International Journal of Agriculture, Environment and Biotechnology
Citation: *IJAEB*: **14**(1): 51-62, March 2021
DOI: 10.30954/0974-1712.01.2021.6

ASSOCIATION FOR AGRICULTURE
ENVIRONMENT AND BIOTECHNOLOGY

AGRICULTURAL STATISTICS

Forecasting of Rice Production using the Meteorological Factor in Major States in India and its Role in Food Security

P. Mishra[1], P.K. Sahu[2], Monika Devi[3], Chellai Fatih[4] and A.J. Williams[5]

[1]College of Agriculture, JNKVV, Powarkheda, Hoshangabad, Madhya Pradesh, India
[2]Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia West Bengal, India
[3]Department of Mathematics & Statistics, CCSHAU, Hisar, Haryana, India
[4]Department of Based Education, University of Ferhat Abbas, Algeria
[5]BTC College of Agriculture and Research Station, IGKV, Bilaspur, Chhattisgarh, India

*Corresponding author: pradeepjnkvv@gmail.com (**ORCID ID:** 0000-0003-4430-886X)

Paper No. 881 **Received:** 11-01-2021 **Revised:** 27-02-2021 **Accepted:** 13-03-2021

ABSTRACT

The world as well as in India, rice is playing a major role in food security. Production factors (like rainfall, minimum temperature, fertilizer consumption, an area under irrigation for a particular crop) are very crucial for crop productivity. Forecasting is always important for policy implication and planning purposes of the country. In the present investigation, the projection has been made using simple ARIMA and ARIMAX (with the inclusion of crop inputs in ARIMA models). In terms of less error in model and projection, wise ARIMAX model was found better compared to simple ARIMA. In this present study, forecasting has been attempted with the inclusion of meteorological factors using ARIMA modeling up to the year 2022. This study reveals the future trend of rice production as well as a factor affecting productivity. Among the major states, West Bengal would lead the state in India in rice production, with a productivity of 4758 kg/ha, while Punjab will be the leader in productivity in the year 2022. This prediction would be helpful for policy implication and food security of the country.

Highlights

- By and large, there has been an expansion in the area, production, and yield of rice in major growing states, including the whole of India. In the majority of the states, including the whole of India, the average rainfall has increased during the study period.
- In rice area, production, and productivity data, none of the series is stationary; Ist differencing with original data makes all the series stationary.
- Inclusions of different factors of productions in the best fitted time series model increase the accuracy and forecasting power compared to the simple ARIMA model.

Keywords: Modelling, forecasting, rainfall, temperature, food security, ARIMAX

World population, in general, and the population of developing countries, in particular, has experienced a sharp increase in recent decades. Provide food to these increasing human populations tough task to the planners of the individual countries and also the world bodies. The vast majority of rice producers are in Asia, which is the emblematic cereal. There are 200 million rice farms on our planet, and 50% of the world's population depends on it for work or consumption. Agriculture in India is an important economic sector of India and one of the most important agricultural productions on the planet. The sector employs almost two-thirds of the workforce (49% of jobs in 2012), out of a total population of 1.1 billion, of which 72% are rural.

How to cite this article: Mishra, P., Sahu, P.K., Devi, M., Fatih, C. and Williams, A.J. 2021. Forecasting of Rice Production using the Meteorological Factor in Major States in India and its Role in Food Security. *IJAEB*, **14**(1): 51-62.

Source of Support: None; **Conflict of Interest:** None

Agriculture contributes approximately 17-18 percent to the country's GDP. Rice belongs to a family of starchy foods that provide vegetable protein and contain complex carbohydrates. These are gradually released into the body and provide energy as and when needed (Ramesh 2020). Nutritionists recommend increasing carbohydrate intake to contribute more than 50% of daily energy intake. An increase in the consumption of starchy foods (complex carbohydrates) is therefore recommended. (Khush 2003). Crop yield is influenced by several meteorological factors. Weather factors play a vital role in production of crops, and if weather factors (viz. rainfall, maximum temperature, minimum temperature, etc.) are not favorable to the crop, then the crop cannot be grown well. The best climatic conditions for growing rice are heat and heavy rainfall (monsoon: rain for several days); any modification of these two factors could affect the productivity of these crops (Kaul 2006 & Raiger *et al.* 2019). Therefore, studying and analyzing the effects of associated metrological factors is also necessary for effective crop prediction. To identify important factors affecting crop diversification, and by using time series data, multiple regression analysis was performed (Kebebe 2000; Joshi *et al.* 2004; 2006). Mishra *et al.* (2015) forecasted the wheat as well as total food grain production using meteorological factors likes (Rainfall & temperature). Biswas *et al.* (2018) studied the impact of temperature on rice yield on *kharif* season. With this backdrop, the present study attempts to examine the trend of factors of production on rice in major growing states of India, modeling, and projecting of rice production in these states along with the whole of India taking in to account the weather factors which are important in the production process. The present investigation has following Aims.

To see the nature of methodological factors (Like, Rainfall, Minimum temperature, maximum temperature) and other factors of production (like Fertilizer use, Irrigation area, and pesticide consumption).

To compare the model between ARIMA (Without factors) and ARIMAX (With factors) used for forecasting.

To show the effect of factors (like Rainfall, temperature, Irrigation, Fertilizer consumption) in productivity and use in projection.

MATERIALS AND METHODS

Selection of study area

The following three states: West Bengal, Uttar Pradesh, and Punjab, are contributing around 37% of total production in India from 33% area under rice in India during 2017-18. In order to select the major growing states in India for rice, production, percentage share of each and every state are examined. The information is presented in table 1.

