INNOVARE JOURNAL OF AGRICULTURAL SCIENCE

Vol 12, Issue 2, 2024

ISSN - 2321-6832

Research Article

# FORECASTING WHEAT AREA AND PRODUCTION IN NEPAL USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL (ARIMA)

BIBEK GAUTAM[1]*, SUSHMA ADHIKARI[2]

School of Agriculture, Agriculture Economics, Om Sterling Global University, Hisar, Haryana, India. [2]Department of Himalayan College of Agriculture Science and Technology, Kathmandu, Nepal. Email: gautam53.bg@gmail.com

*Received: 28 December 2023, Revised and Accepted: 06 February 2024*

## ABSTRACT

Wheat is one of the major staple crops of Nepal grown from Plains to high hills. Nepal has observed the increased production of wheat in the last decade but is unable to produce enough to meet country demand. Autoregressive Integrated Moving Average Model (ARIMA) was implied to forecast the Wheat area and production in Nepal from 2021 to 2030 using available data from FAO. Rstudio software with forecast package using "auto.arima ()" was used for selecting the suitable model. On analysis, it was observed that ARIMA model (2,1,3) and (0,1,0) were found appropriate for the forecasting of production and area of wheat with lowest Akaike's information criterion 682.01 and 537.76 respectively among competitive models. Results from the model suggested the increase in the area and production of the wheat by 1.32% and 1.72%, respectively, but on decreasing rate which suggests to act in the productivity increasing traits for achieving food security.

**Keywords:** Forecasting, Area, Production, ARIMA.

 DOI: http://dx.doi.org/10.22159/ijags.2024v12i1.50055. Journal homepage: https://innovareacademics.in/journals/index.php/ijags

## INTRODUCTION

Wheat (*Triticum aestivum*) is being grown worldwide in 219 million hectares with average productivity of 3.4 t/ha (FAO, 2021). Wheat has been one of the essential food from the initiation of Civilization and Wheat accounts for the largest share in cereal production after maize and rice worldwide (CIMMYT, 2013). Wheat ranks 1st in the source of protein supply in underdeveloped countries with 21% and 20% contribution in food calories and protein respectively to 4.5 billion people from 94 underdeveloped nations (Devkota and Phuyal, 2015).

Wheat is one of the staple winter-season cereal crop after rice and maize in both area and production in Nepal (Kharel *et al*., 2021). Although Wheat is cultivated on all agro-climatic regions (65–4000 masl) of Nepal, the terai region contributes production of 65.2% of total wheat production with coverage of 57.8% (Pandey *et al*., 2017). In present, the total production of wheat is 2185289 Metric ton with an area covered 707505 ha (FAO, 2021).

The area as well production of wheat in Nepal has been distinctly increased in recent years (Gairhe *et al*., 2017) as an impact of the introduction of semi-dwarf varieties (Pandey *et al*., 2017) but the average productivity of wheat is low in Nepal as compared to neighboring countries India and China (Sendhil *et al*., 2019). The main cause of low wheat productivity in Nepal is due to the lack of proper scientific knowledge on the application of fertilizers, insecticides, pesticides and management practices (Bhatta *et al*., 2020). Besides the Practice Climate change have been one of the major causes on decrease of wheat production. It was observed that increased temperature has a negative impact on the wheat production (Luigi *et al*., 2008) while (Asseng *et al*., 2015) revealed that increase in temperature up to certain range can create positive impact on high altitude regions. Although the production and productivity of wheat are increasing the domestic production is not enough to fulfill its demand and therefore a large amount is imported from neighboring countries. Nepal imported 90 thousand tons of wheat in 2020 (MOALD, 2020). Regarding the wheat trade balance the export quantity and value have decreased significantly since 1961 but after that import quantity has been increasing with very less or no export (Gairhe *et al*., 2017).

Analyzing the data over time are called time series analysis. The main objective of time series analyses is to generate a model that forecasts the future based on past data (Revels *et al*., 2020). Auto-Regressive Integrated Moving Average (ARIMA) model has been one of the common methodologies to interpret time-series data and to forecast. While ARIMA models are commonly used for analyzing volatile time series data, these data do not necessarily need to fluctuate with a certain intensity.

