www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(4): 1947-1956

www.thepharmajournal.com
Received: 19-02-2022
Accepted: 21-03-2022

Sudhir Paswan
Ph.D. Scholar, Department of Mathematics and Statistics, SHUATS, Prayagraj, Uttar Pradesh, India

Anupriya Paul
Department of Mathematics and Statistics, SHUATS, Prayagraj, Uttar Pradesh, India

Ajit Paul
Department of Mathematics and Statistics, SHUATS, Prayagraj, Uttar Pradesh, India

Ashish S Noel
Department of Agricultural Economics, SHUATS, Prayagraj, Uttar Pradesh, India

Corresponding Author
Sudhir Paswan
Ph.D. Scholar, Department of Mathematics and Statistics, SHUATS, Prayagraj, Uttar Pradesh, India

Time series prediction for sugarcane production in Bihar using ARIMA & ANN model

Sudhir Paswan, Anupriya Paul, Ajit Paul and Ashish S Noel

Abstract
Agriculture is important to the Indian economy and employment; India is the world's second-largest producer of sugarcane, which is a key cash crop in Bihar. The main objectives of this paper are to look into the stability and long-term viability of sugarcane production in the state of Bihar. For the study period 1939-40 to 2019-20, secondary data on sugarcane production was obtained from the Directorate of Economics and Statistics, Bihar and Sugarcane Industries Department, Bihar. Sugarcane production was predicted using the Box-Jenkins ARIMA model and artificial neural network (ANN) approach. To create the model and estimate the forecasting behaviour, the Box-Jenkins ARIMA Model was employed, as well as the ANN method. The ARIMA (1,1,0) model is the most suitable for forecasting based on the minimal value of AIC (649.781) and BIC (654.545). The model ARIMA (1,1,0) was presented for forecasting sugarcane production from 2020 to 2025, which indicated a significant increase from 126.03 lakh to 131.67 lakh ton. These forecast values are useful for collecting information and planning resources for the government, sugar mills, researchers, and businesspeople, as well as farmers making key decisions about sugarcane crop production in Bihar.

Keywords: ARIMA, ANN, model, error, production, forecast

1. Introduction
Sugarcane is one of the world's most valuable crops due to its strategic location and diverse variety of uses in daily life as well as industrial applications targeted at nutritional and economic sustainability. Around 60% of global sugar output comes from sugar cane with the remaining 40% coming from sugar beet (Onwueme I. C. and Sinha, T.D., 1991) [29]. It's a tropical crop that takes between 8 and 12 months to mature. The sugar content of matured cane reaches its highest when it turns green, yellow, purple, or reddish in colour (Onwueme I. C., 1978) [30]. A person's or a group's socioeconomic status is determined by their social rank or class. It's usually assessed by taking into account criteria like education, wealth and occupation. Sugarcane has had a significant impact on poverty reduction and social development. The price of sugarcane has an impact on socioeconomic development (Bharati *et al*., 2018) [8]. Sugarcane was cultivated on 45.67 lakh hectares in India in 2019-20 with production of 3557.00 lakh ton and a yield of 77.88 ton per hectare (Annonymous, 2020). Sugarcane is grown all over India, with the most important states being Uttar Pradesh, Maharashtra, Karnataka, Tamil Nadu, Bihar, Gujarat, Haryana, Punjab, and Andhra Pradesh. Sugarcane has a significant role in India's economy. Cane agriculture employs around 6 million farmers and a high population of agricultural labourers. Furthermore, India's largest agro-processing industry, the sugar industry, employs nearly half a million skilled and semi-skilled workers, the most of whom are from rural areas. White sugar production accounts for roughly 60% of total cane production in the country on average. For gur and khandsari manufacture about 15-20% sugarcane is used. In Bihar during 2019-20 sugarcane was cultivated in an area 3.04 lakh hectares and production was 182.84 lakh tones with productivity of 60.15 tonnes per hectares (Annonymous, 2020) [3]. In comparison to some of the country's major sugarcane producing states, the yield per hectare is poor. This results in low overall yield and a scarcity of sugarcane for sugar mills. Efforts are being undertaken to tackle this problem by creating high yielding, early maturity and high sucrose content sugarcane types that are both disease and pest resistant.
It has evolved into a complicated scientific study aiming at generating the highest amount of agricultural output with the lowest amount of time, space, and energy in order to meet the demands of an expanding population and economy. Higher sugar production with the available area is required due to exploding population increase and rising per capita sugar consumption.

For future planning and policymaking, accurate forecasting of such key commercial crops is required. These inaccurate estimates led to a prohibition on sugar exports, resulting in a loss of export business at a time when sugar prices were high on the worldwide market. Accurate early warning of crop failures can help mitigate the negative implications of public policy, such as price rises and agricultural distress. An attempt has been made to construct an appropriate forecasting model because sugarcane productivity forecasting could help in estimating production and making decisions about export and import policies, distribution, price policies, storage and marketing strategies (Priya and Suresh, 2009) [31]. The ARIMA model is a forecasting extrapolation approach that like any other requires simply historical time series data on the variables to be forecasted. This is one of the most advanced extrapolation strategies since it encompasses the future of all such approaches and does not require the investigator to prioritise the beginning values of any variables or the values of the many parameters. It can handle any data pattern with simplicity. The advantage of ARIMA modelling over other univariate time series models is that it delivers the fewest mean squared forecast error variations in addition to revealing the intrinsic behaviour (producing process) in the time series variables (Rankja *et al*., 2017) [32]. In an economic system, accurate forecasting of the production of such major commercial crops is critical. Crop production and prices are inextricably linked. Unexpected decreases in production limit farmers' marketable surplus and revenue, resulting in price increases. A surplus of production can result in price drops, which lowers farmers' profits. Inflation, wages and salaries as well as other economic policies are all affected by the price of a basic item. In the case of commercial crops like sugarcane, the quantity of output has an impact on the raw material costs of user companies as well as their market competitiveness (Suresh and Priya, 2011) [39]. In this work, Auto Regressive Integrated Moving Average (ARIMA) models were used to forecast sugarcane area, production, and productivity in Bihar. Artificial neural networks (ANNs) have prompted a lot of interest in agricultural economics and research recently. A collection of interconnected basic computing units makes up an artificial neural network (ANN). ANNs have been applied in a range of situations where traditional statistical approaches have been used. Artificial Neural Networks (ANN) have the ability to discover complicated nonlinear correlations between dependent and independent variables implicitly, therefore they can be utilised to more accurately predict future demand. Because of their ability to accommodate non-linear data and capture subtle functional relationships among empirical data, Artificial Neural Network (ANN) algorithms have been found to be useful techniques for demand forecasting, even when the underlying relationships are unknown or difficult to describe. These are currently preferred tool in predicting future crop yield and production. Statisticians and statistic users often employ traditional statistical methodologies to handle these problems, such as discriminant analysis, logistic regression, Bayes analysis, multiple regression, and time-series models. As a result, it's past time for ANN to be recognised as a powerful data analysis tool. The author provides a great overview of many elements of ANN (Cheng and Titterington, 1994) [13]. Neural networks have the ability to change their settings while dealing with non-stationary and dynamic input (weights), hence this model is used in this paper to compare the strength to deal with non-stationary data as well as ARIMA model. The capacity to forecast future sugarcane production based on historical data is a major tool for agricultural decision-making. Despite rising urbanisation around the world, agriculture remains the primary source of income for a huge percentage of the people. Although technological developments have resulted in more accurate weather predictions and increased yields, much work remains to be done to provide farmers with a bankable strategy that will prepare them for future yields. ANN models have been created for sugarcane production data in India during the last 81 years in this study (1939 - 2019). Sugarcane production is influenced by a multitude of factors, including weather and climate, agricultural policy, and socioeconomic issues, much like any other crop (Houghton *et al*., 1996) [22]. As a result, forecasting sugarcane crop yield becomes a challenging task and here is where ANN's reliability comes into play as it can effectively handle multivariate non-linear non parametric statistical approaches (Cheng and Titterington,1994) [13], (White, 1989) [40]. (Ripley, 1993) [34].