Table 1: Contribution of different states towards rice production ('000 tons) in India (2017)

State	Rice	% to India
Andhra Pradesh	8051.30	7.25
Assam	5158.00	4.65
Bihar	7296.40	6.57
Chhattisgarh	6910.60	6.23
Gujarat	1762.00	1.59
Haryana	3946.80	3.56
Himachal Pradesh	130.50	0.12
Jammu and Kashmir	596.30	0.54
Jharkhand	3775.20	3.40
Karnataka	2359.00	2.13
Kerala	418.70	0.38
Madhya Pradesh	3908.40	3.52
Maharashtra	2660.50	2.40
Odisha	7619.00	6.86
Punjab	12283.30	11.07
Rajasthan	428.80	0.39
Tamil Nadu	6395.90	5.76
Telangana	5825.00	5.25
Uttar Pradesh	13345.90	12.02
Uttarakhand	663.00	0.60
West Bengal	14990.00	13.50
Others	2483.00	2.24
All India	111007.80	100.00

Source: *Ministry of Agriculture and Cooperation, Govt. of India.*

Thus, these three states have got tremendous importance in all India's production of rice vis-à-vis food and dietary security of the nation. As such, this study has taken these three states along with that of the whole of India. The data have been obtained for the area, production, productivity of rice, as well as a factor effecting of rice production like weather factors, fertilizer, and pesticide use from different sources for the period as provided below.

Information on	Period
Area, Production and yield of rice	1950-2009
Average rainfall (mm)	1961-2009
Average maximum temperature	1961-2009
Average minimum temperature	1961-2009
Crops wise irrigation	1965-2009
Fertilizer consumption	1975-2009
Pesticide consumption	1989-2009

Source of data

The information on production, yield of rice, area in the above major states and India for the period 1950-2010 was collected from www.indiastat.com and Agriculture Statistics at a glance. The information on pesticide consumption could be collated for the period from 1989 to 2010 and that of fertilizer consumption from 1975-2010 from the book of fertilizer statistics 1978-2005. Meteorological data could be obtained for 1961-2010 from the website www.indiawaterportal.com. Weather data is compiled into year data with the help of monthly data. Information on irrigation for different crops or different states were collated from different issues of "Agriculture at glance" Anonymous. (2017).

Summary Statistics: Study of historical or time-series data is a challenging task to a researcher in terms of data handling and data analysis. Present investigation summary statistics have been worked out for factors of production. Generally, analytical tool starts with description and summarization of the information collected in any study. As such, this study also describe the characteristics of each series via the central tendency (average arithmetic, percentages ...) and the measurements of dispersion (standard deviation, variance, ...).

As one of the objectives of the present study is to forecast the parameters using ARIMA models and it is advised to have at least thirty data points for effective and successful modeling, the study prefers not to consider pesticide consumption data (having data point for 21 years only) for the purpose excepting for descriptive study.

Grubb's Test for detecting outliers

Time series data, in economic, financial fields are characterized by the presence of outlier. An outlier in this context is an observation, somewhat different from the rest of the observations in the series. There are several ways to detect outliers. Grubbs' test is one of such tests particularly suitable under a large set of data points, Grubb (1950, 1969, 1972). To use this test, we have to estimate the mean of observation $\underline{X}$, the standard error of the data set, and a distance D, between the value, suspected x_i of being an outlier and the mean value, $\frac{|x \quad \underline{X}|}{}$ then, we compare it with a distance limit D_{lim}, so, we reject the hypothesis of no outliers if,

$$D > D_{\lim} = \frac{N-1}{\sqrt{N}} \sqrt{\frac{\left(t_{\frac{\alpha}{2N},N-2}\right)^2}{N-2\left(t_{\frac{\alpha}{2N},N-2}\right)^2}}$$

With, $t_{\left(\frac{\alpha}{2N},N-2\right)}$ the critical value of the student distribution.

Test of Randomness

Test of randomness is an initial step to know the behaviour of any time series data as to whether the observations are following any pattern or changing randomly. This is to test:

$$\{H_0\text{: } X_t \textit{ are random.variable.iid}$$
$$H_1\text{: } X_t \textit{ are not random.variable.iid}\}$$

A time series X_t has a turning point at time t if one of the following two situations is true:

$X_t > X_{t-1}$ and $X_t > X_{t+1}$

Or $X_t < X_{t-1}$ and $X_t < X_{t+1}$

Let H be the number of turning points of a time series of size n. Under H_0, $H \sim N\left(\mu_H, \sigma_H^2\right)$ with:

$$\mu_H = \frac{2(n-2)}{3}, \sigma_H^2 = \frac{16(n-9)}{90}$$

If $\frac{|H-\mu_H|}{\sigma_H} < Z_{1-\frac{\alpha}{2}}$ we accept the null hypothesis; where $Z_{1-\frac{\alpha}{2}}$ is the quintile of order $\left(1-\frac{\alpha}{2}\right)$ of the standard normal distribution (Gouri 1984).

In reality, this test is used to test the *lack of autocorrelation*, and it is a non-parametric test. The parametric tests best known in the literature are those of Box and Pierce (1970) and Ljung and Box (1978).

Box–Jenkins models

The Box–Jenkins (1976) method is based on the principle of *parsimony* for building linear time-series models. Intriligator *et al.* (1996). Under this principle, in model construction, we employ the minimal possible number of parameters that mimics the properties of the sample. It is a methodology for the systematic study of time series from their characteristics in order to determine, in the family of *ARMA* models, the most suitable to represent the studied phenomenon. Technically, the Box–Jenkins method is an iterative cycle: Model Identification ⇆ Model Estimation ⇆ Model Validation ⟹ Forcasting.