Arima model has been extensively used in forecasting in various sectors such as carbon emission (Nyoni and Bonga, 2019), stock market (Dhyani *et al*., 2020), trade flow (Tyagi and Shah, 2021), and so on. The ARIMA model is frequently employed when predicting agricultural outcomes. (Senthamaraikannan and Karuppasamy, 2020) predicted the output of paddy in the Indian states of Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu. Similarly, (Iqbal *et al*., 2015) predicted Pakistan's wheat output and area. Hamja (2014) Predicted the yield of fruits in Bangladesh. Padhan (2012) used ARIMA model for forecasting the productivity of 34 different products in India.

2030 is the year of major concern the millennium development goals declaration by the United Nations has set a foundation for sustainable development goals (SDGs) to be achieved by 2030. The 17 goals of the SDGs are highly interlinked. Among them, Goal 2-"Ending hunger" is a more important for the agriculture sector.

Under SDG target 2, Nepal aims to double the 2015 production of 1975 thousand tons of wheat by 2030 (NPC, 2017). Similarly (Prasad *et al*., 2022) predicted that Nepal will have a Consumption demand of 1158.70 thousand tons of wheat in the year 2030. Considering the following facts this article forecasts the area and Production status of wheat in Nepal by the end of 2030 as per the SDG target and capacity to fulfill the consumption demand needs.

## METHODS

### Data collection

Data on the Production of Wheat were obtained from the FAO Stat website. Data ranging from 1961 to 2020 i.e. 59 years was used for the analysis.

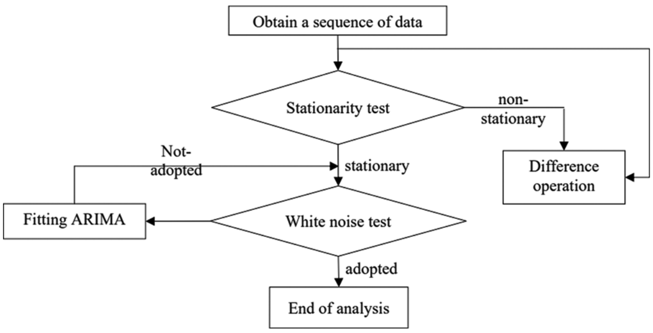
**Empirical model**

ARIMA model is used for forecasting using time series data. This model is comprised of three separate components: Autoregressive (AR) (p), I(d), and moving average (MA) (q) process. The initial component is the autoregressive, or AR (p), process. This component reveals that future predictions or forecasting will be based on the previous data available. The second component of the ARIMA model is known as an integrated, or I(d), process. In the process of analysis time series data available should be taken as stationary to account the accurate model for the future predictions. Data are considered to be stationary if the average and variance of the data do not vary over the period. If the available data time series shows that data is not stationary it needs to be converted to stationary which is called differentiation of the data or preparing an integrated time series. The final component of the ARIMA model is the MA, or MA (q), process. This term is applied for error terms in the series data. If errors are not found constant, the average of the current and previous observation for every observation of the time series must be considered. (Revels *et al*., 2020).

The major advantage of this model is that it does not depend on any other explanatory variables rather it depends on its previous value (Gujarati and Porter, 2009). The major drawback of this model is that ARIMA model does not use other variables that might affect production other than the previous values.

ARIMA model analysis follows 4 steps of Box-Jenkin methodology: model identification, parameter estimation, diagnostics checking, and forecasting. The first step involves the determination of p, d, q values for the ARIMA model where d was determined, order of the difference at which the data will be stationary. AR order p was identified by partial autocorrelation function (PACF) and then MA order q was identified using autocorrelation function (ACF). After the model specification, parameter estimation followed by diagnostic operation was done (Gujarati and Porter, 2009).

The data analysis was performed using Rstudio software with a forecast package using "auto.arima ()" command and Microsoft Excel.




Source: (Dai and Chen, 2019)

**Data analysis**

The data analysis was performed using Rstudio software. The R software has a "forecast" package which instead of examining different models manually, using "auto.arima()" command performs several iterations and come up with an appropriate model having the lowest Akaike's information criterion (AIC) among competitive models. Therefore, for the study "auto.arima ()" was used.

## RESULTS AND DISCUSSION

**Present trend analysis of area and production**

Fig. 1 represents the Trend of Wheat production and Consumption in Nepal from 1961 to 2020 respectively.

It was observed that both the area and Production have been increased in time duration considered. During these 50 years, the area has been increased by 84% while production has been increased by 93%. Release of the modern variety, increase access to irrigation, extension, high return, and market price have driven the increased wheat area and production during the period. (Gairhe *et al*., 2017).