1.1. Forecasting using Time Series Techniques: Theoretical Concepts

A variable to be forecasted (in this case, production) is modeled as a function of time in a time series analysis. i.e.

$$Y_t = f(t) + \varepsilon_t \qquad \text{...(1)}$$

Where, Y_t signifies production for the time t in year, f (t) denotes a function of time t, and ε_t denotes production error (i.e., the difference between observed and forecasted production for time t year). Once a functional link between production and time (in other words, a time series model) has been built, production for year t+1 can be forecasted. The first stage in creating this model is determining whether the time series under consideration is stationary or non-stationary.

1.2 Stationary or non-stationary series categorization

A time series is characterised as (i) stationary or (ii) non-stationary depending on whether it is stationary or non-stationary. If the statistical features of the series (mean, variance, and autocorrelation) are time-independent, it is regarded as stationary; otherwise, it is classified as non-stationary. Before choosing one of the time series groupings approaches that can be used to represent a series, it must first be categorised. When determining whether a series is stationary or non-stationary, the Augmented Dickey-Fuller test is frequently utilised (Gujarati, 1995) [21]. Create a regression equation with two variables: (i) Y_t as the dependent variable and (ii) Y_{t-1} as the explanatory variable to run this test. The series is next checked to see if the derived regression coefficient can be statistically considered as 1, suggesting that it is stationary. To achieve this goal, the absolute Dickey-Fuller t statistic is compared to the absolute critical t statistic (i.e., the estimated regression coefficient divided by its standard error). It is possible to convert a non-stationary series to a stationary series. Such a transformation will open the door to the use of techniques that are not normally applied to non-stationary series, allowing for the use of a greater number of approaches to model the series. The more models constructed to describe a series, the more likely it is that the appropriate technique for forecasting the series will be selected. The differencing method is a straightforward and extensively used technique for transforming a non-stationary series to a stationary one. When this method is applied to the current situation, a non-stationary series is transformed to another series by subtracting Y_{t-1} from Y_t (when level of differencing is 1), Y_{t-2} from Y_t (when level of differencing is 2), … and Y_{t-n} from Y_t (when level of differencing is n). To begin the process of the transformation, the level of differencing is chosen as 1 and the transformed series is tested for stationarity. If the series is determined to be non-stationary, the process of transformation is repeated at the next higher level of differencing. In most cases, a third degree of differencing is sufficient to modify a series. In Tamil Nadu, the Box-Jenkins ARIMA model was applied to forecast sugarcane production (Devaki and Mohideen, 2021) [14]. Sugarcane forecasting in Pakistan utilizing an appropriate measure such as an ARIMA model proved to be useful and appropriate for policymaking (Muhammad *et al*., 1992) [28]. Farmers, governments, and agribusiness corporations could benefit from agricultural output and price forecasts (Allen, 1994) [2]. Researchers investigated the range of forecasting approaches in a case study of wheat output forecasting in Canada and discovered that quadratic smoothing was the most efficient (Boken, 2000) [9]. To construct the model and verify the prediction behaviour, researchers used the Box-Jenkins ARIMA model, artificial neural network (ANN) technology and a hybrid technique. The acreage, production, and yield of wheat crops in Haryana are all increasing, according to the results of the experiment (Devi, *et al*., 2021) [15]. They forecasted sugarcane production in Pakistan using the ARIMA model and they believe the forecast values will be quite similar to the actual values (Yaseen *et al*., 2005) [41]. The relevance of sugarcane in the KPK region of Pakistan was examined and it was discovered that the overall elasticity were more than one, showing that the agricultural sector was producing at a higher rate of return to scale, reflecting that the sector's input allocation isn't appropriate (Azam and Khan, 2010) [7]. In a study on pre-harvest sugarcane yield forecasting using climatic factors in India, they developed a prediction model that used weather variables as a regressor and found that the forecast model could explain 87 percent of the variation in sugarcane output before two months harvest (Krishna and Priya, 2011) [39]. In Pakistan, they use trend techniques to forecast maize acreage and yield. They discovered that the quadratic trend model was the most accurate forecasting approach, that the forecasted values are extremely near to real values and that the trend is positive (Ayesha and Nusrat, 2013) [6]. According to the conclusions of this study, climate and other environmental changes have become a major threat to agricultural economies in the developing world, particularly on the African continent implying that the ARMA model is chosen over other time series models (Askar and James, 2014) [5]. When it comes to forecasting the yields of Pakistan's primary crops, sugarcane and cotton, the ARIMA model was declared to be suitable for forecasting sugarcane yield and the ARIMA model for forecasting cotton yield after multiple diagnostic tests on the fitted model were confirmed (Sajid *et al*., 2015) [36]. The ARIMA (univariate linear time series model) was proposed by Box-Jenkins in 1970. In many locations of Mexico, the hybrid ARIMA–ANN model has been utilised to forecast wind speed (Cadenas and Rivera, 2010) [12]. In various studies, hybrid ARIMA was used for forecasting (Dia-Robles *et al*., 2008 and Faruk, 2010) [16, 18]. The ARIMA model and the ARIMA-WNN hybrid model were employed in wheat crop forecasting [Saeed *et al*. (2000) [35], Ray *et al*. (2016) [23], Mishra *et al*. (2015) [27] and Mishra *et al*. (2021) [21]]. After making stationary at first difference, non-stationary data was used for the ARIMA methodology. Brown's smoothing and Holt's exponential smoothing model produced very accurate results that were useful to policymakers (Akin and Eyduran, 2017) [1]. They discovered that the ARIMA model's results were satisfactory when they used secondary data and the ARIMA forecasting technique to estimate the main food crop output in Khyber Pakhtunkhwa (Shah *et al*., 2017) [37].

