RESEARCH PAPER International Journal of Agricultural Engineering | Volume 8 | Issue 1 | April, 2015 | 97–102
e ISSN–0976–7223 ■ Visit us : *www.researchjournal.co.in* ■ **DOI: 10.15740/HAS/IJAE/8.1/97-102**

Multiplicative seasonal ARIMA modelling of monthly stream flows of Choriti river

B.L. AYARE AND B.S. DHEKALE

Received : 21.02.2015; **Revised :** 05.03.2015; **Accepted :** 16.03.2015

See end of the Paper for authors' affiliation

Correspondence to :

B.L. AYARE
AICRP on Water Management, C.E.S. Wakawali, Dr. Balasaheb Sawant Konkan Krishi Vidyapeeth, Dapoli
RATNAGIRI (M.S.) INDIA
Email : blayare@yahoo.co.in

■ **ABSTRACT :** The multiplicative seasonal ARIMA (p,d,q) × $(P,D,Q)_s$ models of different orders were tried for modelling of monthly inflow of Choriti river of Konkan region of Maharashtra, based on 20 years data. The parameters of seasonal ARIMA models were estimated by fitting ARIMA models to differenced series (d=0 and D=1) at different lags. The goodness of fit of models was tested by Box-Pierce Portmanteau lack of fit test and comparison of historical and forecasted monthly inflows. The forecasted performance of the model was evaluated by using goodness of fit tests. Lower values of root mean squared error; mean relative error and integral square error for multiplicative seasonal ARIMA (0,0,1) × $(0,1,1)_{12}$ model indicated closer agreement between forested and historical monthly inflow series.

■ **KEY WORDS :** Autoregressive integrated moving average model, Forecasting, White noise, Akaike information criteria

■ **HOW TO CITE THIS PAPER :** Ayare, B.L. and Dhekale, B.S. (2015). Multiplicative seasonal ARIMA modelling of monthly stream flows of Choriti river. *Internat. J. Agric. Engg.*, **8**(1) : 97-102.

The world today faces an expanding population and an accompanying increasing demand for water resources. Effective design and operation of water resource systems and for efficient use of available water requires the ability to forecast inflows. A mathematical model of monthly stream flow is a very useful tool for generation of synthetic data and forecasting. Monthly inflows exhibiting non stationarity and seasonality can be modelled using ARIMA models described by Box and Jenkins (1976). Such models have been used in past, to limited extent to model monthly and annual stream flows (Hipel *et al.*, 1977). Mohan and Vedula (1995) developed multiplicative seasonal ARIMA model for long term forecasting of monthly inflows based on 25 years data with logarithmic transformation. Salas and Chung (2001) developed a method for determining the autocorrelation function of discrete series as a function of original continuous series of low order autoregressive moving average and low order discrete autoregressive moving average (ARIMA) processes and analyzed the relationship of autocorrelation and crossing rates. Chen and Rao (2002) explained that hydrological monthly series are stationary; a segmentation algorithm is applied so that non-stationary series are identified and partitioned into stationary segments. Phoon *et al.* (2002) presented a practical inverse approach for forecasting non-linear hydrological time series. Patil (2003) made stochastic modeling for water deficit by using 24 years (1976-1999) data. Sharma *et al.* (2003) developed a stochastic model for forecasting inflow of a non seasonal river in Rajasthan. Chhajed (2004) developed stochastic model for Mahi river inflow series. He suggested that ARMA (3,1) model can be used for one-time step ahead monthly forecasting of Mahi river inflows. Minimum mean square error (MMSE) criterion was used for selection of best model. Verma (2004) developed stochastic model on monthly rainfall of Kota, Rajasthan. Yurekli *et al.* (2005) developed a linear stochastic model to monthly data of Kelkit stream, North Anatolia (Turkey). The ARIMA model were used to simulate monthly data Diagnostic checks were done for all the model selected from the autocorrelation function (ACF) and partial autocorrelation function (PACF). The result shown that, the generated data preserve the basic statistical properties of the original series. Kourosh Mohammadi *et al.* (2006) developed method for parameter estimation of an ARMA model for river flow forecasting using goal programming. The results when compared with usual method of maximum likelihood estimation were favorable with respect to the new proposed algorithm. Kumar and Kumar (2006) used the seasonal autoregressive integrated moving average (SARIMA) model of different orders for modeling of monthly stream flow of Betwa river of Matatila dam site in Jhansi district of Uttar Pradesh. Wagh and Devendra Kumar (2006) developed a autoregressive (AR) model for annual stream flow of Godavari river. They have used AR models of order 1, 2, 3, 4 and 5. AR(1) model was found suitable based on Box-Pierce Portmanteau test and Akaike Information Criteria. Chimirala developed a time series model for forecasting of rainfall at Saidapet rain gauge station, Chennai (Tamil Nadu). Eno Rai and Arpan Sherring (2007) developed autoregressive (AR) time series model for prediction of rainfall and runoff for Manshara watershed for lower Gomati catchment of Utter Pradesh. Heung Wong *et al.* (2007) developed non-parametric time series models for hydrological forecasting of river inflows. Machiwal and Jha (2008) evaluated twenty nine statistical tests for detecting time series characteristics, to analyze 46 years of annual rainfall, 47 years of 1-day maximum rainfall and consecutive 2, 3, 4, 5 and 6 day maximum rainfall of Kharagpur, West Bengal, India. Andrea *et al.* (2009) studied mean monthly nitrate concentration from the Des Moines river, IOWA, U.S.A. for a 30 years period (1977-2006) using time series analysis. Theodoros *et al.* (2009) developed non linear time series and forecasting model and applied to the mean monthly Nestos river discharge data for the 1966-2008 periods.