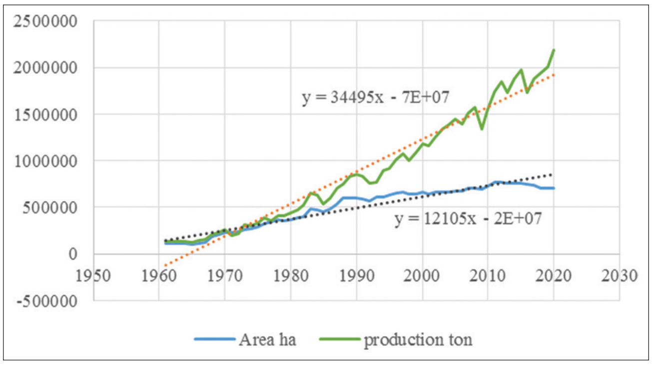



**Fig. 1: Trend analysis of wheat area and production in Nepal**

**Forecasting wheat area and production**

Box- Jenkins Method is a tool used to identify the model in ARIMA model. Initially lags is determined for the autoregressive, integrated, and moving average components. Second, the parameters are estimated for the models. Third, validation of the model is accomplished and fourth it involves the forecasting using the prior validated model (Revels *et al*., 2020).

*Checking stationarity*

As per the steps for ARIMA for model identification initially, we need to identify weather data is stationary or not. In case data are not stationary or shows the trends we need to differentiate the data to make it constant. ACF and PACF was used which is one of the simple ways to test for stationarity (Gujarati and Porter, 2009).

Lags demonstrate the correlation among successive observations. The ACF plot above in Zero and declines slowly indicates the non-Stationary time series or in case of the stationary time series the lag drops to zero rapidly. When the usage of ACF plots, there's no precise threshold or charge of degradation to decide the differencing needed (Trottier *et al*., 2006).

On analysis of the production and Area data, it was revealed that the lags in ACF Stay above Zero and decline slowly indicating the non-stationary time series (Figs. 2 and 4). Therefore, the difference was used so that data becomes stationary (Figs. 4 and 7) for production and area, respectively.

*Selection of optimum model*

The second order or process of the methodology is to determine the optimum model. The "auto.arima ()" commands in the "forecast" package in R software was used to determine the optimum model identification. The "auto.arima ()" suggested that the optimal model for wheat production was ARIMA (2,1,3) with drift. The AIC and BIC and Log likelihood of the ARIMA (2,1,3) was found 682.01, 696.55, and-334, respectively and least among all competitive models.

Similarly optimal model for the wheat area was ARIMA (0,1,0) with drift. The AIC and BIC of the ARIMA (0,1,0) were found least, i.e., among all competitive models. The AIC, BIC, and Log likelihood were estimated 537.76, 541.92, and -266.88 respectively for the wheat area.

The ACF and PACF of residuals for ARIMA model (2,1,3) is presented on Figs. 5 and 6 respectively. Similarly, ACF and PACF of residuals for ARIMA (0,1,0) are presented in Figs. 7 and 8, respectively.

**Table 1: Forecasted wheat production and area in Nepal (2021-2030)**

| Year | Wheat Production (thousand tons) | | | | Wheat Area (Thousand hectare) | | | |
|---|---|---|---|---|---|---|---|---|
| | Point Forecast | Hi95 | Lo95 | Difference (%) | Point Forecast | Hi95 | Lo95 | Difference (%) |
| 2021 | 2176.36 | 2036.39 | 2316.35 | - | 717.63 | 673.55 | 761.71 | - |
| 2022 | 2174.58 | 2009.84 | 2339.33 | −0.08 | 727.76 | 665.42 | 790.09 | 1.39 |
| 2023 | 2238.14 | 2073.32 | 2402.98 | 2.84 | 737.89 | 661.54 | 814.23 | 1.37 |
| 2024 | 2279.59 | 2109.76 | 2449.42 | 1.82 | 748.01 | 659.86 | 836.17 | 1.35 |
| 2025 | 2327.9 | 2153.99 | 2501.81 | 2.08 | 758.14 | 659.58 | 856.70 | 1.34 |
| 2026 | 2373.48 | 2193.41 | 2553.56 | 1.92 | 768.27 | 660.30 | 876.24 | 1.32 |
| 2027 | 2419.54 | 2232.23 | 2606.86 | 1.90 | 778.40 | 661.78 | 895.01 | 1.30 |
| 2028 | 2465.00 | 2269.00 | 2661.02 | 1.84 | 788.52 | 663.85 | 913.19 | 1.28 |
| 2029 | 2510.25 | 2304.28 | 2716.22 | 1.80 | 798.65 | 666.42 | 930.88 | 1.27 |
| 2030 | 2555.15 | 2337.98 | 2772.33 | 1.76 | 808.78 | 669.39 | 948.16 | 1.25 |

