Journal of Agriculture, Biology and Applied Statistics
Vol. 4, Nos. 1-2, 2025, pp. 1-9

URL : www.arfjournals.com
https://DOI:10.47509/JABAS.2025.v04i1-2.01

Ameliorating Ragi Production Forecasting Accuracy using ARIMA based on EEMD Decomposition Model in Odisha

Abhiram Dash[1] and Madhu Chhanda Kisan[2]
[1]Asst. Prof. (Agril. Statistics), Department. of Agril. Statistics, College of.Agriculture, (OUAT), Bhubaneswar, Odisha University of Agriculture and Technology
[2]Asst. Prof. (Agril. Statistics), College of.Agriculture, (OUAT), Bhawanipatna, Odisha University of Agriculture and Technology

To cite this article

Dash, A., & Kisan, M. C. (2025). Ameliorating Ragi Production Forecasting Accuracy using ARIMA based on EEMD Decomposition Model in Odisha. Vol. 4, Nos. 1-2, pp. 1-9.

Abstract: Odisha is one of the agricultural dependent states of India. Due to Odisha Millet mission millets have been popularized throughout the state. Ragi is now one of the super food that is being consumed by majority of the population which has a scope of growing in a large scale in the future. In this study ARIMA model was compared with hybrid Ensemble Empirical Mode Decomposition (EEMD)-ARIMA to evaluate the past behavior of the time series data regarding the production of Ragi, in order to make inference about its future behavior. The forecasting performance of these models are evaluated and compared by using the Root Mean Square Error, Mean Absolute Percentage Error and Mean absolute error. The findings reveal that superiority of Hybrid EEMD-ARIMA model than the ARIMA model for forecasting.

Keywords: ARIMA, EEMD, Forecasting, Hybrid EEMD-ARIMA, RMSE

Introduction

Ragi, also known as finger millet has always been an integral part of the diet of the people of Odisha. Previously, it was primarily consumed by individuals engaged in manual and physical labour since it digests slowly and allows them to work for extended period of time. Odisha Millet mission has ameliorated the all-round millet cultivation throughout Odisha, giving Ragi a spotlight to thrive as a super food. It is considered as a super food because of it's exceptional nutritional profile as well as additional health benefits. Ragi is rich in dietary fibers, protein, vitamins and minerals like Ca, Fe, P. It is adopted as safe gluten free alternative as well as it's high fiber content makes it a supplement for weight loss and also good for the dietary health. It contains antioxidants like ferulic acid and tannins which reduces the risk of

Article History Received : 04 January 2025; Revised : 3 February 2024; Accepted : 14 February 2024; Published : 02 June 2025

cardio vascular diseases. Ragi has a low glycemic index which is beneficial for people affected with diabetes or who have a risk of diabetes.

In Odisha, Koraput has the biggest share of area that is under Ragi and it is one of the biggest producer of Ragi all over India. As per 2019-20 data Ragi contributes about 116,800 ha, producing about 128,730 tonnes of the crop every year. The present study is focusing on forecasting the status of Ragi as crop which will help the Government, farmers and policy markers to adopt new technology for a better status of the crop.

Generally agricultural data are known to be non-stationary, so in this study two models are used one is Box-Jenkin's ARIMA model and the other is EEMD embedded hybrid ARIMA to analyze the timeseries data of production of Ragi. ARIMA is a traditional time series model that focuses on linear trend and stationary data, while EEMD-ARIMA is an enhanced approach that incorporates EEMD to handle nonlinearity and non-stationarity. The models were compared with each other, there forecasting is done in order to make inference about the future behavior.

Methodology

The data relating to the production of ragi in Odisha for the period from 1976-77 to 2019-20 are used here for the study. The data are collected from "Five Decades of Odisha Agricultural Statistics 2020" published by the Directorate Agriculture and Food production, Government of Odisha, 2020. The data of production is expressed in MT. The data from 1976-2016 is used for model building i.e. training data set and 2017-2020 is used for model validation i.e, testing data set.