2. Materials and Methodology

Secondary information of Bihar sugarcane production collected in Bihar from Directorate of Economics and Statistics, Bihar and Sugarcane Industries Department, Bihar for the study period 1939–40 to 2019-20 have been employed.

2.1. Box-Jenkins modeling procedure

In ARIMA, historical observations and random mistakes (errors) are combined to form a linear relationship that can be used to estimate future variables. Box-Jenkins for forecasting the future value of a time series, ARIMA modelling employs a three-step iterative technique: model identification, parameter estimation and residual diagnostics testing (Box, *et al*., 1967 and Box, *et al*., 2009) [10-11].

Autoregressive integrated moving average model (ARIMA (p, d, q).

$$\varphi_{p(B)}\Delta^d h_t = c + \theta_q(B)\ g_t \qquad \dots (2)$$

Where, h_t is variable under forecasting at time t, B is lag operator, g is error term (Y-Ŷ in which Ŷ is the estimated value of Y), $\varphi_{p(B)}$ is non-seasonal AR i.e. the autoregressive operator, represented as a polynomial in the back shift operator, $(1 - B)^d$ is non-seasonal difference, $\theta q(B)$ is non-seasonal MA i.e. the moving average operator, represented as a polynomial in the backshift operator, φ′ s and θ′ s are the parameters to be estimated.

Table 1: Summary statistics of Area, Production and Yield of Sugarcane (1939-40 to 2019-20)

Parameter	Area (Lakh Ha)	Production (Lakh Ton)	Yield (Lakh/Ton)
Mean	1.53	58.11	37.29
Standard Deviation	0.47	30.02	9.37
Kurtosis	2.59	3.09	-0.70
Skewness	1.65	1.99	0.11
Range	2.08	123.00	42.64
Minimum	0.94 (2000-01)	26.00 (1948-49)	16.56 (1948-49)
Maximum	3.02 (2014-15)	149.00 (2014-15)	59.20 (2017-18)