■ METHODOLOGY

The seasonal autoregressive integrated moving average (ARIMA) models are useful for modelling seasonal time series in which mean and other statistics for a given season are not stationary across the year. The differencing of time series is used to remove its non-sationarity and trend. It is possible to take first, second or in general d-th difference, which leads to simple non-seasonal ARIMA (p,d,q) models. It is possible to take periodic or seasonal differences at lag w, such as 12th difference of monthly series, which leads to seasonal ARIMA (p, d, q) × (P, D, Q)$_s$ models.

A general ARIMA (p, d, q) × (P, D, Q)$_s$ models has following form (Box and Jenkins, 1976);

$$\phi_p (B)\, \Phi_P (B^s)\, \nabla^d\, \nabla_s^D Z_t = \theta_q (B)\, \Theta_Q (B^S)\, a_t \quad .....(1)$$

where,

ϕ_p (B) = 1- ϕ_1 B - ϕ_2 B^2 - ….. - ϕ_p B^p

Φ_p (B^S) = 1 - Φ_1 B^S - Φ_2 B^{2S} - …… - Φ_P (B^{PS})

Θ_Q (B^S) = 1 - Θ_1 B^S - Θ_2 B^{2S} - …. - Θ_Q B^{QS}

θ_q (B) = 1 - θ_1 B - θ_2 B^2 - ….. - θ_q B^q

B is the backward shift operator (*i.e.* $BZ_t = Z_{t-1}$), s is the number of seasons per year (s=12 for monthly data), Φ's and ϕ's are seasonal and non-seasonal AR parameter, respectively, $\Theta_{,s}$ and θ's are seasonal and non-seasonal MA parameters, p and q are orders of non-seasonal AR and MA parameters, respectively, P and Q are order of seasonal AR and MA parameter, respectively, d and D are non-seasonal and season differences, respectively, a_t is the residual series (t =1,2,….) which has mean zero and variance, σ^2.

The box and jenkins formalized the modelling process through following steps :

– Model identification : The orders of models are determined.

– Parameter estimation : The linear co-efficients of the model are estimated based on maximum likelihood or minimum least square.

– Model validation: Certain diagnostic checking methods are used to test the suitability of the models.

– Foresting: The best models chosen are used for forecasting.