Arima

Box and Jenkins developed an outstanding and simple technique to time series forecasting in the year 1976, and it became one of the most used approach until the late 90's. The future forecasted values are assumed to be a linear combination of past time series and past errors in this method of forecasting.

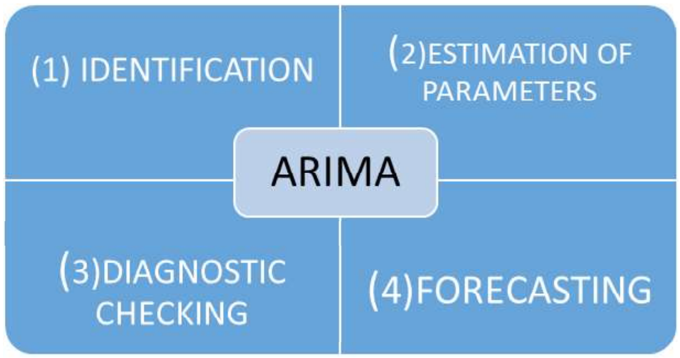
ARIMA stands for Autoregressive Integrated Moving Average models. It is one of the univariate timeseries forecasting model that projects the future values of a series, based on its own inertia. An ARIMA model is usually stated as ARIMA(p,d,q), where p shows the order of the autoregressive components, d shows the degree of differencing and q is the order of the moving average. The general form of ARIMA is :

$$X_t = \theta_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \ldots . - \theta_q e_{t-1} \quad (1)$$

Where, X_t= the actual time series data at time t

e_t= random error at time t (Dash and Mahapatra, 2020)

θ_j(j = 1, 2, 3, ... q) and φ_i (i = 1, 2, 3, ... p) are the model parameters, p shows the autoagressive and q shows the moving average terms are in polynomial. (Singh,2021).



ARIMA modelling is done under four stages. These are model identification, parameters estimation, diagnostic checking and forecasting.

EEMD

EMD (Empirical Mode Decomposition) is an adaptive time-space analysis method suitable that are non-stationary and non-linear. EMD parts the series into 'modes'(IMFs: Intrinsic Mode Functions) without leaving the time domain. EMD decomposition of the time series data can be presented by, $X(t) = \Sigma_{i=1}^{n} Yi(t) + r(t)$ (Debnath *et al*. 2013).

Where r(t) is the reminder term, $Y_i(t)$ is the components of IMF, its frequency ranks from high to low.

EMD faces one of the significant problem known as "mode mixing phenomenon" where the extreme points are distributed unevenly. To overcome this EEMD(Ensemble-EMD) is used to overcome this problem proposed by Wang *et al*. (2015). The principle involved in EMD and EEMD is almost similar. EEMD use white noise's statistical property which is distributed evenly, and adds it to the original signals changing the characteristics of the original signals extreme points.