*Value in parenthesis is Year

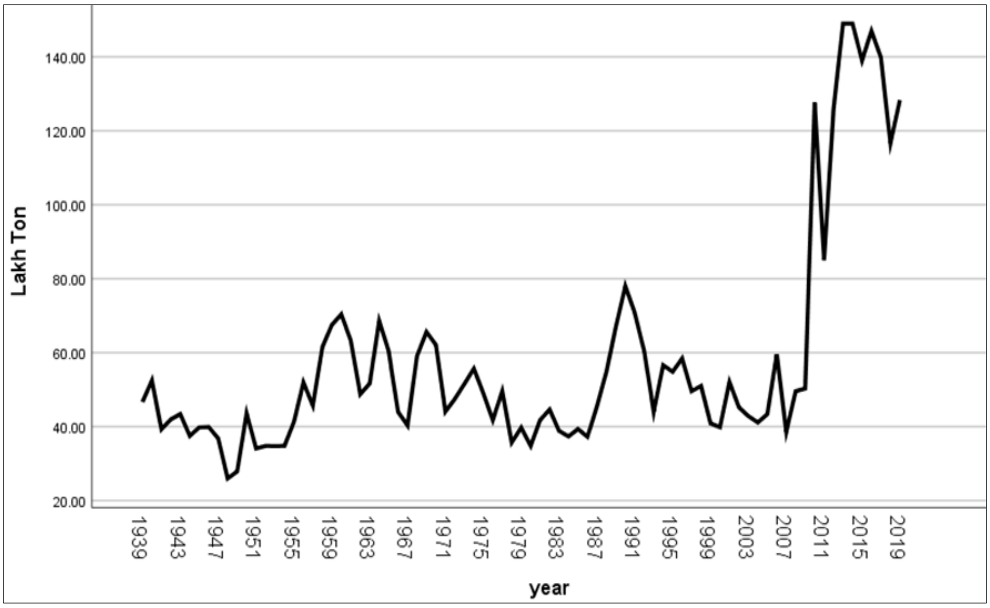

Fig 1: Sugarcane Production of study period

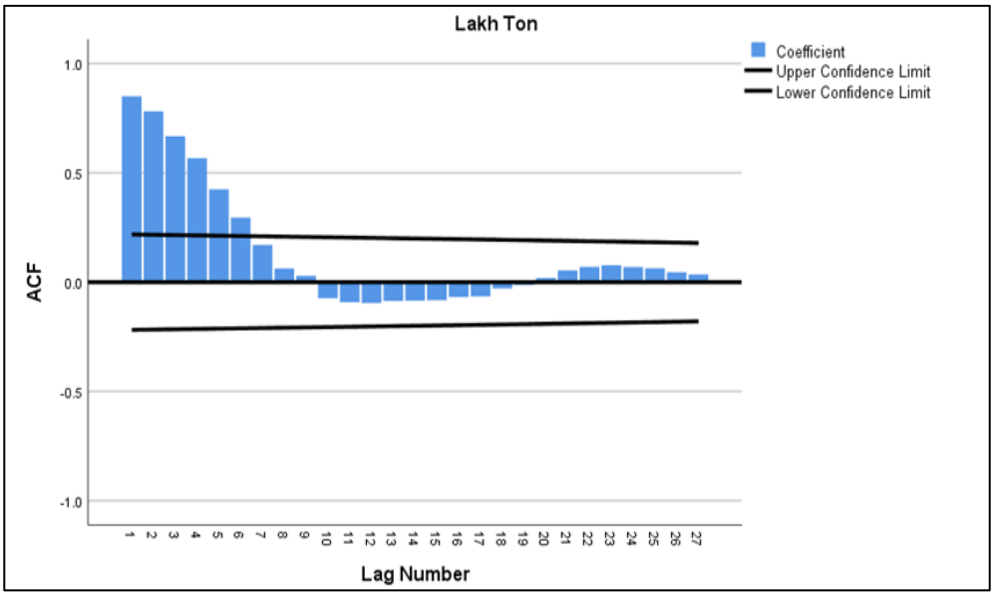

Fig 2: Sugarcane Production ACF Plot

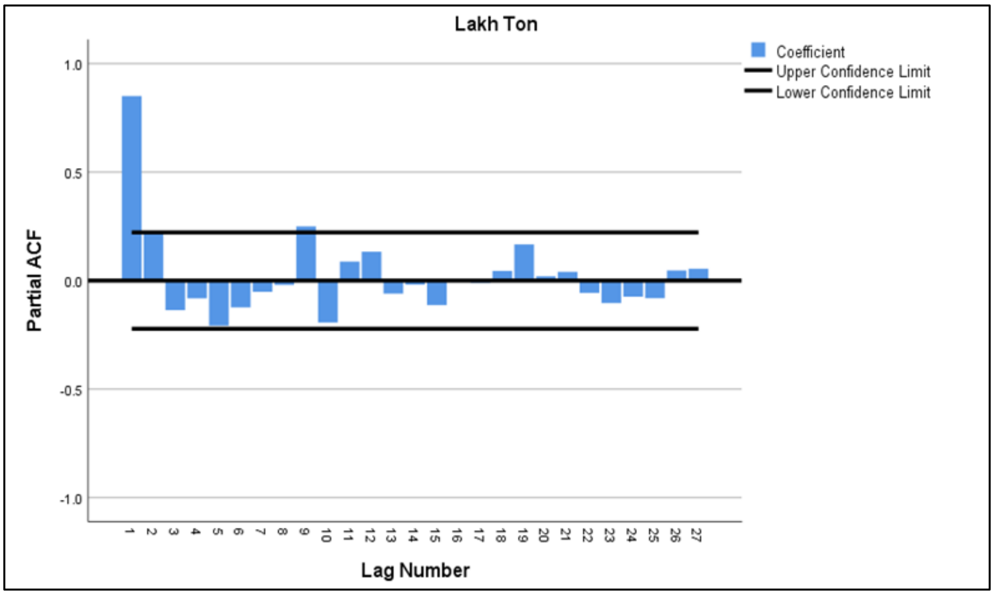

Fig 3: Plot of Sugarcane Production (PACF)

The unit root test series graphs of Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), and Augmented Dickey-Fuller (Dickey and Fuller, 1979) [17] of Sugarcane crop production showed a zig-zag increasing trend to reach the stationary of the series graphs (Figs.1, 2 and 3). As a result, the observed series was non-stationary. Graphs and the Augmented Dickey-Fuller (ADF) unit root Test were used to determine if the data was stable. Based on the graphs of the Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) and Augmented Dickey-Fuller (ADF) of the unit root test, first order Integration (Difference) was proposed to make the above series Stationary.

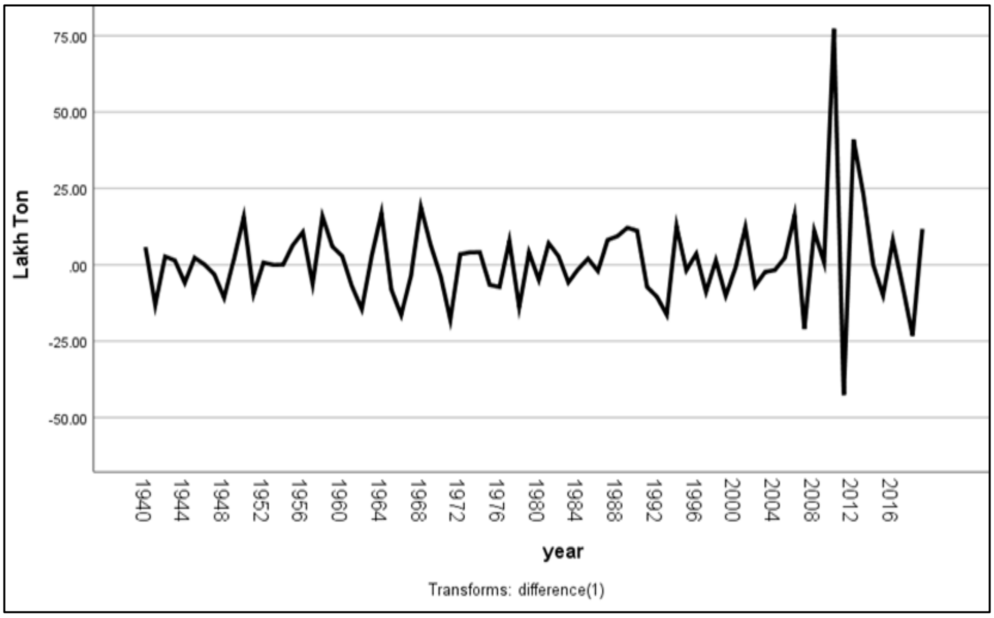

Fig 4: Time Graph of the First Difference in Production The data is stationary in the graph (fig.4) of the first difference of the original Sugarcane crop Production.