■ RESULTS AND DISCUSSION

The findings of the present study as well as relevant discussion have been presented under following heads :

Data used in the study:

The daily inflow data of 20 years converted into monthly data, were used in this study *i.e.* from the year 1988 to 2007 of the Choriti river at Natuwadi dam site of Konkan region, Maharashtra. In the present study, the multiplicative seasonal ARIMA model has been built following the above steps. The details of procedure adopted in each step are explained below:

Model identification :

The sample autocorrelation function (ACF) and partial autocorrelation function (PACF) of the original inflow series at different lags (upto 70) and their 95 per cent confidence limits were computed and plotted in Fig. 1 and 2, respectively.

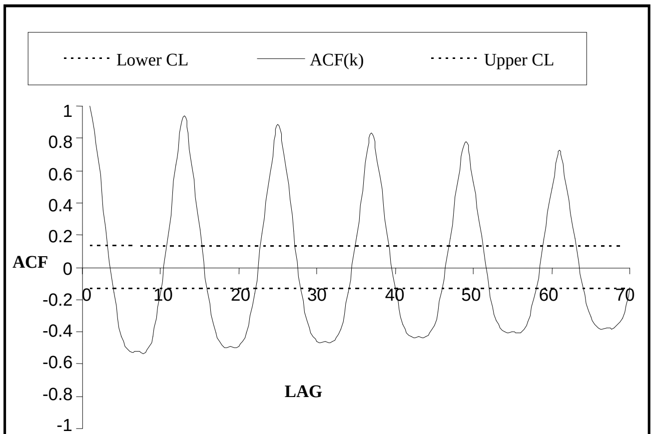

Fig. 1: Sample ACF of original inflow series with 95 per cent confidence limits

The ACF and PACF of monthly inflow of differenced series (d=0 and D=1) at different lags along with their confidence limits were developed for the identification of proper type and orders of seasonal

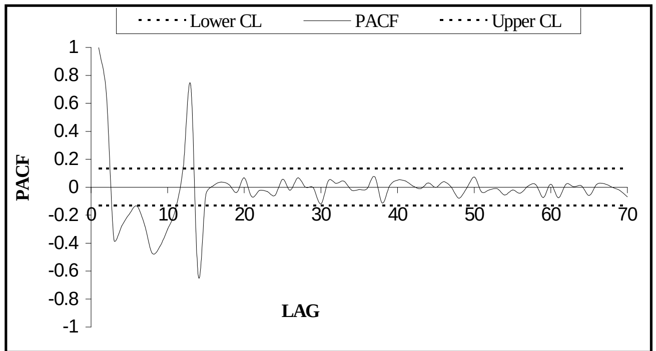

Fig. 2: Sample PACF of original series with 95 per cent confidence limits

ARIMA models and shown in Fig. 3 and 4, respectively. After examining the ACF and PACF values of the above differencing scheme, eight candidate models (Table 1) are selected for the study.

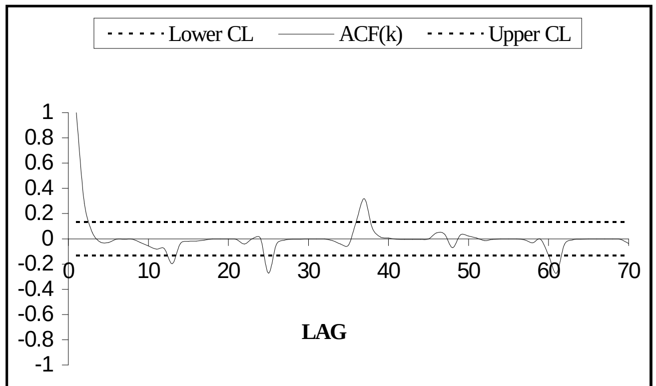

Fig. 3: ACF of differenced series (d=0 and D=1) with 95 per cent confidence limits

Table 1 : Goodness of fit of different seasonal ARIMA (p,d,q) ×(P,D,Q)$_s$ models