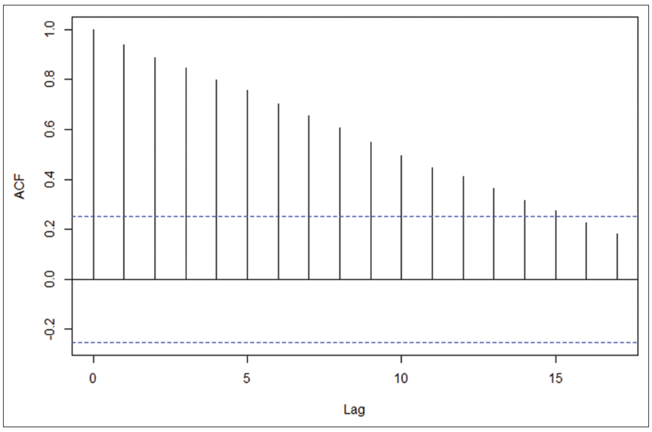



**Fig. 2: ACF of wheat production**

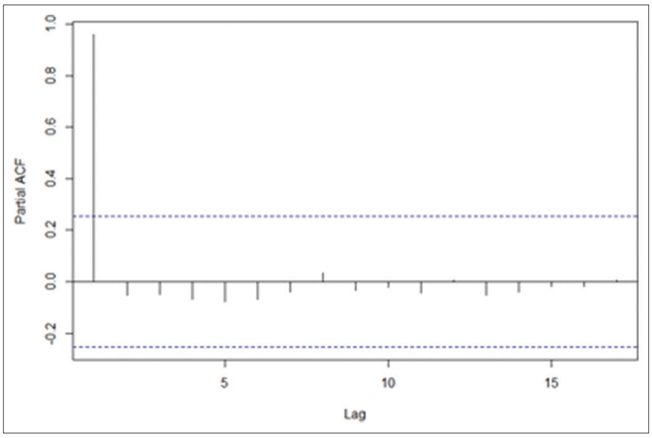



**Fig. 5: ACF of wheat area**

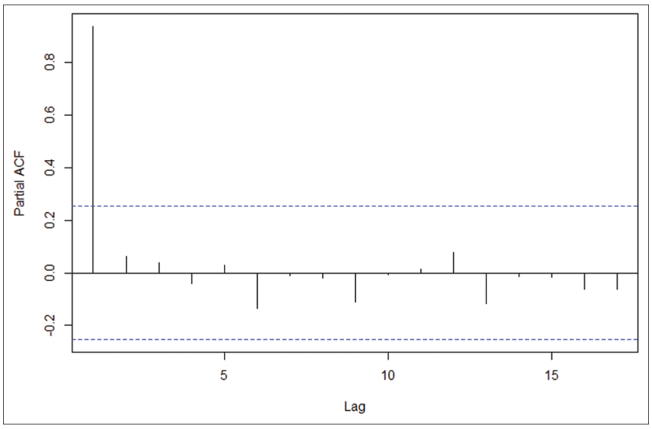



**Fig. 3: PACF of wheat production**

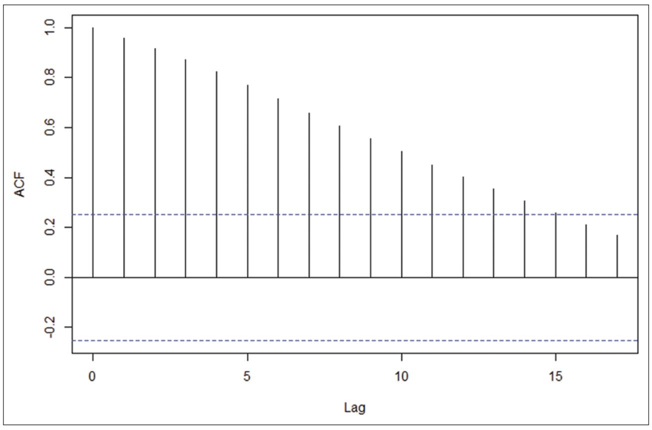



**Fig. 6: PACF of wheat area**

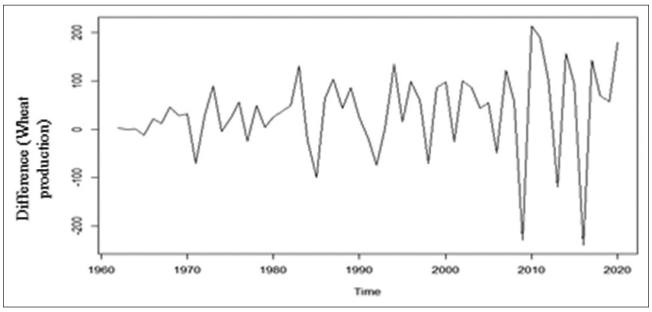



**Fig. 4: First difference of wheat production**

### *Validating the model*

Validation of these identified models should be done before forecasting. In time series, validation can be done by checking the model for white noise. White noise means that the residuals are not auto-correlated (Gujarati and Porter, 2009). In the presence of autocorrelations, a model selected is not always sound sufficient to make correct predictions. The Ljung-Box check has a null speculation that the residuals are white noise