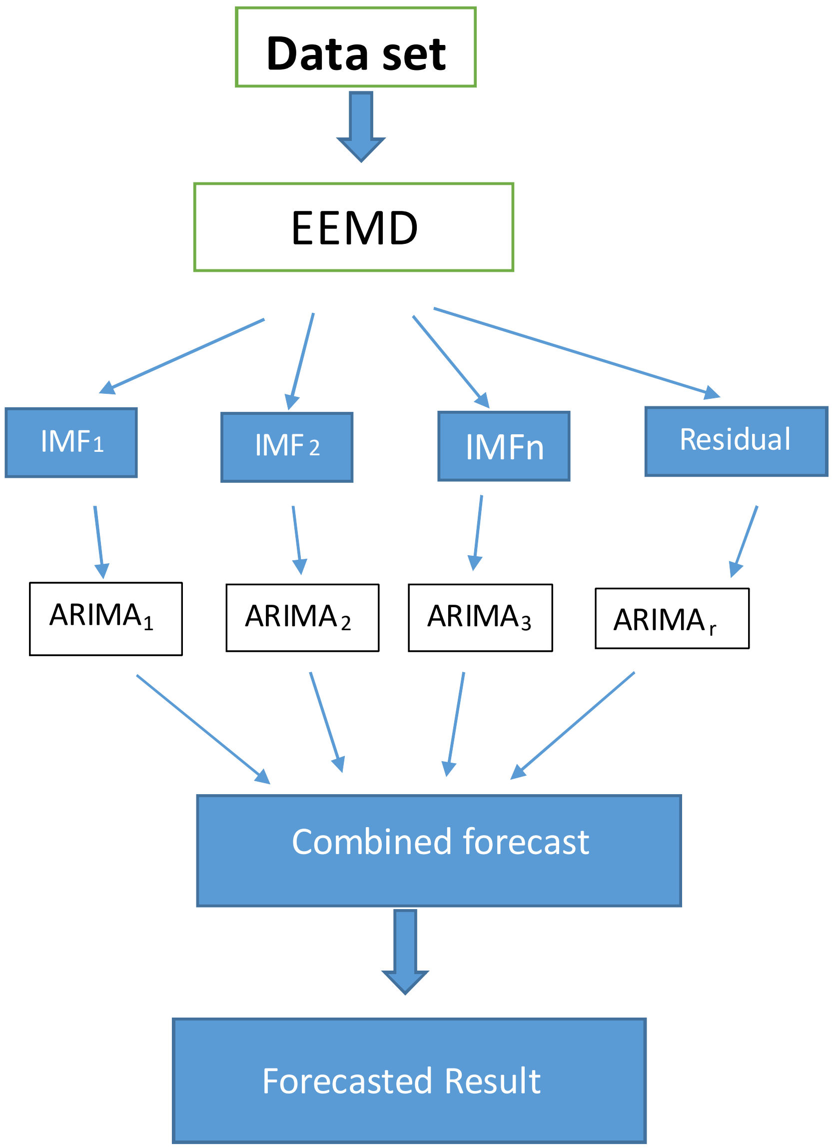
Steps of EEMD are as follows:

1. Define a noise signal Y(t).
2. Obtain X(t)= Y(t)+ White noise
3. Define the local extrema of Y(t) time series
4. Aggregate all the local maxima and the local minima to form the upper $e_{max}(t)$ and lower envelope $e_{min}(t)$
5. Calculate the average, $m_i(t)= (e_{max}(t)+ e_{max}(t))/2$
6. Calculate the difference $h_i(t)= X(t)- m_i(t)$
7. Check if $h_i(t)$ satisfy the stopping criteria condition (IMF is symmetric with zero mean and have the same number of zero crossing) if $h_i(t)$ satisfy then, $h_1(t)=IMF_1(t)$, other wise the first 4 steps are replicated until the stopping criteria is met

8. The residual is computed, $r(t)=X(t)- IMF_i(t)$
9. Consider r(t) is a new time series Y(t) until getting a constant residual rn(t).

Hybrid EEMD-ARIMA modelling Strategies

Work flow



(WORKING OF EEMD-ARIMA FORECASTING)

As seen from the work flow of EEMD-ARIMA, the main steps involved in the forecasting are listed below

(1) Using EEMD approach, the original timeseries data x(t)(t=1,2,3..n) are decomposed into n numbers of IMFs and one residual component.

(2) Then the ARIMA model is utilized as a forecasting tool for each extracted IMF component and residual component to model the decomposed IMFs and residual components and produce the appropriate prediction for each component.

(3) Finally, the forecasting results of all extracted IMFs and residual components acquired by ARIMA are aggregated to produce an aggregated output that may be implemented as the final forecasting result for the original time series data.

In its simplest form, the hybrid EEMD-ARIMA forecasting model combines the decomposition and ensemble principles. The decomposition streamlines the forecasting work, where as the ensemble produces a consensus forecast based on the original dataset. (Singh, 2021)

Evaluation Methods for Forecasting Performance

The most popular forecasting evaluation methods like root mean squared error (RMSE), Mean Absolute Error(MAE) and Mean Absolute Percentage Error (MAPE) were used to evaluate above models.

