.112	1.361
Pineapple production	Training	28.157	0.018	16.833	19.226	18.683
	Testing	22.339	15.270	17.245	22.073	24.963

Table 9. Forecast results generated by the best forecasting models of the two series

Variables	Kernel function	c	γ	ϵ
Pineapple area	RBF	1	0.5	0.1
Pineapple production	RBF	1	0.5	0.1

**Fig. 4.** Graphical representation of all models of pineapple area.**Fig. 5.** Graphical representation of all models of pineapple production.

based on root mean square error (RMSE) and mean absolute percentage error (MAPE) criteria. The findings indicate that incorporation of non-linear modelling approaches can substantially improve forecasting accuracy for agricultural time series data.

The forecast results suggested a future increase in both pineapple area and production in Assam during the forecast period. These findings may assist agricultural planners, policymakers and stakeholders in developing efficient production planning, resource allocation strategies and market forecasting frameworks for pineapple cultivation in the state.

Future research may focus on incorporating climatic variables, economic indicators and advanced deep learning approaches to further improve forecasting accuracy and understand the factors influencing pineapple production dynamics in Assam.

Authors' contributions

AS carried out the research work as part of the M.Sc. (Agri.) thesis, including data collection, analysis and manuscript preparation. BN being the research guide, conceptualised the study, provided guidance in data analysis, interpretation of results and manuscript refinement. HG and PB as members of the advisory committee, contributed through technical suggestions, critical review and overall supervision of the research work. UB, MP, TS and TN reviewed the manuscript critically for intellectual content. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interest to declare.

Ethical issues: None

References

1. Ministry of Agriculture and Farmers Welfare. Horticultural statistics at a glance 2020-21. New Delhi: Government of India; 2021.
2. Rathod S, Mishra GC, Singh KN. Hybrid time series models for forecasting banana production in Karnataka State. *J Indian Soc Agric Stat.* 2017;71(3):193–200.
3. Alam W, Sinha K, Kumar RR, Ray M, Rathod S, Singh KN, et al. Hybrid linear time series approach for long-term forecasting of crop yield. *Indian J Agric Sci.* 2018;88(8):1275–9. <https://doi.org/10.56093/ijas.v88i8.82573>
4. Sujatha P. Hybrid statistical models for forecasting yield of mango and banana in Tamil Nadu, India. *Asian J Agric Ext Econ Sociol.* 2021;39(11):168–74. <https://doi.org/10.9734/ajaees/2021/v39i1130738>
5. Directorate of Economics and Statistics. Statistical handbook of Assam 2023–24. Guwahati: Government of Assam; 2024.
6. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. 5th ed. Hoboken (NJ): John Wiley & Sons; 2015.
7. Vapnik VG. Support vector method for function approximation, regression estimation and signal processing. In: Mozer M, Jordan M, Petsche T, editors. *Adv Neural Inf Process Syst.* Cambridge (MA): MIT Press; 1997. p. 281–7.

8. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing.* 2003;50:159–75. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)
9. Brock WA, Dechert WD, Scheinkman JA, LeBaron B. A test for independence based on the correlation dimension. *Econom Rev.* 1996;15(3):197–235. <https://doi.org/10.1080/07474939608800353>
10. Chatfield C. The analysis of time series: an introduction. 6th ed. Boca Raton (FL): CRC Press; 2004.
11. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. 3rd ed. Melbourne (AU): OTexts; 2018. <https://doi.org/10.32614/CRAN.package.fpp3>
12. Gujarati DN, Porter DC. Basic econometrics. 5th ed. New York (US): McGraw-Hill Education; 2009.
13. Granger CWJ, Newbold P. Forecasting economic time series. 2nd ed. San Diego (US): Academic Press; 1986.
14. Neog B, Gogoi B, Barua S, Mahanta, Sarma K. Development of hybrid time series models for forecasting autumn rice production in Assam. *Pharma Innov J.* 2023;12(2):3376–84.
15. Zhang GP, Patuwo BE, Hu MY. Forecasting with artificial neural networks: the state of the art. *Int J Forecast.* 1998;14(1):35–62. [https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7)
16. Khashei M, Bijari M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Appl Soft Comput.* 2011;11(2):2664–75. <https://doi.org/10.1016/j.asoc.2010.10.015>
17. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Netw.* 1989;2(5):359–66. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
18. Varghese E, Gills R, Anuja AR, Jayasankar J, editors. Manual on advanced analytical tools for social science research. Vol. 2. CMFRI Training Manual Series No. 29/2023. Cochin: ICAR-CMFRI; 2023. p. 324.
19. Pai PF, Lin CS. A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega.* 2005;33(6):497–505. <https://doi.org/10.1016/j.omega.2004.07.024>
20. Vapnik VN. Statistical learning theory. New York (US): Wiley; 1998.
21. Neog B, Gogoi B, Patowary AN. Development of hybrid time series models for forecasting autumn rice using ARIMAX-ANN and ARIMAX-SVM. *Ann For Res.* 2022;65(1):9119–33.

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonpublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonpublishing.com/journals/index.php/PST/indexing_abstracting

Copyright: © The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited (<https://creativecommons.org/licenses/by/4.0/>)

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.