Indian Journal of Weed Science **47**(4): 408–410, 2015

Forecasting of herbicide consumption using autoregressive integrated moving average model

Yogita Gharde*

ICAR-Directorate of Weed Research, Jabalpur, Madhya Pradesh 482 004

Received: 16 October 2015; Revised: 28 November 2015

ABSTRACT

A study was conducted on modelling and forecasting the time series data of total herbicide consumption in India. Among many time series methodology, Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model was used for modelling and forecasting purposes using data from 1990 to 2010. Before the modelling, stationarity of the data was checked using Augmented Dicky Fuller test. Best model was chosen using two criterion *viz.* Akaike information criterion and Schwarz's Bayesian criterion. ARIMA (0, 1, 1) model was found to the best among many models from ARIMA family. Forecasting was done using the best model and prediction for total herbicide consumption in India was made for next three year (2011, 2012, 2013) as 6624, 6581 and 6562 tonnes respectively.

Key words: ARIMA, Forecasting, Herbicide consumption, Modelling, Time series

Nowadays, it is very difficult for farmers to employ sufficient hand labour to keep their fields clear of weeds, making herbicides the most logical alternative. According to the Federation of Indian Chambers of Commerce and Industry (FICCI), per capita consumption of agrochemicals in India is still much lower at 0.6 kg/ha compared to 13 kg/ha in China and 7 kg/ha in USA. Herbicides are the largest growing segment and currently account for 16% of the total market (FICCI 2013). Among many agrochemicals, herbicide consumption is increasing day by day with increase in necessity and awareness among farmers in India. Therefore, total consumption of herbicides is very important figure for policy makers and companies to predict the demand and supply of herbicides for the next year. Companies try to forecast demand, plan production, set sales targets and compete to push stocks to distributors based on primary sales, without having an idea of actual demand and consumption by farmers. Use of available technology to track sales, if at all, is ineffective. Further, it is also observed that yearly data on herbicide consumption is not easily available even after 2-3 years, as the main source of the data is only the concerned industry. In this situation, it is difficult for policy makers and industry persons to take appropriate measures in this regard. Therefore, it becomes necessary to predict the consumption of herbicides in advance.

***Corresponding author:** yogitagharde@gmail.com

In many scientific areas, data is observed in the form of time series. Hence to analyze the time series data, many models are fitted either to better understand the data or to predict future points in the series (forecasting). In this context, Autoregressive Integrated Moving Average (ARIMA) models are the most popular and widely used forecasting models for the analysis of uni-variate time series data. Applying ARIMA models, Iqbal *et al.* (2005) forecasted area and production of wheat in Pakistan. In India, Sarika *et al.* (2011) modelled and forecasted the pigeonpea production, Padhan (2012) predicted agricultural productivity and Kumari *et al.* (2014) projected rice yield. In past, no study has been conducted to predict the total herbicide consumption in India. In view of this, an attempt has been made to model and forecast the herbicide consumption in India for consecutive three years based on the past data using ARIMA models.

MATERIALS AND METHODS

India's herbicide consumption data from 1990 to 2010 were collected from www.faostat.fao.org. Data from 1990 to 2008 were used for model building and 2009 and 2010 for model validation. The present study applies Box-Jenkins ARIMA model which is an extrapolation method, and requires past time series data of study variable. The model may be expressed as follows. Let Y_t is a discrete time series variable which takes different values over a period of time. Then ARIMA (p, d, q) model can be given as

$$Y_t = \mu_0 + \mu_1 Y_{t-1} + \mu_2 Y_{t-2} + \cdots + \mu_p Y_{t-p} + \epsilon_t + \delta_1 \epsilon_{t-1} + \cdots + \delta_q \epsilon_{t-q} + \gamma t$$

where $Y_t, Y_{t-1}, \dots Y_{t-p}$ are the values of response variable at time t, t-1, t-2 and so on, $\mu_0, \mu_1, \dots \mu_p$ are the coefficients associated with response variable, $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-q}$ are the error term, $\delta_1, \dots, \delta_q$ are the coefficients of the estimated error term and γt is the constant mean of the series, p and q are the parameters of the model and d denote the degree of differencing applied to the data. In order to identify the order of p and q, the ACF (Autocorrelation function) and PACF (Partial autocorrelation function) is applied.