Sr. No.	Model structure	Goodness of fit			
		t > 2	AIC	Ljung box statistic	Box pierce
1.	(0,0,1) × (1,1,1)$_{12}$	No	884.24	85.70	72.14
2.	(0,0,1) × (1,1,0)$_{12}$	Yes	906.67	116.0	96.88
3.	(0,0,1) × (1,0,0)$_{12}$	Yes	973.29	109.50	91.30
4.	(0,0,1) × (0,1,1)$_{12}$	Yes	883.51	73.84	62.09
5.	(0,0,1) × (0,0,1)$_{12}$	Yes	1451.87	1635.18	1393.9
6.	(0,0,1) × (1,0,1)$_{12}$	No	889.80	81.0	68.32
7.	(0,0,1) × (2,0,0)$_{12}$	Yes	963.51	113.82	95.26
8.	(0,0,1) × (2,1,0)$_{12}$	Yes	853.02	49.93	42.12
9.	(0,0,1) × (2,1,1)$_{12}$	No	854.17	50.05	42.34
10.	(0,0,1) × (2,1,2)$_{12}$	No	854.60	50.56	43.30
11.	(0,0,1) × (0,0,2)$_{12}$	Yes	1222.70	1116.35	950.75
12.	(0,0,1) × (0,1,2)$_{12}$	Yes	877.37	65.90	54.96
13.	(0,0,1) × (1,1,2)$_{12}$	No	871.23	59.70	49.79

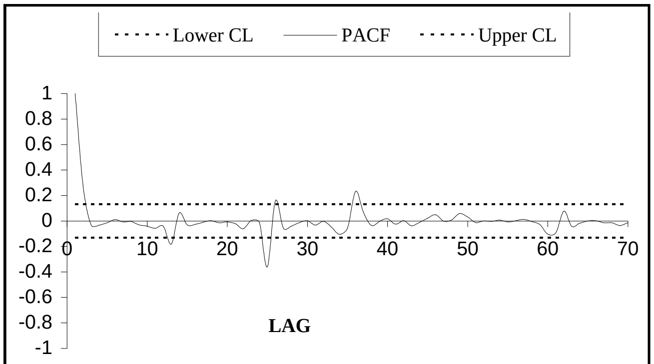

Fig. 4 : PACF of differenced series (d=0 and D=1) with 95 per cent

Parameter estimation :

The parameters of the each of the models structure were estimated using maximum likelihood estimation procedure.

Model validation :

To test the goodness of fit, t-test has been applied to each selected model. Out of 13 models, eight ARIMA models passed the t-test. The ACF and PACF of residuals (Kottegoda, 1980) were calculated with their confidence limits. Out of eight models, three seasonal ARIMA models with structure $(0,0,1 \times (0,1,1)_{12}$, $(0,0,1) \times (0,1,2)_{12}$, and $(0,0,1) \times (2,1,0)_{12}$ passed the ACF and PACF residual test. The residuals obtained from all these models are mutually independent as they are not significantly different from zero and are within 95 per cent confidence limits, confirming the residuals are white noise. The Akaike Information Criteria (AIC) is used to measure the goodness of fit. Selected 3 models are having lowest AIC values.

The Box –Pierce Portmanteau lack of fit test was used to check the adequacy of SARIMA models. The values of test statistics were computed (Table 1) using following equation:

$$Q = N \sum_{k=1}^{L} r_k^2(\varepsilon) \quad \text{......(2)}$$

where, $r_k(\varepsilon)$ is the correlogram of the residuals ε_t

L is the maximum lag considered.

N is the number of observations in the series.

The test statistics were compared with the tabulated values of Chi-square. The comparison of test statistics for all three models is found to be less than the tabulated value. Therefore, all three models *viz.*, ARIMA $(0,0,1) \times (0,1,1)_{12}$, ARIMA $(0,0,1) \times (0,1,2)_{12}$, ARIMA $(0,0,1) \times (2,1,0)_{12}$ are giving good fit and are acceptable.

Forecasting with seasonal ARIMA model :

The selected three ARIMA models were used to forecast the inflows for one year ahead. The comparison between actual and forecasted inflows is shown in Table 2. It is observed that, the seasonal monsoon pattern of river inflow series is maintained in forecasted values by all the three models.

Forecasting performance of the models was evaluated quantitatively using following tests of goodness of fit.