and the opportunity is that as a minimum one residual is auto-correlated. If the p-price of the Ljung-Box check is underneath 0.05, then null speculation of white noise could be rejected and the version will now no longer be beneficial for forecasting (Gujarati and Porter, 2009).

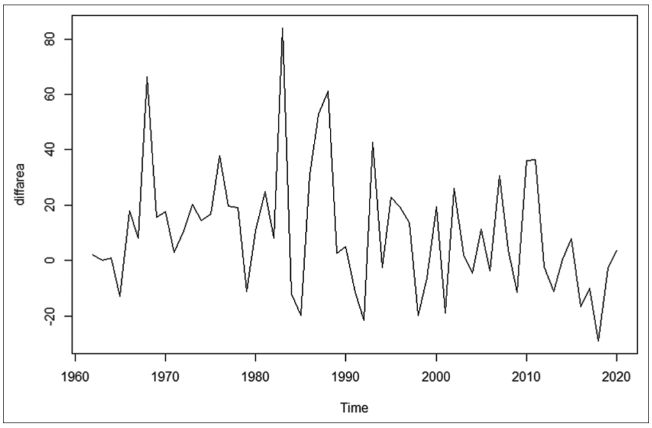



**Fig. 7: First difference of wheat area**

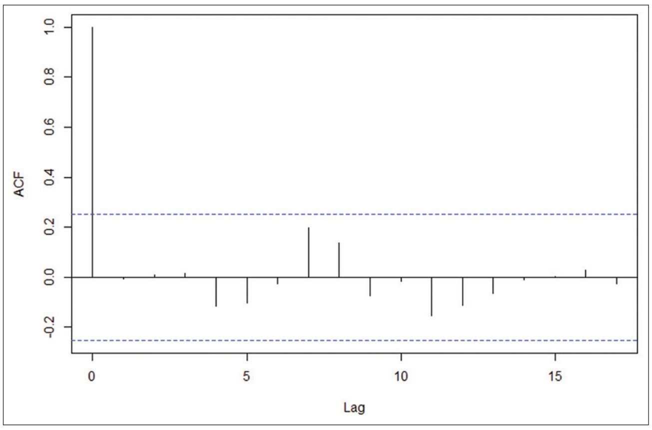



**Fig. 8: Differentiated ACF of wheat production**

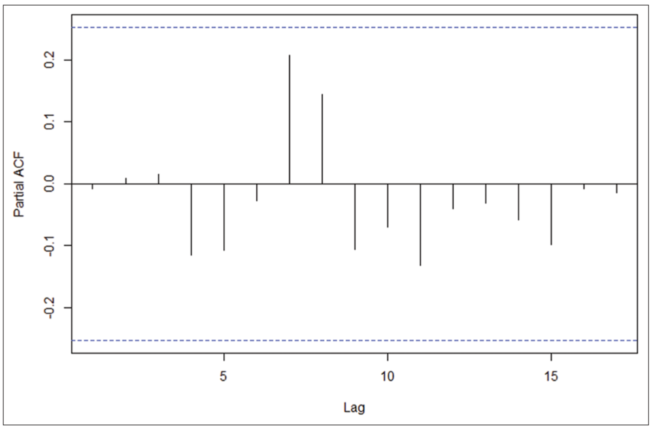



**Fig. 9: Differentiated PACF of wheat production**

On analysis, it can be observed that average of the residuals is close to zero in all cases with no significant correlation in the residuals series i.e. data is white noise for the both models (Figs. 5 and 7). Similarly, p value was observed to be 0.359 and 0.5337 for the Wheat Production and area.