- **MAPE**(Mean absolute percentage Error)
  Here, the error is measured in percentage terms.

$$\text{MAPE}=\frac{\dfrac{\sum_t \left|P_t - A_t\right|}{A_t}}{n} * 100$$ (Dash, *et al*. 2017)

- **MAE**(Mean absolute error)
  AME measures the average magnitude of the errors in a set of predictions without considering their directions.

$$AME = \frac{\Sigma \,|\, P_t - A_t \,|}{n}$$

- **RMSE**(Root mean square error):

$$\text{RMSE} = \sqrt{\frac{\Sigma (P_t - A_t)^2}{n}}$$ (Dash *et al*. 2020)

Where, P_t = Predicted Value
A_t = Actual value
n = no. of observations

Results and Discussion

The necessary data were gathered, analysed and interpreted in the setting of the study's objective. The acquired results are presented and discussed further below.

Fig. 1. Shows the time series plot of production of Ragi in Odisha from 1976-2020.

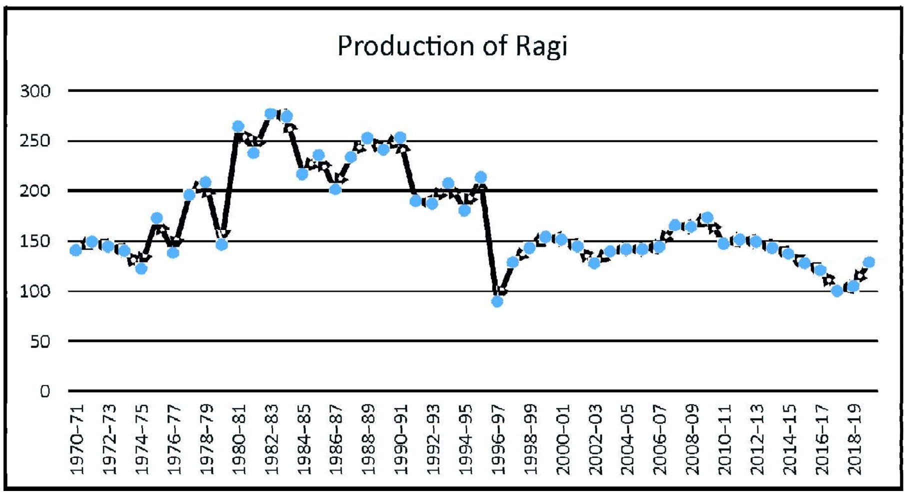


Figure 1: Time series plot of Ragi

ARIMA model was fitted to the production timeseries of Ragi and it was found that ARIMA (1,1,0) was found to be the best. The whole data set is divide into 2 parts period, one is the training data set i.e. 1976-2015 and the other one is the testing data set i.e. 2016-20. Here 2 models are taken for the analysis and forecasting of the data, the first one is Box-Jenkins ARIMA model and it is compared with the hybrid EEMD-ARIMA.

Table 1: Forecasted values of Ragi production with 95% confidence level

Production of Ragi(in MT)			
Year	*Forecasted values (ARIMA)*	*95LCL*	*95UCL*
2020-21	117.4551	56.9749	177.9353
2021-22	122.7949	54.4444	191.1439
2022-23	120.2659	38.2106	202.3212
2023-24	121.4631	30.6196	212.3067
2024-25	120.8962	20.8044	220.9880
2025-26	121.1647	13.1590	229.1703