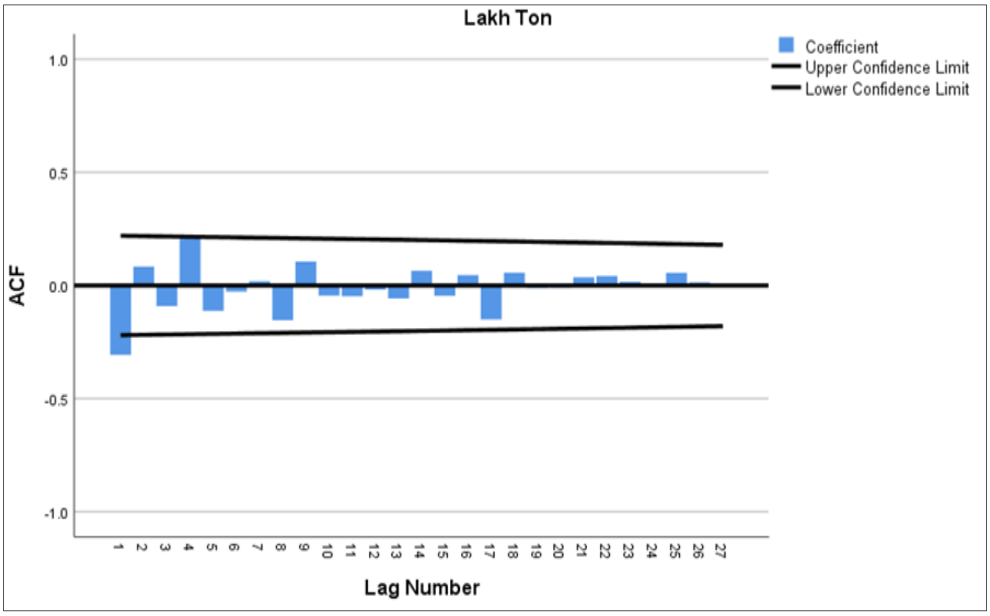

Fig 5: ACF Plot of the First difference in Production

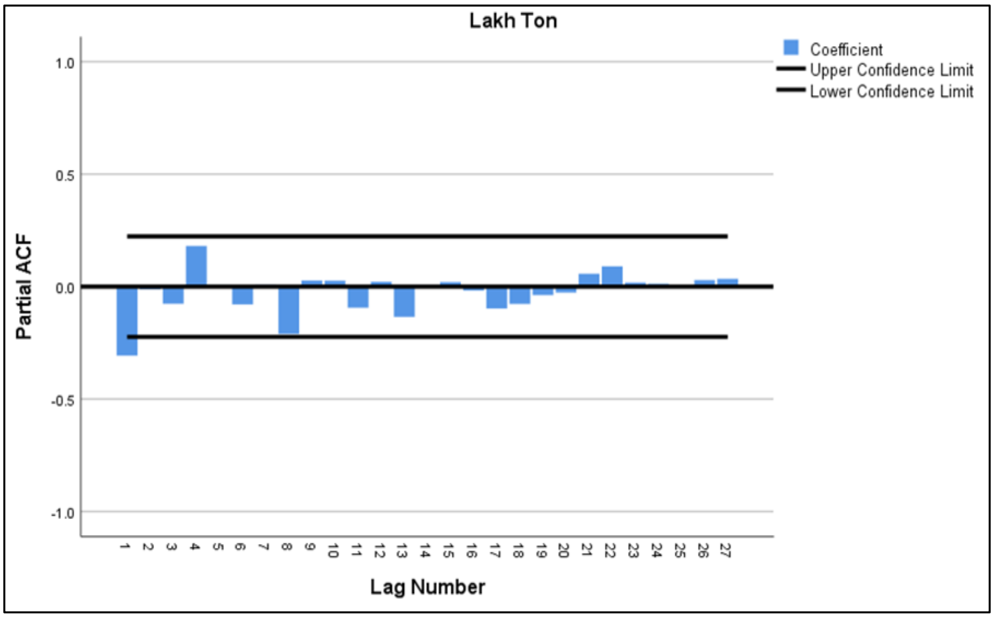

Fig 6: PACF Plot of the First difference Production

The first order moving average (MA) model was adopted because the autocorrelation function graph (Figs. 5 and 6) had only one negative spike and the partial autocorrelation function graph had just one negative spike. A tentative autoregressive integrated moving average (ARIMA) model was used to build the ARIMA model (1,1,0). We begin with this model and perform a diagnostic comparison of all viable fitted models to determine the best forecast ARIMA (p,d,q) model.

Table 2: Examine for stationarity

Level Series		First Differencing	
D F	P- Value	D F	P- Value
-1.758360	0.7155	-12.10817	0.0001

2.2. Brief descriptions of ANN

The human brain is made up of ANN, also known as neural networks, which are biologically inspired by artificial intelligence AI models. Artificial neural networks connect nerve cells in numerous layers of networks, analogous to how neurons in the human brain connect to each other (Remus, W., and O'Connor, M. 2001) [33].

Artificial Neural Networks can be applied in three different ways:

(a) As biological nerve system and intelligence models.

(b) In hardware as real-time adaptive signal processors or controllers for applications such as robots.

(c) In terms of data analytic approaches

Nodes are the names given to the neurons. Without a lot of data or knowledge, ANN is a self-adaptive nonlinear nonparametric data-driven statistical method for regulating nonlinearity and approximating complex relationships. It is employed in the forecasting of time series. The functional form of ANN is as follows if the input nodes have lagged values:

$$y_t = h(y_{t-1}, y_{t-2}, y_{t-3}, \ldots\ldots\ldots\ldots, y_{t-1}) + \varepsilon_t$$

or above functional form can be expressed mathematically.

$$y_t = w_{ij} \sum_{j=1}^{q} w_j .\, g\left(w_{oj} + \sum_{i=1}^{p} w_{ij} y_{t-1}\right) + \varepsilon_t$$

The bias of input is defined as y_{t-1} transformed data (input), w_{ij} is weights associated with input nodes, and w_j is weights associated with hidden nodes, where p is the number of input nodes and q is the number of hidden nodes. In their respective linear and non-linear domains, ARIMA and ANN models have both been successful (Makridakis *et al*.1998 and Ghosh *et al*. 2005) [24, 19]. When ARIMA models and ANNs are compared, the results are mixed in terms of forecasting performance strength.

3. Results and Discussions

Over the research period, Table 1 presents descriptive statistics for sugarcane acreage, production and productivity in Bihar. In terms of area, the kurtosis (2.59) and skewness (1.65) values indicate that the distribution is platykurtic, with a highly positive skew in the right thinner tail, which is longer than the left tail, which has a thinner tail than a normal distribution. The distribution is leptokurtic with a highly positive skew, with the right heavier tail being longer than the left, indicating a tail that is heavier and longer than a normal distribution, according to the kurtosis (3.09) and skewness (1.99) values in production. The kurtosis value (-0.70) in yield is negative, the nature is platykurtic with a lighter, thinner tail, and the skewness value (0.11) is approximately symmetric. Production increased by more than five times during the research period, from 26 lakh tonnes to 149 tonnes. The tables clearly show that yield in Bihar increased from 16.56 ton/ha to 59.20 ton/ha in 2019, a threefold increase from the previous year. Table 2 depicts the stationary test. Various AR and MA combinations were investigated in order to determine the best model for sugarcane production, and ARIMA (1,1,0) was selected based on the lowest AIC and BIC values (table 3). The chosen model is summarised in Table 4, and the residuals were examined using the Box-Ljung test because the coefficients were determined to be highly significant. Table 4 shows that the parameter estimations for the ARIMA (1,1,0) model were all significant, indicating that this model was appropriate for estimating sugarcane production. Figure 7 depicts a graphical diagnostic evaluation of the ARIMA (1,1,0) Model based on the residual plot for autocorrelations and partial autocorrelations functions. The 95 percent confidence interval includes all autocorrelations and partial autocorrelations in the plot. As a result, the model's parameters were accurate. Figure 8's normal probability plot shows that the residuals are symmetric and lie along a straight line indicating that the model was correctly defined. Figure 9 displays the residuals histogram and figure 10 shows the residuals graph, which depicts the best fit of ARIMA(1,1,0) for sugarcane production with plots of actual and anticipated sugarcane production using ARIMA and ANN approaches below figure 11.