### *Forecasting*

Using ARIMA (2,1,3) with drift forecasting of production was done at a 95% level of confidence (Fig. 9). Similarly forecasting of wheat area in Nepal was done using the ARIMA (0,1,0) model with a 95% confidence level (Fig. 10). It was observed that there will be positive growth in the both area and Production of the Wheat in Nepal during analysis.

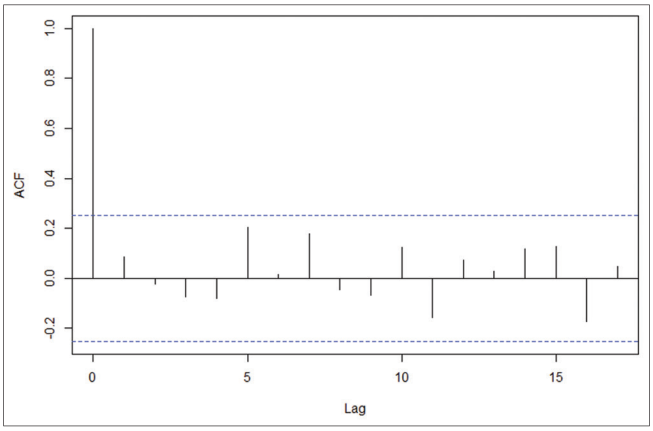



**Fig. 10: Differentiated ACF of wheat area**

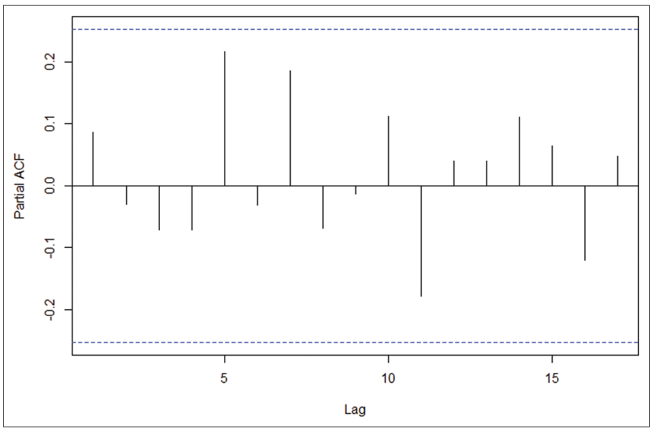



**Fig. 11: Differentiated PACF of wheat area**

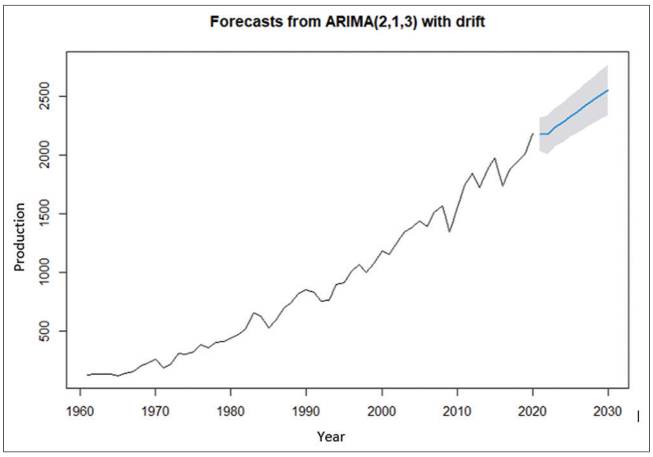



**Fig. 12: Forecast of Wheat Production using ARIMA (2,1,3) model**

Table 1 presents the forecasted production and area of wheat in Nepal from 2021 to 2030. Data reveal an average growth of 1.76% in production and 1.32% in wheat area. It was predicted that on the year 2030 Area under wheat will be 808.78 thousand hectares and Production will be 2555.15 thousand tons. However, the growth of both area and production observed is decreasing indicating serious chances of food insecurity leading to the higher import to meet the needs. It was also observed that there is a strong positive correlation between the forecasted area and production with correlation Coefficient of 0.98 which means that the area has very strong contribution in the production of the wheat.