ARIMA modelling and forecasting is performed in four steps: model identification, model estimation, model validation and forecasting. Preliminary, it is very important to check about the stationarity of the data. Since ARIMA modelling is applied only to stationary data (data with constant mean and variance). Therefore, to check the stationarity of the series, Augmented Dickey Fuller (ADF) test was applied. Modelling and forecasting was done using SAS Enterprise Guide 4.3 (SAS Institute Inc., USA). Many models from ARIMA family were used to fit the time series data with different values of p, d and q. Among these, best model was selected using two criterion, *viz.* Akaike information criterion (AIC) and Schwarz's Bayesian criterion (SBC). After identifying the model, parameter estimates were obtained for chosen model. Residual analysis was also done to check the independency of the white noise residuals. If the model is appropriate representation of the time series, then the white noise residuals should be independent of each other. Thus, any significant autocorrelation in the estimated residuals at ACF and PACF indicates model inadequacy and suggests the model modification.

RESULTS AND DISCUSSION

ARIMA modelling starts with the checking for stationarity of the series. Therefore, Augmented Dickey Fuller test (ADF) was applied which revealed the non-stationarity of the data. Hence, first differencing was applied to the data. Later, data series became stationary. Among many models from ARIMA family, ARIMA (0, 1, 1) model was found to be best after taking minimum AIC and SBC criteria for selecting model (Fig. 1). Parameter estimates were obtained using conditional least square method for the selected model and presented (Table 1) along with standard error for the fitted ARIMA (0, 1, 1) model. Selected model is given by: $\hat{Y}_t = \mu_0 + Y_{t-1} - \delta_1\, e_{t-1}$, where μ_0 is the constant, δ_1 is the constant associated with error term. $\hat{Y}_t$ is the estimated value of herbicide consumption at time t and Y_{t-1} is the data point associated with time t-1. Forecast values were obtained for the year 2011, 2012 and 2013 as 6624, 6581 and 6562 tonnes respectively.

Table 1. Parameter estimates along with standard error

Parameter	Estimate	Standard Error	t Value	Pr > |t|
μ_0 (constant)	-24.83	30.81	-0.19	0.85
δ_1	0.51	0.22	2.33	0.03

Thus model is obtained as $Y_t = -24.83 + Y_{t-1} - 0.51\, e_{t-1}$.

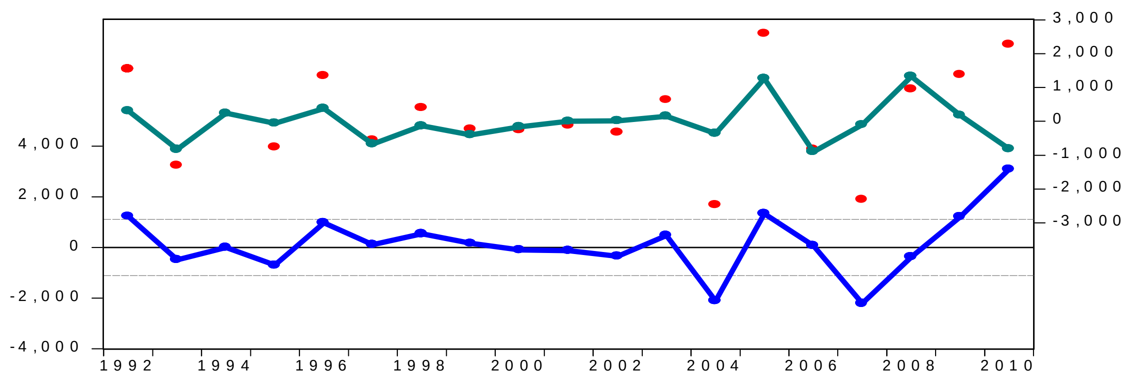

Fig 1. Fitted ARIMA (0, 1, 1) model along with original value* and forecast error

Table 2. Autocorrelation check of residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	1.68	4	0.7936	-0.078	0.030	-0.054	0.065	0.106	-0.178
12	4.77	10	0.9059	0.113	-0.176	0.062	-0.129	0.006	-0.107
18	11.82	16	0.7560	0.061	0.087	-0.195	-0.050	-0.042	0.104