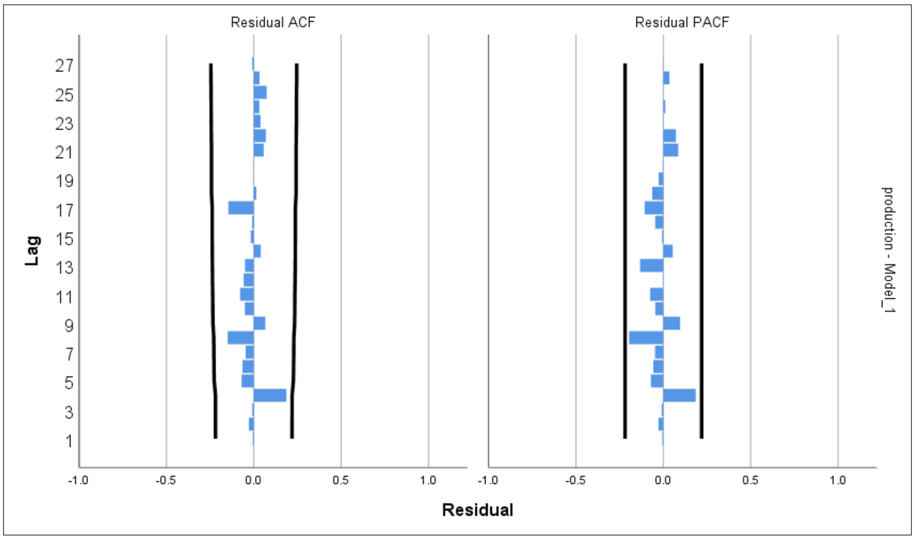

Fig 7: ACF and PACF of a Production Residual lot

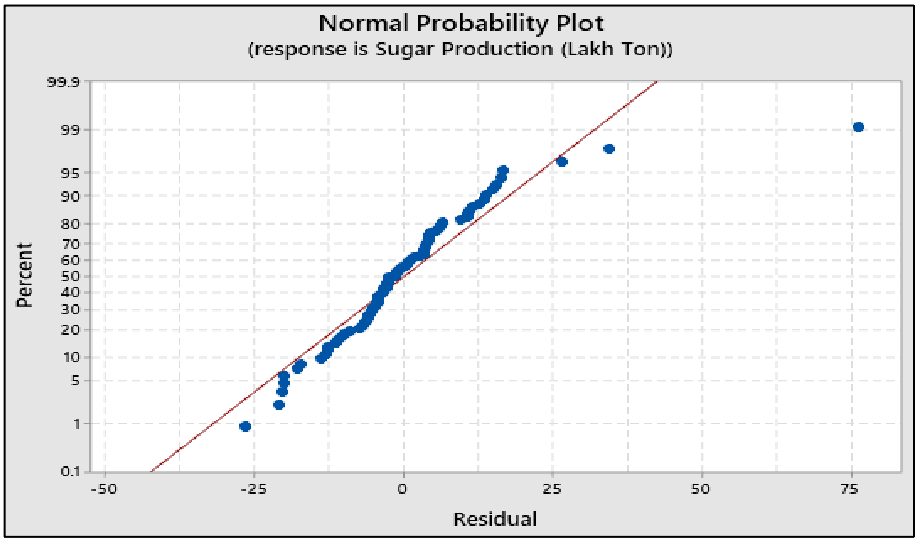

Fig 8: Residuals Plot of Normal Probability

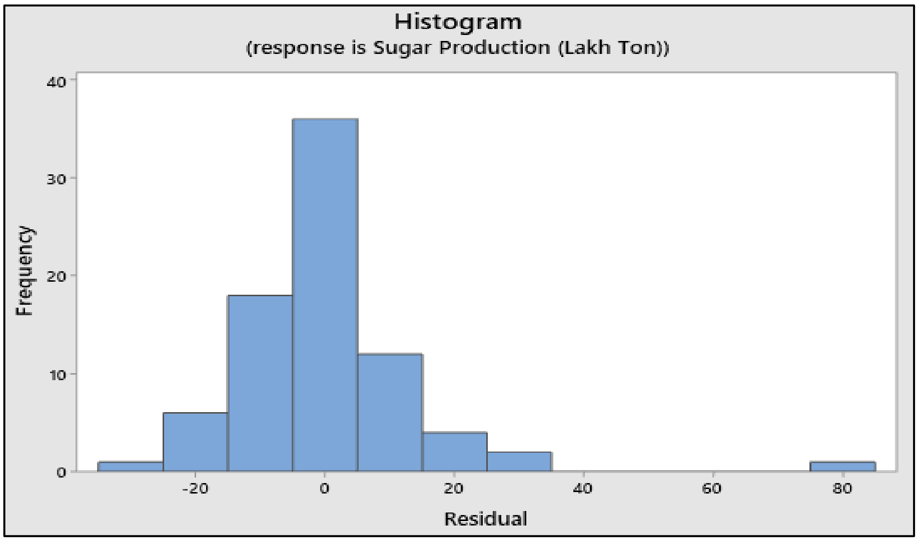

Fig 9: Histogram (Response is Production)

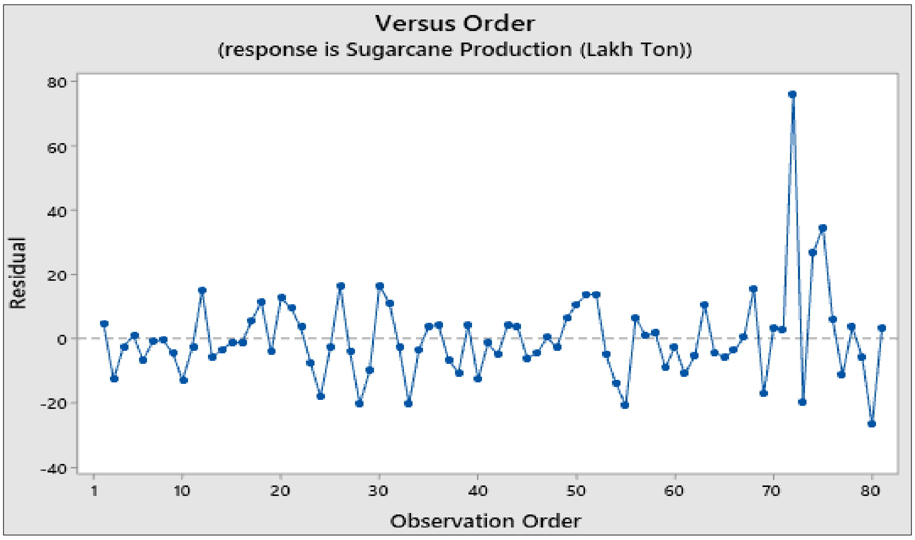

Fig 10: Residuals from ARIMA (1, 1, 0)

Table 3: AIC and BIC values of various combination fitted ARIMA models in sugarcane output in Bihar

Arima Model	Coefficient		AIC	BIC
	MA1	AR1		
Arima (0, 1, 0)	-	-	655.146	657.528
Arima (1, 1, 0)	-	-0.305	649.302	654.066
Arima (0, 1, 1)	0.287	-	649.781	654.545
Arima (1, 1, 1)	0.036	-0.273	651.328	658.474