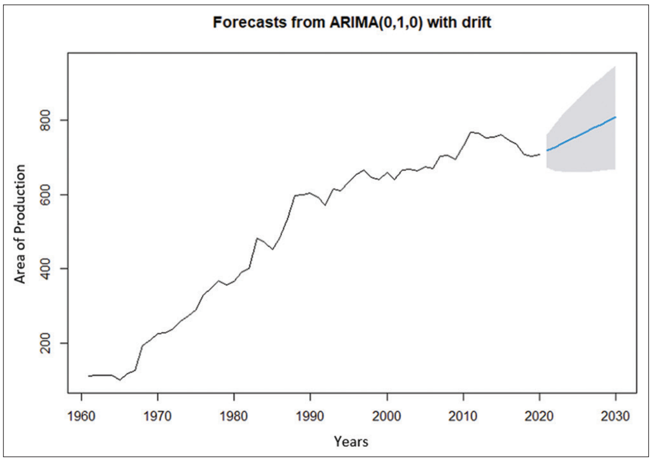



**Fig. 13: Forecast of Wheat Area using ARIMA (0,1,0) model**

## CONCLUSION AND POLICY IMPLICATION

Time series data from 1961 to 2020 were obtained from FAO database and analyzed using ARIMA and Box-Jenkins Methodology to forecast the area and production of wheat in Nepal. On analysis of the data, it was observed that the average growth rate of wheat cultivation area is 1.32% and average growth of Wheat production is 1.72%. Despite the positive growth rate growth is in decreasing trend. It was predicted that Nepal will have a total production of 2555.15 thousand tons with a coverage of 808.78 thousand hectares on end of 2030 which is around 80% more than the demand for 2030 as predicted by Prasad *et al.* [19]. Similarly, it was found that Nepal will achieve only 22% increase in the production compared to 2015, a starting year for the SDG.

It was also observed the production and area increases in decreasing rate which might be due to decrease in land availability for cultivation due to urbanization and land fragmentation. Since the correlation of the area is observed strongly positively correlated, considerations should be taken to maintain area under the wheat cultivation. Based on the findings of the study following points are suggested to ensure the production of wheat in Nepal:

- The government needs to take proper steps for controlling the land fragmentation and unplanned settlement in cultivable land.
- As land always has been the scare resource, productivity-related attributes could be areas of improvement for increasing the wheat production.
- Focus should be made on developing the varieties with more production potentials considering the climate change issues.
- Developing the irrigation facility, access to the quality seeds, agriculture loans, and farming equipment's followed by store facility and price regulations are subsequently necessary for Quality production.

## AUTHOR CONTRIBUTION

In this collaborative effort, each author played a crucial role in shaping the research. Author 1-Bibek Gautam led the conceptualization and design, providing the framework for the study. Author 2- Sushma Adhikari conducted data collection and analysis, contributing valuable insights. Both played a pivotal role in drafting and revising the manuscript, ensuring clarity and coherence. All authors actively participated in discussions and revisions, contributing their expertise to enhance the overall quality of the work.

## CONFLICTS OF INTERESTS

We both authors affirm commitment to transparency and integrity in academic pursuits. In accordance with ethical standards, I declare that I have no conflicts of interest that could unduly influence the objectivity of my research or its publication. Full disclosure is paramount to maintaining the credibility and trust of the scholarly community.

## AUTHORS FUNDING

There was no any funding received by authors for this study. The authors themselves were fully responsible for required cost and logistics.

## REFERENCES

Asseng, S., Ewert, F., Martre, P., Rotter, R. P., Lobell, D. B., Cammarano, D., & Zhu, Y. (2015). Rising temperatures reduce global wheat production. *Nature and Climate Change*, *5*(1), 143-147.

Bhatta, R. D., Amagain, L. P., Subedi, R., & Kandel, B. P. (2020). Assessment of productivity and profitabilty of wheat using nutrient expert®-wheat model in Jhapa District of Nepal. *Heliyon*, *6*(6), e04144.

CIMMYT. (2013). *Annual Report-2013*. International Maize and Wheat Improvement Center.