The *identification* phase is the most important and the most difficult: it consists in determining the appropriate model in the family of *ARMA* models. It is based on the study of simple and partial *correlograms*. At the *estimation* stage, different methods can be used: (Least Squares Method, Yule-Walker Equations, Maximum of Likelihood Estimation...). Then *validation* (diagnostic checks) is applied to establish if the chosen model adequately represents the given data set. If the model is well validated, we past to the forecasting process. If any inadequacy is found, we return to the identification phase until the best representation is found.

Autoregressive model

In an autoregressive process, *AR*(p) of order, p the present observation X_t is generated by a weighted average of past observations until the p^{th} period, *AR*(p) is written as follow:

$$X_t = c + \sum_{i=1}^{p} \theta_i X_{t-i} + \varepsilon_t, \theta_i \in R, \varepsilon_t \leadsto N(0, \delta_\varepsilon).$$

Where θ_1, θ_2, … θ_p model of the parameter to be estimated, c is a stable and ε_t: is *i.i.d* with mean zero. The constant c does not change the stochastic properties of the process; generally we normalize it to be $c = 1$.

Moving Average model: (sub-section)

The moving average process, with notation *MA*(q) refers to series of order q, which suppose that an observation X_t is generated by a weighted average of noise ε_t until the q^{th} period, we represent it by:

- *MA*(1): $X_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1}$.
- *MA*(2): $X_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2}$ with: $E(\varepsilon_t) = \sigma_\varepsilon^2$
- …
- *MA*(q): $X_t = \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \ldots - \alpha_q \varepsilon_{t-q}$

As a general presentation of such process, we reformulate it as:

$$X_t = \mu + \sum_{i=1}^{q} \alpha_i \in_{t-1}$$

Where α_1, α_2, …, α_q the are the parameters of the process (q), μ is the expectation of X_t, and the ϵ_t, ϵ_{t-1}, ϵ_{t-2},…, ϵ_{t-q} error terms. ϵ_0 is often normalized to be equal to 1.

Autoregressive Moving Average Process (ARMA)

We can combine the two previous models *AR*(p), *MA*(q), to obtain a mixed autoregressive moving average model of order p and q, *ARMA*(p,q):

$$(L) X_t = \alpha(L) \in_t$$
$$(L) = \theta_0 + \theta_1 L + \ldots + \theta_p L^p$$
$$(L) = \alpha_0 + \alpha_1 L + \ldots + \alpha_q L^q$$

Where: $\theta_0 = \alpha_0 = 1$, and $\upsilon(\in_t) = \sigma_\varepsilon^2$ and $E(\in_S \in_t) = 0,\ t \neq S$.

So, *ARMA* an model represents a process generated by a grouping of the historical values of X_t and the past errors ϵ_t, (or disturbances). The models: *AR*(p), *MA*(q), *ARMA*(p,q), are representative only of time series that are: stationary in tendency; seasonally adjusted (Mishra *et al.* 2021 & 2020).

ARIMAX Model *or* Transfer Function Model

In practice, we fit a relationship between a dependent variable X_t (or output) and an independent variable Z_t, (or input); also, when estimating such relationship, in reality, the impact of change in Z_t may not appear (realized) instantly, but takes time (*e.g.*, effect on climate changes on agriculture production), this impact is called a *dynamic response*, and the model that takes into account this dynamic effect is called a *transfer function model.* Intriligator *et al.* (1996). We can formulate this relationship as:

$$X_t = \gamma_0 Z_t + \gamma_1 Z_{t-1} + \gamma_2 Z_{t-2} + \ldots = \gamma(L) Z_t$$

Generally, $\gamma(L) = \frac{\beta(L)}{\varphi(L)}$ we put:**,** this transformation allows the dynamic relation between X_t and Z_t to be represented as:

$$\varphi(L) X_t = \beta(L) Z_t$$

We add ϵ_t a noise to the relationship (which is affected by noise, omitted variables...) between Z_t and X_t:

$$X_t = \gamma(L) Z_t + \mu_t$$

We assumed ϵ_t follow an *ARMA(p,q)* process:

$$\theta_t = \frac{\alpha(L)}{\theta(L)} \in_t$$

The final model of ARIMAX is a combination of: a system part and an ARMA noise model,

$$X_t = \frac{\beta(L)}{\varphi(L)} Z_t + \frac{\alpha(L)}{\theta(L)} \in_t .$$

The difference between a transfer function model and an ARIMA model is that in the former case, the input variables (*exogenous*) Z_t is observable, while in the latter case, the innovation (or error) ϵ_t is not observable. The main interest (advantage) of ARIMAX model is to use it to forecast output variable X_t from future values of the input variables Z_t.

Checking for model adequacy

Among the competitive Box- Jenkins model best model is selected on the basis of maximum R^2, minimum root mean square error (RMSE), minimum mean absolute percentage error (MAPE), minimum of maximum average percentage error (MaxAPE), minimum of maximum absolute error (MaxAE), minimum of Normalized BIC. Any model which has fulfilled most of the above criteria is selected. This part of the study provides some goodness-of-fit tests and measures used in econometric modeling.

Coefficient of determination R^2: is the proportion of the total of the dependent variable Y_t that is explained by the regression; it has different expressions:

$$R^2 = \frac{\sum_t \left(\hat{Y}_t - \underline{Y}\right)^2}{\sum_t \left(Y_t - \underline{Y}\right)^2} = 1 - \frac{\sum_t e_t^2}{\sum_t \left(Y_t - \underline{Y}\right)^2}$$

There is an inverse relationship between the sum of squared errors Σ_t and e_t^2 the value of R^2**,** under some conditions R^2, provide us is an estimation of the explanatory power of the estimated model.