Table 4: Selected model ARIMA (1, 1, 0)

Coefficient	Estimate	Std. Error	t- value	T(>ItI)
AR1	-0.305	0.108	-2.825	0.006

3.1. Testing for Residuals

Box-Ljung test: X-squared = 10.319, df = 17, p-value =0.89

Table 5: Sugarcane production in Bihar: Accuracy measurements of forecasted models

Year	Observed	ARIMA	ANN	ARIMA RD (%)	ANN RD (%)
2015-16	139.00	150.27	133.62	-8.11	4.16
2016-17	147.00	143.32	136.16	2.50	7.37
2017-18	139.82	145.83	137.7	-4.30	1.52
2018-19	116.61	143.28	138.61	-22.87	-18.87
2019-20	128.33	124.97	139.15	2.6	-8.43
Total				-30.18	-14.25
Average RD (MAPE)				6.036	2.85
RMSE				13.821	12.99

RD (%) ={(Observed Production – Predicted Production)/ Observed Production}*100

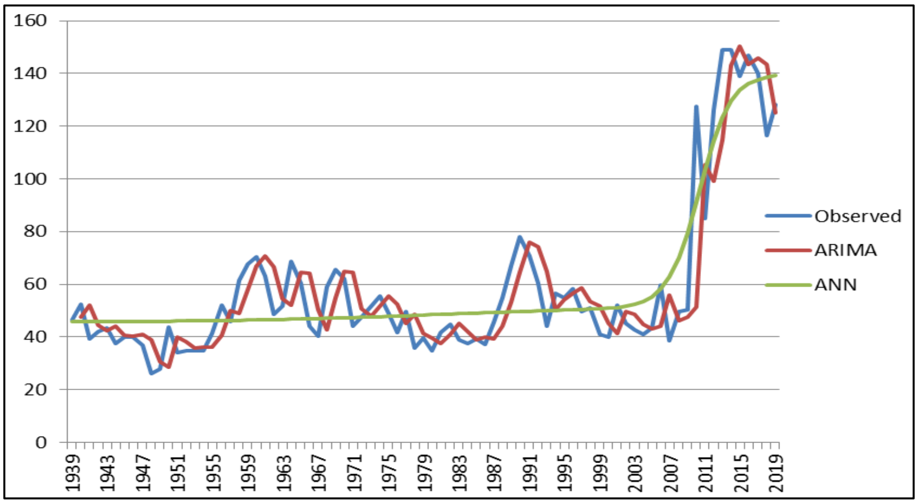

Fig 11: ARIMA and ANN methods graph for observed and predicted sugarcane production

For the specified models, it displays the accuracy measurements of all used techniques, such as ARIMA and ANN. The first 76 years of data were used to develop the model, and the final five years of data were used to ensure that the models chosen were accurate. In terms of forecasting precision, the ANN method outperformed ARIMA with an average relative deviation of 2.85 and a root mean square error (RMSE) of 12.99. The observed and projected values in the case of ARIMA had a relative deviation of 6.036 and a root mean square error (RMSE) of 13.821. As a result, only one ARIMA was used for predicting, and the outcomes for the next six years are reported in Table 6.

Table 6: Forecast sugarcane production (Lakh ton) through ARIMA (1, 1, 0) from year 2020-21 to 2025-26 at 95% Limits

Period	Forecast	Lower	Upper
2020-21	126.03	98.51	153.54
2021-22	128.00	94.50	161.50
2022-23	128.67	88.77	168.57
2023-24	129.74	84.69	174.78
2024-25	130.68	80.93	180.40
2025-26	131.67	77.64	185.70

It's utilised to figure out when the right model has been fitted. The purpose of our forecasting is to project future Sugarcane Crop Production values from 2020-21 to 2025-29. Figure 12 shows a time plot of the Sugarcane Production Predicted numbers, while Table 6 shows the forecast values for the next six years, with 95 percent confidence intervals.

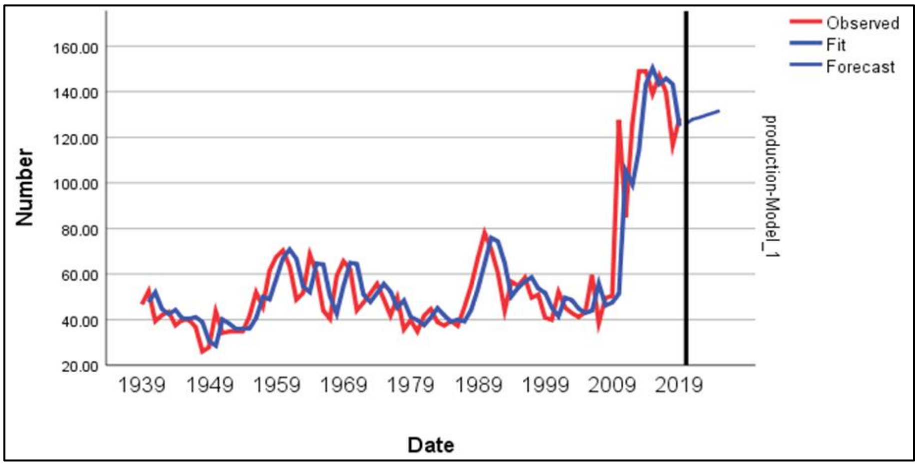

Fig 12: Forecast of Sugarcane Production for the year 2020-2025 through ARIMA (1, 1, 0) model

4. Conclusions and Recommendations

Among the numerous attempted ARIMA models, ARIMA (1, 1, 0) was shown to be the best for modelling and forecasting sugarcane production. In Bihar, ANN was proven to be capable of forecasting sugarcane output behaviour. In the instance of sugarcane production forecasts, ANN outperforms the other two methodologies and is recommended for ongoing use. A comparison plot of observed and anticipated values generated by the specified model reveals that the model is well-fitting. The ARIMA (1, 1, 0) model is the best forecasting model for sugarcane crop output for the years 2020-21 to 2025-26, according to Box- Jenkin's Methodology. Table 5 shows the predicted sugarcane crop output for the years 2020-21 to 2025-26, which shows a significant increase from 126.03 lakh to 131.67 lakh tonnes. Actual sugarcane production in 2014-15 was 149.00 lakh tonne, the highest ever recorded, and will be reduced to 26.00 lakh tonne as anticipated production in 1948-49 before rapidly increasing in subsequent years. This research could aid Bihar policymakers in developing macro-level policies for food security, a fair support price, and better sugarcane production planning. It also assists sugar mill operators with micro-level requirements such as high-yield varietals, fertilisers, insect control tactics, and proper irrigation and other sugarcane crop inputs. The proposed ARIMA model for forecasting for information and resource planning should be used by researchers, businesses, and farmers in Bihar who are making decisions about sugarcane crop yield.