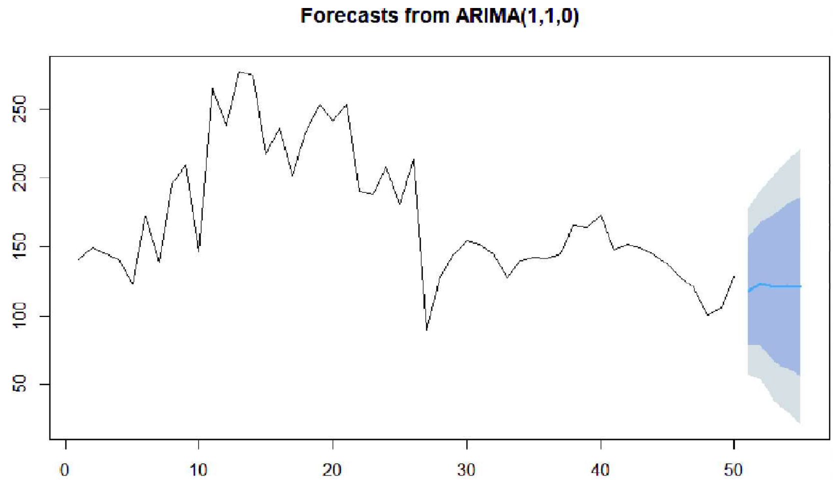


Figure 2: Forecasts of Ragi Production using ARIMA

Table 1 shows the projected values of Ragi production using ARIMA (1,1,0). According to the table the Ragi production will rise in the subsequent year. In the year 2025-26, it will rise to 121.16 MT. Along with the prediction values, lower and higher confidence values at 95% are displayed in this table. In the second phase of the experiment the hybrid EEMD-ARIMA model was applied to the timeseries data of the Ragi production data. Figure 3 shows that the total timeseries data is decomposed into 4 IMFs and ARIMA model is applied to the individual IMFs. Then the result is combined and presented in table number 2.

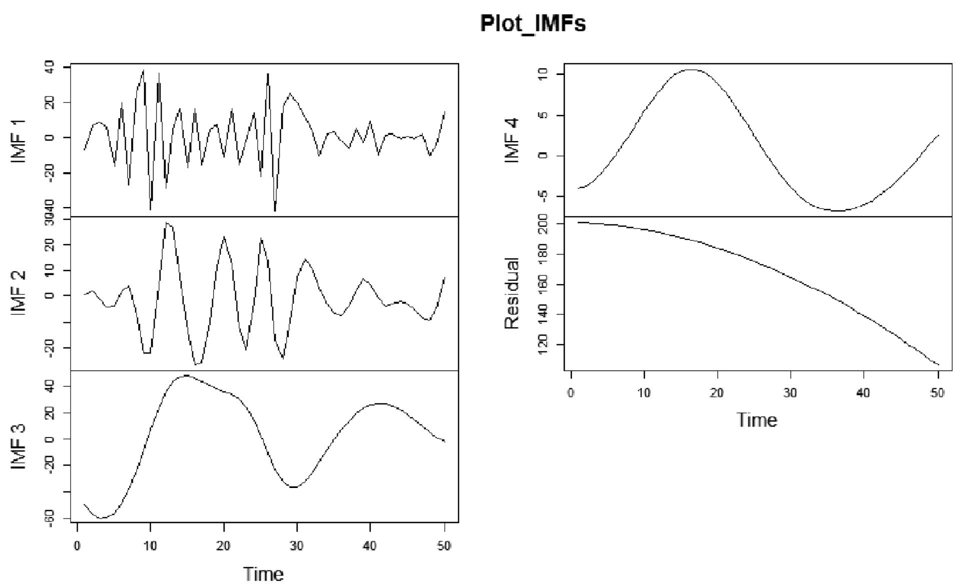


Figure 3: IMFs of the EEMD-ARIMA model

Table 2: Production forecast of Ragi with EEMD-ARIMA model

Year	*Forecasted values (EEMD-ARIMA)*
2020-21	149.6274
2021-22	151.1844
2022-23	145.1607
2023-24	149.2646
2024-25	148.5095
2025-26	145.5095

As a result we can see, the production of Ragi will increase up to 151.1844 MT in coming year and we can see it will have a decreasing trend in coming years. The Ragi production will be at 145.5095 in 2025-26, according to the EEMD-ARIMA model forecast. Production may increase as a result of the application of inputs and effective packages of practices, but production will remain within the anticipated limitation.