Root Mean Squared Error, **(RMSE):** is a measure characterizing the "precision" of the estimated model.

The best model is simply the one with the smallest value of *RMSE*.

$$RMSE = \sqrt{\frac{\sum_t \left(Y_t - \hat{Y}_t\right)^2}{n}} = \sqrt{\frac{\sum_t e_t^2}{n}}$$

Mean Absolute Percentage Error, (**MAPE):** also called: mean absolute percentage deviation (**MAPD**), allows us to measure the accuracy of a statistical forecasting method. The accuracy under this measure is expressed as a percentage:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|$$

Mean absolute error, (MAE). In time series analysis, mean absolute error (MAE) is a measure of forecast error, which is the difference the observed data set Y_i and the regression part $\hat{Y}_i$:

$$MAE = \frac{\sum_{i=1}^{n} \left|Y_i - \hat{Y}_i\right|}{n} = \frac{\sum_{i=1}^{n} | e_i |}{n}$$

Maximum Absolute Percentage Error. (*MaxAPE*). Is a measure of the extreme forecasted errors. These statistics give us an idea about the outlier among the forecasted errors: $\hat{e}_i : i : 1, 2, \ldots n.$

$$MaxAPE = 100\%.Max\left(\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|\right).$$

Maximum Absolute Error, (*MaxAE*). At contrast of, maximum absolute error, is expressed in the same unit of measure as the data set Y_t, the *MaxAE*, estimate (determine) the maximum of the forecasted errors:

$$MaxAE = Max\left(\left|Y_i - \hat{Y}_i\right|\right) = Max\,|\,e_i\,|\,.$$

Normalized Bayesian Information Criterion: Model selection is a well-known problem in statistics, and information theory provides a rigorous framework for the development of efficient estimators. Among the answers to this problem: is the minimization of a penalized criterion. The first criteria that appear in the literature are Akaike Information Criterion (*AIC*, Akaike (1973), Bayesian Information Criterion (*BIC*, Schwaz (1978).

$$BIC = -2\ ln\ ln\ (L) + k.\ ln\ ln\ (N).$$

To obtain the normalized *BIC*, we just divide the *BIC* by 2 *N*,

$$NBIC = -\frac{1}{N}\ln\ln\left(L\right) + \frac{k}{2N}.\ln\ln\left(N\right).$$

For *N* large, the second term converges to zero. The model that will be selected is the one that minimizes the BIC criterion.

RESULTS AND DISCUSSION

Weather factors have a very crucial role in rice plant growth. So, descriptive statistics tell the story of these factors of production. From table 1, one can see the pattern of climatologically factors (rainfall, maximum and minimum temperature) and factors affecting the production. The average annual rainfall in West Bengal is estimated at 1523 mm, with the highest being at 2329 mm, while the minimum annual rainfall is 961 mm. There has been a tendency of increased rainfall over the years as depicted by the simple growth rate of 16% during 1961 -2009. But in comparison to the growth rate of whole India (38%) it is far below. The positive skewness and kurtosis values reveal that during the initial period under study, the change has taken place. The mean annual rainfall of Uttar Pradesh is 894 mm, with maximum annual rainfall (1260 mm) received during 1980 that is almost 25% higher than the mean. While the annual minimum rainfall of 534 mm was recorded during 2001. The state has achieved a simple growth rate of 371 % during the period. The ß$_1$ value (-0.2) of coefficient of skewness suggests maximum asymmetry of data during the later part of the period under investigation. The value of ß$_2$ coefficient of kurtosis is near about (- 0.8) during the study. Similarly, in Punjab, the annual rainfall has varied by three folds, from a mere 290 mm to 785 mm during the period under study. During the period under study, an overall growth rate of more than 38% in average annual rainfall for the whole of India is recorded.

The average maximum temperature of West Bengal was highest at 32.7 ^{o}C and the lowest at 29.4 $^{\circ}$C, registering a change of almost 3.4 %. The ß$_1$ value (0.8) of coefficient of skewness suggests the asymmetric nature of data during the study period under investigation. In Uttar Pradesh average maximum temperature (32.9 ^{o}C) is received during 1977 which was almost equal to the mean maximum temperature and registering virtually no growth during the period under consideration. In Punjab, the average maximum temperature is similar to the whole of India during the period under investigation. The highest average maximum temperature for the whole of India was 32.7 ^{o}C, and that of the lowest average maximum temperature was at 30.9 ^{o}C with an average of 31.7 ^{o}C. In Punjab, the average minimum temperature varies between 16 to 18 ^{o}C during the period with an average of 17.1 ^{o}C and virtually registering no growth rate (only). The estimated average of the irrigated area under rice in West Bengal is 3011 × 10^3 hectare. The highest irrigated area under rice was 615 × 10^4 the hectare, while the lowest was at 1251 thousand hectares. West Bengal is the state which has a registered maximum growth rate (160%). In Uttar Pradesh, the irrigated area under rice has the highest at 608×10^4 hectare in 1999 and lowest at 633×10^3 hectare in 1964; recording a growth rate of only 1.01%. In Punjab, irrigation for rice recorded the highest rate in 2008 at 2734×10^3 hectare and the lowest rate in 1965 at 240×10^3 hectare. For the whole India, irrigation under rice has increased from 12146×10^3 hectare to 25465×10^3 hectare registering a growth of almost 100%. Expansion in irrigation potential is most warranted in augmenting rice production, particularly under high input-intensive production regime.

The mean total fertilizer consumption of Uttar Pradesh is estimated at 2059 thousand tonnes.