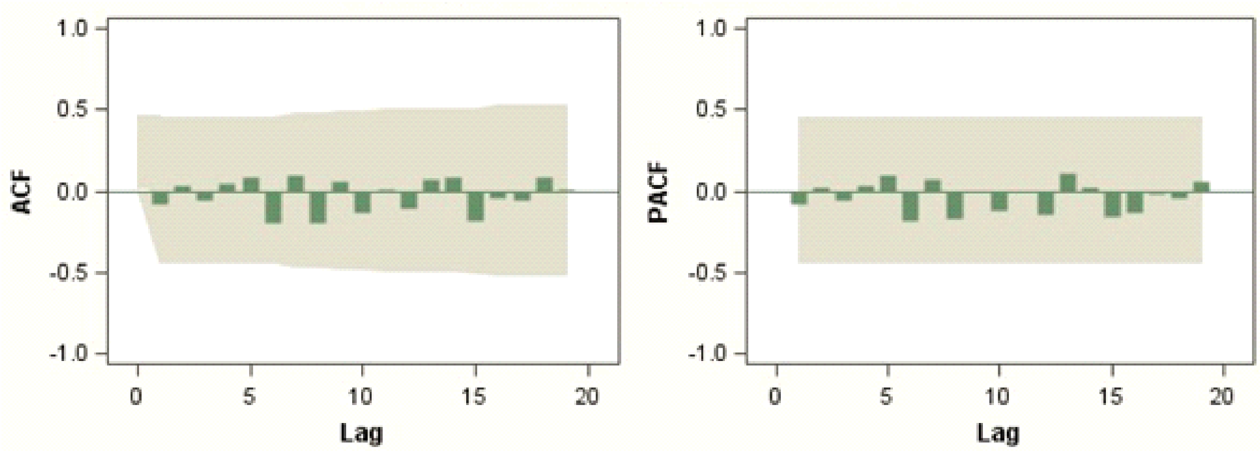

Fig. 2. Residual ACF (autocorrelation function) and PACF (partial autocorrelation function) plots for ARIMA (0, 1, 1) model

Forecast values obtained from the model shows decreasing trend in the herbicide consumption of India. It might be due to the development of low dosage and high potency molecules which require less volume of chemical per unit treated area in comparison to older chemicals.

Residual analysis was done to check the independency of white noise residuals. Non-significant chi-square values led to rejection of hypothesis that residuals are autocorrelated. Hence, it revealed that residuals are independent to each other which was also confirmed by ACF and PACF study as no significant spikes were present in the plots and thus assumption of independence of residuals was not violated (Table 2).

Modelling and forecasting of herbicide consumption data was done using ARIMA modelling. Among the many models from ARIMA family, ARIMA (0, 1, 1) model is found to be best. After an appropriate time series model was decided, its unknown parameters estimated, it was established that the model fitted well. Using the selected model, forecast values for herbicide consumption were obtained for consecutive three years. These are 6624, 6581 and 6562 tonnes for the year 2011, 2012 and 2013, respectively. Forecasting may be done for further many more years but since forecast is associated with standard error, which gradually increase with increase in number of forecast values, therefore, it must be done for short term only.

REFERENCES

Box G and Jenkins G. 1970. *Time Series Analysis: Forecasting and Control. San Francisco*: CA: Holden-Day. 1976. 575 p.

Dickey DA and Fuller WA. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association* **74**: 427-431.

FICCI - KPMG report 2013.

http://www.faostat.fao.org.

Iqbal NBK, Maqbool Asif and Shohab Abid Ahmad. 2005. Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan. *Journal of Agriculture and Social Sciences* **1**(2): 120-122.

Koutroumanidas T, Ioannou K and Arabatzis G. 2009. Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and hybrid ARIMA-ANN model. *Energy Policy* **37**: 3627-34.

Kumari P, Mishra GC, Pant AK, Shukla G and Kujur SN. 2014. Autoregressive Integrated Moving Average (ARIMA) approach for prediction of Rice (*Oryza Sativa L.*) yield in India. *The Bioscan* **9**(3): 1063-1066.

Padhan PC. 2012. Application of ARIMA Model for Forecasting Agricultural Productivity in India. *Journal of Agriculture & Social Sciences* **8**(2): 50–56.

Sarika, Iquebal MA and Chattopadhyay C. 2011. Modelling and forecasting of pigeonpea (*Cajanas cajan*) production using autoregressive integrated moving average methodology. *Indian Journal of Agricultural Sciences* **81** (6): 520-3.