Root mean squared error (RMSE) :

Root mean squared error (RMSE) was used for evaluation of forecasting performance of the SARIMA models. Lower the value of RMSE, better is the model. The RMSE can be predicted by following mathematical relationship:

Table 2 : Comparison of actual and forecasted monthly Choriti river inflow (M cu.m) series by SARIMA models

Sr. No.	Model structure	June	July	August	September	October
1.	Actual	3.23	19.35	27.78	28.08	28.08
2.	ARIMA $(1,0,0) \times (0,1,1)_{12}$	8.81	17.60	24.46	28.02	28.0
3.	ARIMA $(1,0,0) \times (0,1,2)_{12}$	11.70	17.89	24.80	27.92	27.97
4.	ARIMA $(1,0,0) \times (2,1,0)_{12}$	11.13	18.58	25.10	28.06	28.08

Table 3 : Estimation of forecasting performance of model with statistical errors

Sr. No.	Model structure	Statistical errors		
		RMSE	MRE	ISE
1.	ARIMA $(0,0,1) \times (0,1,1)_{12}$	1.67	0.88	0.57
2.	ARIMA $(0,0,1) \times (0,1,2)_{12}$	2.46	1.17	0.83
3.	ARIMA $(0,0,1) \times (2,1,0)_{12}$	2.27	1.03	0.77

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{Observed} - \mathrm{Forcasted})^2}{\sqrt{n}}} \qquad .....(3)$$

where, n is the number of observations in a year.

Mean relative error (MRE) :

Raghuwanshi and Wallender (2000) have given the criterion of mean relative error (MRE) for evaluation of the forecasting performance of the models. The mean relative error was computed by the following equation:

$$\mathrm{MRE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\frac{\mathrm{observed} - \mathrm{Forecasted}}{\mathrm{Forecasted}}\right)^2}{n}} \qquad ......(4)$$

where, n is the number of observations in a year.

Then select the model that results in the least value of the mean relative error. That model is the selected as the most appropriate model.

Integral square error(ISE) :

Singh *et al.* (1991) used the integral square error (ISE) as a measure of goodness of time series model for air temperature. The integral square error was computed to evaluate the foresting performance of the model for monthly inflow series by the following equation :

$$\mathrm{ISE} = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{Observed} - \mathrm{Forcasted})^2}{\sqrt{\sum_{i=1}^{n}\mathrm{Forecasted}}}} \qquad ......(5)$$

Lower the value of ISE, better is the model.

The values of RMSE, MRE and ISE for SARIMA models given in Table 3.

From the comparison to test the goodness of fit in Table 3, it revealed that, the seasonal ARIMA (0,0,1) × $(0,1,1)_{12}$ has lowest values of RMSE, MRE and ISE, so this model provided good fit.

Conclusion :

The following conclusions are drawn from the present study.

– The values of root mean squared error, mean relative error and integral square error for seasonal ARIMA (0,0,1) × $(0,1,1)_{12}$ model were found to be 1.67, 0.88 and 0.57, respectively.

– The most appropriate model for forecasting monthly inflow of Choriti river at Natuwadi medium irrigation project was found to be seasonal ARIMA (0,0,1) × $(0,1,1)_{12}$.

Acknowledgement :

Authors sincerely acknowledge the kind assistance rendered in the form of data by the official and personnel of the Irrigation Department, Konkan region, Government of Maharashtra.

Authors' affiliations:

B.S. DHEKALE, Bindan Chandra Krishi Vishwavidyalaya, MOHANPUR (W.B.) INDIA

■ REFERENCES

Andrea, L.S., Keith, E.S. and Kung, S.C. (2009). Time series modelling of reservoir effect on river nitrate concentrations. *Adv. Water Resou.*, **32** (8) : 1197-1205.

Box, G.E.P. and Jenkins, G.M. (1976). *Time series analysis : Forecasting and control,1st Edn*. Holden Day San Francisco, CA. 1-57pp.

Chen, H.L. and Rao, A.A. (2002). Testing hydrologic time series for stationary. *J. Hydrologic Engg.*, **7** (2) : 129-136.

Chhajed, N. (2004). Stochastic modeling for forecasting Mahi river inflows. M.E. Thesis, Maharana Pratap University of Agriculture and Technology, Udaipur, RAJASTHAN (INDIA).