Table 1: *Per se* performance of climatological and some input factors of production of rice

	West Bengal						**Uttar Pradesh**					
	ARF	**Tmax.**	**Tmin.**	**IA**	**TFC**	**PC**	**ARF**	**Tmax.**	**Tmin.**	**IA**	**TFC**	**PC**
Minimum	961	29.5	19.4	1251.0	129.7	3000.0	534.3	31.3	17.9	633.0	223.1	1855.0
Maximum	2329.526	32.8	21.3	6150.0	1519.3	4785.0	1260.8	33.0	19.7	6080.0	4032.8	9563.0
Mean	1523.154	30.8	20.4	3011.6	783.8	3937.3	894.7	32.2	18.9	2951.0	2059.8	7533.9
S.E.	38.728	0.1	0.1	312.9	74.0	109.9	26.7	0.1	0.1	325.3	183.5	356.9
Kurtosis	0.772	6.9	0.4	-1.6	-1.4	-0.5	-0.8	-0.8	0.6	-1.6	-0.8	8.4
Skewness	0.619	0.9	-0.2	0.6	-0.2	-0.2	-0.2	0.0	-0.4	0.4	-0.1	-2.4
SGR(%)	16.522	3.5	-2.0	160.5	364.4	1.8	371.4	0.2	2.1	1.0	248.6	1.5
	Punjab						**India**					
Minimum	290.8	30.9	16.3	240.0	311.3	5610.0	848.9	30.9	16.3	12146.0	2080.0	35674.0
Maximum	785.4	32.8	18.0	2734.0	2100.6	7400.0	1314.2	32.8	18.0	25465.0	24909.3	75033.0
Mean	512.9	31.8	17.1	1574.4	1280.1	6694.6	1100.5	31.7	17.1	18062.9	11386.0	53436.5
S.E.	19.2	0.1	0.1	131.7	81.9	126.6	15.2	0.1	0.1	594.7	1119.0	2742.6
Kurtosis	-0.9	-0.6	-0.3	-1.4	-0.5	-1.0	-0.6	-0.5	-0.4	-1.2	-1.2	-1.1
Skewness	0.1	0.2	0.4	-0.3	-0.6	-0.6	-0.4	0.2	0.3	0.4	0.4	0.5
SGR(%)	457.2	11.7	0.1	-0.1	159.1	-0.8	38.7	-4.5	0.1	0.3	254.1	-7.8

Annual rainfall (ARF) in mm, Average maximum and minimum temperature (T_{max} and $T_{min.}$) in ^{0}C, Irrigated area(IA) in '000 ha, Total fertilizer consumption (TFC) in '000 tonnes, Pesticide consumption (PC) in MT (Metric tonnes).

Maximum (4032 thousand tonnes) fertilizer consumption (is recorded during 2008). Punjab is the state which has registered a minimum growth rate (159%) in average total fertilizer consumption during the period. For whole India total fertilizer consumption has an average of 11386 thousand tones, with a minimum amount of 2080 thousand tones and maximum of 24909 thousand tones. Thus, from the study of fertilizer consumption, it is clear that there has been tremendous growth in the consumption of fertilizer during the period under study. Crop protection is one of the important aspects in augmenting food grain production, as such use of pesticides plays vital role in this aspect. West Bengal is the state which has registered a minimum growth rate (1.7%) in average pesticide consumption during the period (1975-2009). In West Bengal average total pesticide consumption is estimated at 3937 MT with a standard error of 109 MT. In Uttar Pradesh pesticide consumption was highest at 9563 MT in 2008 and lowest at 1855 MT in 2003. Thus there is a great fluctuation in the use of pesticides. Similarly in Punjab maximum pesticide consumption was at 7400 MT during 1993 with a minimum of 5610MT. Whole India has doubled pesticide consumption; from mere 35674 MT to 75033 MT during the period under study.

As discussed material and method section each any every series under the climatologically and input factors are subjected to test of outlier and randomness. Table (2) presents the results of the tests. The data further revealed that in some of series one can detect the outliers which were subsequently replaced by median of the respective series before further analysis. From the test of randomness one can see that annual rainfall in all the major states have changed randomly except for whole India. Thus, year to year regional changes in annual rainfall has made agriculture more and more dependent on assured irrigation. In, West Bengal the average maximum temperature has changed randomly. Also the average minimum temperatures, excluding whole India, in other major states are random in nature. Irrigated areas under rice, other major states including whole India have followed definite trends. This is definitely the indication of the clear cut policy and its implication with respect to the expansion of the irrigation potential of the country. Only in West Bengal total fertilizer consumption has changed randomly over the period. The results show the randomness nature of pesticide consumption in West Bengal, Punjab and whole India, in contrast to definite trends for pesticide consumption in Uttar Pradesh. From the above results it is quite clear that during the period under consideration, the rice production in major states, as well as in India have remained on the mercy of the weather parameters.

Table 2: Test of outliers and randomness for climatological and some input factors of rice

	ARF	Tmax.	Tmin.	IA	TFC	PC	ARF	Tmax.	Tmin.	IA	TFC	PC
	West Bengal						**Uttar Pradesh**					
No. of observation	50	50	50	44	34	20	50	50	50	44	34	20
No. of turning point (p)	29	27	33	12	18	9	31	26	28	18	13	8
E (P)	32	32	32	28	21.3	12	32	32	32	28	21.3	12
V(P)	8.5	8.5	8.5	7.5	5.7	3.2	8.5	8.5	8.5	7.5	5.7	3.2
$^{\tau}$*Cal*	-1.0	-1.7	0.3	-5.8	-1.3	-1.6	-0.3	-2.0	-1.3	-3.6	-3.4	-2.2
Inference	R	R	R	T	R	R	R	T	R	T	T	T
Outliers test	No	Yes	No	No	No	No	No	No	No	No	No	Yes
	Punjab						**India**					
No. of observation	50	50	50	44	34	20	50	50	50	44	34	20
No. of turning point (p)	34	24	29	6	12	9	25	25	26	14	14	9
E (P)	32	32	32	28	21.3	12	32	32	32	28	21.3	12
V(P)	8.5	8.5	8.5	7.5	5.7	3.2	8.5	8.5	8.5	7.5	5.7	3.2
$^{\tau}$*Cal*	0.68	-2.7	-1.0	-8.0	-3.9	-1.6	-2.3	-2.3	-2.0	-5.1	-3.0	-1.6
Inference	R	T	R	T	T	R	T	T	T	T	T	R
Outliers test	No	No	No	No	No	No	No	No	No	No	No	No