Reference

1. Akin M, Eyduran SP. Forecasting harvest area and production of strawberry using time series analyses. Journal of Agricultural Faculty of Gaziosmanpasa University. 2017;34(3):18-26.
2. Allen PG. Economic forecasting in agriculture. International Journal of Forecasting. 1994;10(1):81-135.
3. Annonymous. Data Bank of Crops Unit-I, Crops Division (DAC & FW), India, 2020.
4. Annonymous. Sugarcane Industries Department, Bihar, 2020.
5. Askar C, James J. Crop yield prediction using time series model. Journal of Economic and Economic Education Research. 2014;15:53-67.
6. Ayesha T, Nusrat H. Forecasting of maize area and production in Pakistan. ESci J Crop Prod. 2013;2(2):44-48.
7. Azam M, Khan M. Significance of the sugarcane crops with special and reference to NWFP. Sarhad J Agric. 2010;26:289-295.
8. Bharati B, Panta R, Khanal K. Assessing socio-economic condition of sugarcane producers in Nawalparasi district of Western Nepal, Biomedical Journal of Scientific & Technical Research. 2018;12(3):9296-9297.
9. Boken VK. Forecasting spring wheat yield using time

series analysis. Agronomy Journal. 2000;92(6):1047-1053.
10. Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis: Forecasting and Control, third ed., Prentice Hall, Englewood Clifs, 1967.
11. Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis: Forecasting and Control, third ed., Holden-Day, San Francisco, 2009.
12. Cadenas E, Rivera W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model, Renew. Energy. 2010;35(12):272-273.
13. Cheng B, Titterington DM. Neural networks: A review from a statistical perspective. Statistical Science. 1994;9:2-54.
14. Devaki C, Mohideen DAK. Forecasting analysis for sugarcane production in Tamil Nadu using ARIMA Model. Journal of Xi'an University of Architecture & Technology. 2021;13(6):778-790.
15. Devi M, Kumar J, Mallik DP, Mishra. Forecasting of wheat production in Haryana using hybrid time series model. Journal of Agriculture and Food Research. 2021;5:1-5.
16. Díaz-Robles LA, Ortega JC, Fu JS, Reed GD, Chow JC, JG. A hybrid ARIMA and artificial neural networks to forecast particulate matter in urban areas: the case of Temuco, Chile, Atmos. Environ. 2008;42(35):8331-8340.
17. Dickey A, Fuller WA. Distribution of the estimators for autoregressive time series with a unit-root. Journal of the American Statistical Association. 1979;74:427-431.
18. Faruk DO. A hybrid neural network and ARIMA model for water quality time series prediction, Eng. Appl. Artif. Intell. 2010;23(4):586-594.
19. Ghosh H, Prajneshu AK, Paul AK. Study of Nonlinear Time Series Modelling in Agriculture, IASRI. Project Report, 2005.
20. Gujarati DN. Basic Econometrics. 4th ed.McGraw-Hill Book Company Inc. New York, 2003, 465-7.
21. Gujarati DN. Basic Econometrics, McGraw Hill, New York, 1995.
22. Houghton JT, Meira Filho LJ, Callander BA, Harris A, Kattenberg A, Maskell AK. (Eds.). Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1996, 572.
23. Ray M, Rai M, Ramasubramanian AV, Singh KN. ARIMA-WNN hybrid model for forecasting wheat yield time series data, J Indian Soc. Agric. Stat. 2016;70(1):63-70.
24. Makridakis S, SC Wheelwright SC, Hyndman RJ. Forecasting: Methods and Applications, third ed., Wiley, Chichester. 1998.
25. Mishra P, Matuka A, Abotaleb MSA, *et al*., Modeling and forecasting of milk production in the SAARC countries and China, Model. Earth Syst. Environ, 2021, 1-13.
26. Mishra P, Al Khatib MG, Sardar I, *et al*., Modeling and forecasting of sugarcane production in India, Sugar Tech, 2021, 1-8.
27. Mishra P, Sahu PK, Dhekale BS, Vishwajith KP. Modeling and forecasting of wheat in India and their yield Sustainability, J Econ. Dev. 2015;11(3):637-647.
28. Muhammad F, Javed MS, Bashir MM. Forecasting sugarcane production in Pakistan using ARIMA Models. Pakistan Journal of Agricultural Sciences. 1992;9(1):31-36.
29. Onwueme IC, Sinha TD. Field Crop Production in Tropical Africa. CTA, Wageningen, Netherlands, 1991, 401-411.
30. Onwueme IC. Crop science: Tropical agricultural series. Cessel, London, 1978, 89-90.
31. Priya Krishna SR, Suresh KK. A study on pre-harvest forecast of sugarcane yield using climatic variables, Statistics and Application. 2009;7(1):1-8.
32. Rankja NJ, Kalsariya BN, Butani AM. An analysis of growth trends in cotton productivity for Rajkot district of Gujarat, Guj. J Ext. Edu. 2017;28(2):389-392.
33. Remus W, O'Connor M. Neural Networks for Time-Series Forecasting, Springer, New York, 2001.
34. Ripley BD. Statistical aspects of neural networks. In Networks and Chaos –Statistical and Probabilistic Aspects eds, Barndorff, O.E., Nielsen J., Jensen, L. and Kendall, W.S. London: Chapman and Hall, 1993, 40-123.
35. Saeed NA, Saeed M, Zakria TM. Bajwa, Forecasting of wheat production in Pakistan using ARIMA models, International Journal of Agricultural Biology. 2000;2(4):352-353.
36. Sajid A, Badar N, Fatima H. Forecasting production and yield of sugarcane and cotton crops of Pakistan for 2013-2030. Sarhad J Agric. 2015;31(1):1-10.
37. Shah SAA, Zeb N, Alamgir. Forecasting major food crops production in Khyber Pakhtunkhwa, Pakistan. J App. Ad. Res. 2017;2(1):21-30.
38. Suman, Verma U. Autoregressive Integrated Moving Average models for sugarcane production in Haryana. International Journal of Computer & Mathematics Sciences. 2016;5(12):33-38.
39. Suresh KK, Priya Krishna SR. Forecasting sugarcane yield of Tamil Nadu using ARIMA model. Sugar Technology. 2011;13(1):23-26.
40. White H. Learning in artificial neural networks: A statistical perspective, Neural Computation. 1989;1:425-464.
41. Yaseen M, Zakriya M, Shahzad ID, Khan MI, Javed MA. Modeling and forecasting the sugarcane yield of Pakistan. Int. J Agric. Biol. 2005;7:180-183.
42. Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting. 1998;1435-62.