Dai, J., & Chen, S. (2019). The application of ARIMA model in forecasting population data. *Journal of Physics Conference Series*, *1324*(1), 012100.

Devkota, N., & Phuyal, R. K. (2015). Climatic impact on wheat production in terai of Nepal. *The Journal of Development and Administrative Studies*, *23*(1), 1-22.

Dhyani, B., Kumar, M., Verma, P., & Jain, A. (2020). Stock market forecasting technique using Arima model. *International Journal of Recent Technology and Engineering*, *8*(6), 2694-2697.

FAO. (2021). *FAO Stat*. Retrieved from https://www.fao.org/faostat/en/#home

Gairhe, S., Karki, T. B., Uapdhyay, N., & Sapkota, S. (2017). *Trend analysis of wheat area, production and productivity in Nepal: An overview, proceedings of 30th National winter crops workshop*. Nepal: Nepal Agriculture Research Council.

Gujarati, D.N., & Porter, T. (2009). *Basic econometrics*. Boston: McGraw-Hill Education.

Hamja, M. A. (2014). Forecasting major fruit crops productions in Bangladesh using box-jenkins ARIMA model. *Journal of Economics and Sustainable Development*, *5*(7), 96-107.

Iqbal, N., Bakhsh, K., Maqbool, A., & Ahmad, S. A. (2015). Use of the ARIMA model for forecasting wheat area and production in Pakistan. *Journal of Agriculture Social Sciences*, *1*(2), 120-122.

Kharel, M., Ghimire, Y. N., Adhikari, S. P., Subedi, S., & Paudel, H. K. (2021). Economics of production and marketing of wheat in Rupandehi district of Nepal. *Journal of Agriculture and Natural Resources*, *4*(2), 238-245.

Luigi, M., Luigi, B., & Alessandra, F. (2008). Changes in land-use/land-cover patterns in Italy and their implications for biodiversity conservation. *Landscape Ecology*, *22*(4), 617-631.

MOALD. (2020). *Statistical information on Nepalese agriculture 2076/77*. Agribusiness Promotion and Statistics Division, Ministry of Agriculture and Livestock Development, Kathmandu, Nepal.

NPC. (2017). *Nepal sustainable development goals baseline report*. National Planning Commission, Government of Nepal.

Nyoni, T., & Bonga, G.W. (2019). Prediction of $CO_2$ Emissions in India using ARIMA models. *Journal of Economics and Finance*, *4*(2), 1-10.

Padhan, P.C. (2012). Application of ARIMA model for forecasting agricultural productivity in India. *Journal of Agriculture Social Sciences*, *8*(2), 50-56.

Pandey, G., Yadav, L., Tiwari, A., Kahtri, H., Basnet, S., Bhattarai, K., & Khatri, N. (2017) Analysis of yield attributing characters of different genotypes of wheat in Rupandehi, Nepal. *International Journal of Environment*, *Agriculture and Biotechnology*, *2*(5), 2374-2379.

Prasad, S. K., Pullabhotla, H., & Kumar, G. (2022). *Supply and demand for cereals in Nepal, 2010-2030.* Delhi: International food Policy Research Institute. Retrieved from https://www.ifpri.org/sites/default/files/publications/ifpridp01120.pdf; 2011

Revels, S., Kumar, S. P., & Ben-Assuli, O. (2020). Predicting obesity rate and obesity-related healthcare costs using data analytics. *Health Policy and Technology*, *2*(1), 198-207.

Sendhil, R., Kiran, T. M., & Singh, G. P. (2019). Wheat production in India: Trends and prospects. In *Recent advances in grain crops research*. London, UK: IntechOpen Publishers.

Senthamaraikannan, K., & Karuppasamy, K. M. (2020). Forecasting for agricultural production using Arima model. *Palarch's Journal of Archaeology of Egypt Egyptology*, *18*(7), 5940-5949.

Thabani, N., & Bonga, W.G. (2019). Prediction of $CO_2$ emissions in India using ARIMA models. *Journal of Economics and Finance*, *4*(2), 1-10.

Trottier, H., Philippe, P., & Roy, R. (2006). Stochastic modeling of empirical time series of childhood infectious diseases data before and after mass vaccination. *Emerging Themes Epidemiology*, *3*(1), 1-6.

Tyagi, A., & Shah, U. (2021). Modeling the direction and volume of trade flows in global crisis, COVID-19. *Journal of The Institution of Engineers*, *102*(6), 1225-1231.