Annual rainfall (ARF) in MM, Average maximum and minimum temperature (T_{max} and $T_{min.}$) in ^{0}C, Irrigated area(IA) in '000 ha, Total fertilizer consumption (TFC) in '000 tonnes, Pesticide consumption (PC) in MT (Metric tonnes), R=Random, T =Trend.

Table 3: Correlation of factors related with productivity of rice in major states of India

	West Bengal	Uttar Pradesh	Punjab	India
Irrigated area	0.911	0.926	0.92	0.939
	<.0001	*<.0001*	*<.0001*	*<.0001*
Annual rainfall	0.213	-0.296	-0.115	0.329
	0.227	*0.09*	*0.517*	*0.057*
Tmax	0.417	0.476	0.252	0.262
	0.014	*0.004*	*0.151*	*0.134*
Tmin	0.139	0.349	0.223	0.302
	0.433	*0.043*	*0.205*	*0.082*
Total fertilizer consumption	0.979	0.744	0.881	0.879
	<.0001	*<.0001*	*<.0001*	*<.0001*

Italicised letters indicate the significance levels of the corresponding correlation coefficients.

At the time of setting objective for present study, we assumed that the factors (rainfall, maximum and minimum temperature, net irrigated area of crop and total fertilizer consumption) are supposed to have a great role in productivity of rice. In this section attempts have been made to work out the degree of linear associationship and linear relationship among these parameters for rice crop. Correlation, analysis is taken up to find out the extent and actual linear relationship of productivity on these parameters. In all major rice producing states including whole India (table 3) it is found that rice yield is significantly and positively correlated with net irrigated area of rice and total fertilizer consumption. Thus, fertilizer use and irrigation are the two major contributing factors related to rice yield vis-a-vis rice production there was no significant effect of rainfall, average maximum and minimum temperature on rice productivity. This may be due to the fact that rice production in major producing states might have overcome the dependence on monsoon activities rather the irrigation potential and other than *kharif* rice is playing major role in total rice production.

Our next task is to forecast the observation for a short-run time. For the purpose we adopted the Box –Jenkins methodology as discussed in the material and method section. Data over the period 1975-2009 has been used for model estimation, and the data for years 2010 to 2017, are used to validate the estimated models. Best fitted models, as adjudged through model accuracy criteria discussed in material and method section, are used to forecast step for the years to 2022.

Economic data are characterized by instability over time, which produce what we call *a no-stationary time series.* Statistically, a stationarization of the series is necessary; by using: (i) the differencing operation, (in practice, the level of differencing *d* = 2 is *two*) or (ii) transformation, especially in the case of series with a tendency *(called a models*). First order differencing was sufficient to make these stationary. It is found that from *ARIMA* (0,1,1) model to *ARIMA* (1,1,5) are the most appropriate models for modeling and forecasting the level of rice production. For increasing the accuracy and to outline the path of production, and the model building we incorporate different factors of productions in the selected models. The best model among ARIMAX is selected based on the minimum values of RMSE, MAE, MSE, MAPE, and maximum value of R^2, NBIC (Mishra *et al.* 2015). From the model validation period data for the period 2010-2017, it is clear that ARIMAX models are comparatively better than the corresponding ARIMA models; hence the importance of inclusion of auxiliary variables in the model. With the inclusion of factors, on the basis of forecasted value in 2009-10 production of whole India is 96534 thousand tonnes almost similar findings were reported by Ravichandan *et al.* (2012) in rice in India. Thus, there would be an increase in production as well as productivity in major growing states. It is clearly visible that in the case of production of rice West Bengal would be the leading state of India during 2022-23; West Bengal is supposed to produce more than seventeen million tonnes of paddy from an estimated area of slightly above of six million hectare with a per hectare production of around 2.9 quintal. Though Uttar Pradesh will contribute maximum area under rice in India because of its lower productivity it will produce 13.7 million tones of rice from 6.71 million hectare of land at a yield rate of 2.47q/ha. Punjab is another major paddy producing state will continue

Table 4: ARIMA models for area, production and yield of rice in India including with factors and without factors