Chimirala, ch., Srinivas, P., Srinivas, K. and Kumar, Satish, S. Thayumanavan (2007). Box and Jenkins models application for rainfall forecasting of saidpet raingauge station, Chennai National Conference on Watershed Management and Impact of Environmental changes on water Resources, Department of Civil Engineering J.N.T.U., Hyderabad (India). 124-132pp.

Eno Rai and Sherring, Arpan (2007). Development of Autoregressive time series model for prediction of rainfall and runoff for Manshara watershed of lower Gomati catchments. *J. Agric. Engg.*, **44** (4) : 38-42.

Heang Wong, Wai, Ch.lp Zhang, R. and Jun Xia (2007). Non-parametric time series models for hydrological forecasting. *J. Hydrol.*, **332** (3-4) : 337-347.

Hipel, K.W., McLeod, A.I. and Lennox, W.C. (1977). Advances in Box Jenkins modeling: Model Construction. *Water Resour. Res.*, **13** (3) : 567-575.

Kottegoda, N.T. (1980). *Stochastic water resources technology*. The Macmillan Press Ltd., London. 384pp.

Kourosh, Mohammadi, Eslami, H.R. and Rene, Kahawita (2006). Parameter estimation of on ARMA model for river inflow forecasting using goal programming. *J. Hydrol.*, **331** (1-2) : 293-299.

Kumar, Mukesh and Kumar, Devendra (2006). Seasonal ARIMA modelling of monthly stream flows. *J. Appl. Hydrol.*, **4**: 91-98.

Machiwal, Deepesh, and Jha, M.K. (2008). Comparative evaluation of statistical tests for time series analysis: application to hydrological time series. *Hydrological Sci. J.*, **53** (2) : 353-366.

Mohan, S. and Vedula, S. (1995). Multiplicative seasonal ARIMA model for long-term forecasting of inflows. *J. Water Resour. Mgmt.*, **9** (2) : 115-126.

Patil, R.M. (2003). Stochastic modeling of water deficit for Rahuri region. M.E. Thesis Maharana Pratap University of Agriculture and Technology, Udaipur, RAJASTHAN (INDIA).

Phoon, K.K., Islam, M.N., Liaw, C.Y. and Liong, S.Y.(2002). Practical inverse approach for forecasting non linear hydrological time series. *J. Hydrologic Engg.*, **7** (2) : 116-128.

Raghuwanshi, N.S. and Wallender, W.W. (2000). Forecasting daily evapotranspiration for a grass reference crop. *Agric. Engg. J.*, **9** (1) : 1-16.

Salas, J.D. and Change, C. (2001). Relating autocorrelation and crossing rates of continuous and discrete valued hydrologic Engineering. March-April: 109-118pp.

Sharma, A.,Singh, R.V., Narulkar, S.M. and Dashora, P.K. (2003). Identification of Appropriate stochastic model for forecasting Monsoonal river inflows. *J. Agric. Engg.*, **12** (3-4) : 191-208.

Singh, A.K., Singh, J. and Gupta, A.P. (1991). Time series analysis of air temperature of Udaipur. *Hydrol. J.*, **14**: 16-23.

Theodoros, K., Georgios, S., Eleni, Z. and Vassilios, T.A. (2009). Genetic modeling for the optimal forecasting of hydrologic time series; Application in Neston River, *J. Hydrol.*, **368** (1/4) : 156-164.

Verma, A. (2004). Stochastic modeling on monthly rainfall of Kota, Rajasthan. M.E. Thesis, G.B. Pant University of Agriculture and Technology, Pantnagar, U.S. Nagar, UTTARAKHAND (INDIA).

Wagh, R.N. and Kumar, Devendra (2006). Modeling of annual stream flows by autoregressive models. *J. Agric. Engg.*, **43** (4) : 47-51.

Yurekli, Ahmet, Kurunc and Fazli, Ozturk (2005). Application of linear stochastic model to monthly inflow data of kelkit. *J. Ecol. Model.*, **183**: 67-75.