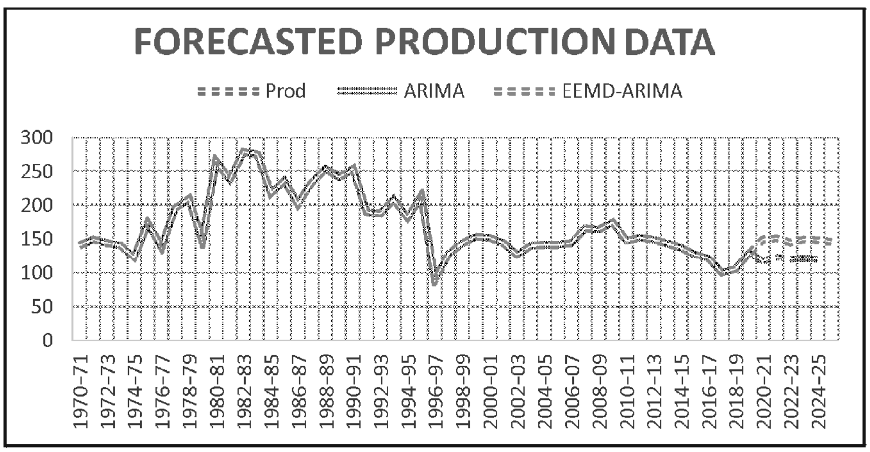


Figure 3: Forecasts of Ragi Production

A comparative analysis was done between the ARIMA and EEMD-ARIMA models were done to show techniques with best results based on the forecasting accuracy measures including Mean Absolute Error(MAE), Root Mean Square Error(RMSE), and Mean Absolute Percentage Error(MAPE).

Table 3
Forecasting performance of the models for the production of Ragi in Odisha

Models	*RMSE*	*MAPE*	*MAE*
ARIMA	31.98	13.56	22.74
EEMD-ARIMA	16.76	0.109	12.53

As we can see in the above model the RMSE,MAPE AND MAE value for the hybrid EEMD-ARIMA model is comparatively less compared to the ARIMA model, making the hybrid EEMD-ARIMA model more suitable for forecasting.

Conclusion

The study has suggested that the hybrid model of EEMD-ARIMA is the best model for the production projection of Ragi. The hybrid method divided the non-stationary and non-linear data into modes or IMFs, then the each component uses the ARIMA model. This method can be an effective way to improve forecasting performance. Based on the results obtained in this work one can say a hybrid EEMD-ARIMA model can increase the forecasting accuracy, which will help the government, farmers and policy makers to adopt new technology for a better status of the crop.

References

[1] Agricultural Statistics at a glance,2020, Government Of Odisha, Department of Agriculture.

[2] Dash A, Dhakre DS and Bhattacharya D (2017) Study of growth and Instability in Food grain Production of Odisha: A Statistical Modelling Approach. *Environment and Ecology*, 35(4D): 3341-3351.

[3] Dash Abhiram and Mahapatra Subrat Kumar (2020). Using ARIMA Model for forecasting of important pulse crops of Odisha, India. *Amazonian Journal of Plant Research*, , 4(3): 646-659.

[4] Dash Abhiram, Mangaraju A, Mishra Pradeep and Nayak, H (2020). Using Autoregression Integrated Moving Average (ARIMA) technique to forecast the production of Kharif cereals in Odisha (India). *Current Journal of Applied Science and Technology*, 39(9):104-113.

[5] Debnath MK, Bera K, Mishra P (2013) Forecasting area, pro duction and yield of cotton in India using ARIMA model, *Research & Reviews: Journal of Space Science & Technology*: 2(1):16-20.

[6] Naveena, K., Subedar Singh, Santosha Rathod and Abhishek Singh. 2017. Hybrid ARIMA-ANN Modelling for Forecasting the Price of Robusta Coffee in India. *Int. J. Curr. Microbiol. App. Sci.* 6(7):17211726.doi:https://doi.org/10.20546/ijcmas.2017.607.207

[7] Singh. Abhishek (2021). Comparison of artificial neural networks and statistical methods for forecasting prices of different edible oils in Indian markets. *International Research Journal of Modernization in Engineering Technology and Science*, 3(10): 1044-1050

[8] Wang, WC., Chau, KW., Xu, DM. and Chen XY. (2015). Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition. *Water Resource Management* 29(8), 2655–2675.