		ARIMA models	$_{R}2$	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
West Bengal									
Without Factors	A	(1,1,5)	0.892	202.943	2.905	148.573	11.285	537.737	11.209
	P	(1,1,5)	0.958	820.083	8.811	622.392	28.843	2212.000	13.903
	Y	(0,1,1)	0.950	127.924	7.479	100.098	27.213	291.975	9.841
With factors	A	(1,1,1)	0.927	101.665	1.259	70.600	4.026	233.697	9.879
	P	(1,1,5)	0.99	371.388	2.496	239.672	8.756	681.14	13.106
	Y	(1,1,3)	0.992	50.324	2.072	34.139	9.428	118.231	8.897
Uttar Pradesh									
Without factors	A	(1,1,4)	0.913	229.874	3.035	149.835	17.978	824.279	11.359
	P	(1,1,1)	0.938	1001.000	13.836	688.727	117.118	2990.000	14.232
	Y	(1,1,5)	0.937	156.628	11.419	109.356	115.763	584.604	10.592
With factors	A	(1,1,4)	0.917	113.611	1.308	71.139	5.371	281.311	10.631
	P	(1,1,1)	0.978	464.778	3.817	311.474	12.685	932.064	13.237
	Y	(1,1,5)	0.985	60.755	2.841	40.867	11.097	120.629	9.379
Punjab									
	A	(1,1,5)	0.996	61.964	6.394	43.018	88.16	180.019	8.737
Without Factors	P	(1,1,3)	0.991	356.194	20.986	238.371	150.37	903.302	12.097
	Y	(1,1,3)	0.974	180.958	5.902	124.755	22.376	526.746	10.742
With factors	A	(0,1,2)	0.997	34.803	1.345	22.287	9.365	79.789	7.947
	P	(1,1,1)	0.995	194.978	2.380	138.892	8.610	373.496	11.393
	Y	(1,1,2)	0.957	85.003	1.826	58.155	6.945	188.343	9.839
India									
Without Factors	A	(1,1,5)	0.946	1066	1.869	716.742	10.078	4150	14.428
	P	(0,1,0)	0.927	1167	2.051	794.612	9.593	3950	14.194
	Y	(0,1,5)	0.956	101.537	5.508	72.241	25.307	284.749	9.656
With factors	A	(0,1,2)	0.956	492.192	0.751	315.849	3.157	1269	13.351
	P	(1,1,5)	0.985	2254	2.115	1453	6.022	4990	16.712
	Y	(0,1,5)	0.990	38.393	1.599	26.203	4.872	81.914	8.355

Table 5: Model validation and forecasting of area, production and yield of rice in India

	Area ('000 ha)			Production ('000 tonnes)			Yield (Kg/ha)		
Years	**Obs**	**Pred 1**	**Pred 2**	**Obs**	**Pred 1**	**Pred 2**	**Obs**	**Pred 1**	**Pred 2**
West Bengal									
2010	4944	5841	5786	13046	13316	13311	2639	2631	2638
2015	5524	5730	5657	15953	17513	15812	2888	2693	2767
2016	5496	5811	5580	15302	17717	15897	2784	2699	2795
2017	5240	5922	5520	14990	16178	15221	2861	2704	2822
2018		6184	6032		18932	16654		2708	2849
2019		6231	6058		19142	16544		2711	2877
2020		6278	6084		19355	17011		2713	2904
2021		6284	6091		19464	17119		2719	2922
2022		6297	6102		19511	17323		2723	2929
Uttar Pradesh									
2010	5657	5082	5753	11992	10779	11750	2120	2369	2172
2015	5862	5144	5862	12501	10450	12609	2133	2568	2306
2016	5992	5222	6103	13754	10084	12770	2295	2615	2334
2017	5814	5295	6245	13345	10053	12937	2295	2661	2361
2018		4463	6587		9695	13100		2705	2389
2019		4327	6629		9590	13266		2748	2416
2020		4186	6670		9229	13430		2791	2444
2021		4193	6688		9198	13511		2798	2456
2022		4206	6711		9179	13703		2814	2471
Punjab									
2010	2831	2666	2729	10837	10904	10950	3828	4157	4071
2015	2975	2801	2851	11823	12223	12019	3974	4407	4315
2016	2898	2824	2878	11586	12494	12208	3998	4462	4377
2017	2926	2844	2905	12283	12782	12392	4198	4517	4444
2018		2864	2933		13069	12572		4514	4572
2019		2881	2962		13366	12750		4588	4627
2020		2897	2991		13667	12925		4664	4682
2021		2903	2997		13721	12965		4722	4735
2022		2906	3001		13778	12991		4731	4758
India									
2010	42862	42493	43198	95970	86957	97021	2239	2216	2248
2015	43499	40981	44593	104408	85526	103320	2400	2276	2374
2016	43993	40642	44812	109698	83353	104561	2494	2285	2399
2017	42949	40291	43031	111007	83633	105808	2585	2292	2425
2018		39928	45250		81512	107060		2299	2450
2019		39552	45469		81463	108317		2305	2475
2020		39165	45688		79364	109577		2310	2500
2021		39122	45699		79131	111123		2316	2512
2022		39079	46112		78987	111521		2320	2521

Obs.: Observed, Pred.1: Predicated values without factors (Simple ARIMA), Pred.2: Predicated values with factors (ARIMAX).

to produce 13.7 million tones of paddy from an estimated area of 3 million hectare at a yield rate of more than 4.7q/ha, All these estimate leads to a total 111 million tones of paddy from an estimated area of 45.7 million hectare at an average yield of 2.52q/ha during the same period.

CONCLUSION

From the present investigation, it is found that rainfall has changed randomly during the study period for West Bengal, Uttar Pradesh, and Punjab. ARIMA modeling has long historical background throughout the world. In the present study, forecasting models have been compared with and without including different factors of production. It is evident from the study the meteorological factor besides helping crop production, are also important in increasing efficiency and accuracy of the crop forecasting models. Non-stationary time series data on different production parameters could be made stationery after Ist differencing and which show better accuracy in the ARIMA models after the inclusions of factors (ARIMAX) of productions; in the process forecasting power of selected models get enhanced. During the model validation period, a model with the factors i.e., the ARIMAX predicted closely to realized values, whereas during the out of sample forecasting, these have got tendencies to predicting lower values compared to other competitive models. For production, when we see the expected values, it is clear that West Bengal, Uttar Pradesh, and Punjab will have to play a vital role in the rice production of India. Punjab, in terms of the productivity of rice, would be the leading state of India. The study reveals that knowing the production potentials in association with factors of production will improve the forecasting accuracy, which in turn may help in formulating effective agriculture policy related to the particular crop.

REFERENCES

Anonymous. 2011. *www.indiastat.com*, Last accessed December 11.2011.

Anonymous. 2011. *www.indiawaterportal.com,* last accessed January 20.2011.

Anonymous. 2017. *Agriculture at a Glance*. Available from http://www.agricoop.nic.in/Last accessed on 20 Oct. 2017.

Anonymous. Various issue. *Fertilizer statistics*, Govt. of India.

Biswas, R., Banerjee, S. and Bhattacharyya, B. 2018. Impact of temperature increase on performance of *kharif* rice at Kalyani, West Bengal using WOFOST model. *J. Agrometeorol.*, **20**(1): 28-30.

Box, G.E.P. and Jenkins, G.M. 1976. *Time Series Analysis: Forecasting and Control*, Holden-Day, San Fransisco.

Diebold, F.X. and Mariano, R.S. 1995. Comparing predictive accuracy. *J. Bus. Econ. Stat.*, **13**: 253-263.

Ďurka Peter, Pastoreková Silvia. 2012.ARIMA vs. ARIMAX – which approach is better to analyze and forecast macroeconomic time series. Proceedings of 30th International Conference Mathematical Methods in Economics.

Food and Agriculture Organisation of the United Nations 2002. Nutritional studies, F.A.O., Rome.

Gideon E. Schwarz. 1978. Estimating the dimension of a model, *Ann. Stat.*, **6**(2): 461-464.

Gouri K. Bhattacharyya. 1984. Tests of randomness against trend or serial correlations Handbook of Statistics, Elsevier, Volume 4, Pages 89-111, ISSN 0169-7161, ISBN 9780444868718, https://doi.org/10.1016/S0169-7161(84)04007-4.

Grubbs, F.E. 1950**.** Sample criteria for testing outlying observations. *Ann. Math. Stat.,* **21**: 27-58.

Grubbs, F.E. 1969. Procedures for detecting outlying observations in samples. *Technometrics,* **11**(1): 1-21.

Grubbs, F.E. and Beck, G. 1972. Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics,* **14**: 847-854.

Harvey, D., Leybourne, S. and Newbold, P. 1997. Testing the equality of prediction mean squared errors. *Int. J. Forecast*., **13**(2): 281-291.

Intriligator, M.D., Bodkin, R.G. and Hsio, C. 1996. Econometric Models, Techniques, and Applications", Prentice Hall edition.

Joshi, P.K., Gulati, A., Birthal, P.S. and Tewari, L. 2004. Agriculture diversification in South Asia: Patterns, determinants and policy implications. *Econ. Polit. Wkly,* **39**(18): 2457-2467.

Joshi, P.K., Tewari, L. and Birthal, P.S. 2006. Diversification and its impact on smallholders: Evidence from a study on vegetable production. *Agric. Econ. Res. Rev.*, **19**(2): 219-236.

Kaul, S. 2006. Economic Analysis of Productivity of Rice Production- State-wise Analysis *presented at the 14th Annual Conference of Agricultural Economics Research Association,* held during Sept. 27-28, 2006 at G.B. Pant Univ. of Agri. & Tech.., Pant Nagar, Uttaranchal.

Kebebe, Ergano, Mehta, V.P. and Dixit, P.N. 2000. Diversification of agriculture in Haryana: An empirical analysis. *Agril. Situation in India,* **57**(8): 459-463.

Khush, G. 2003. Productivity improvements in rice. *Nutr. Rev.,* **61**(6 Pt 2): S114-S116.

Mishra, P., Fatih, C., Niranjan, H.K., Tiwari, S., Devi, M. and Dubey, A. 2020. Modelling and Forecasting of Milk Production in Chhattisgarh and India. *Indian J. Anim. Res.,* **54**(7): 912-917.

Mishra, P., Sahu, P.K., Padmanaban, K., Vishwajith, K.P. and Dhekale, B.S. 2015. Study of Instability and Forecasting of Food Grain Production in India. *Int. J. Agric. Sci.*, **7**(3): 474-481.

Mishra, P., Sahu, P.K., Dhekale, B.S. and Vishwajith, K.P. 2015. Modeling and forecasting of wheat in India and their yield Sustainability. *Soc. Eco. Dev*., **11**(3): 637-647.

Mishra, P., Matuka, A., Abotaleb, M.S.A., Weerasinghe, W.P.M.C.N, Karakaya, K. and Das, S.S. 2021. Modeling and Forecasting of Milk Production in the SAARC countries and China, *Model. Earth Syst. Environ*., pp. 1-9.

Mishra, P. 2015. Study of instability and forecasting of food grain production in India. *Int. J. Agric. Sci., ISSN*, 0975-3710.

Raiger, H.L., Sahu, P.K., Behera, M.P. and Jajoriya, N.K. 2019. Variation and Character Association in Seed Yield and Related Traits in Rice Bean (*Vigna umbellata*). *Int. J. Agric. Environ. Biotech.*, **12**(1): 13-16.

Ramesh, T., Rathika, S., Subramanian, E. and Ravi, V. 2020. Effect of Drip Fertigation on the Productivity of Hybrid Rice. *Int. J. Agric. Environ. Biotech.*, **13**(2): 219-225.

Ravichandran, S., Muthuraman, P. and Rao, P.R. 2012. Time - series modelling and forecasting India's rice production - arima vs stm Modelling approaches. *Int. J. Agric. Stat. Sci.*, **8**(1): 305-311.