Modeling of Ground Water Tables in Three Districts of Red and Laterite Zone of West Bengal

Vishwajith K. P., P. K. Sahu[*], B. S. Dhekale, Md. Noman, L. Narsimhaiah

Department of Agricultural Statistics, Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, West Bengal

Abstract Ground water table is an important indicator not only for sustainable agriculture but also for assured supply of water to industries and other activities. In this regard fluctuations of ground water table plays an important role. Modeling of ground water table and its fluctuation behavior are important in these aspects. The present paper is an attempt to model the ground water table in three districts viz. Bankura, Birbhum and Purulia under drier Red and Laterite zone of West Bengal. Time series ground water level data on seasonal basis (i.e January, May, August and November) for a period 2005 to 2013 are used for the purpose. For obvious reason ground water table demonstrates seasonal periodicity. In the study, the famous Box-Jenkins ARIMA modeling techniques has been used to model the ground water table with the inclusion of seasonality parameters in the model. Among the competitive SARIMA models the best fitted model for respective series is selected based on R^2, AIC, BIC, MAE, MSE, RMSE and other criteria. Selected best model are put under model validation through comparisons of out of sample observed and predicted data. In the last stage, the models are subjected to diagnostic check through the checking of ACF and PACF of model residuals. The best models, selected thus are used for forecasting purposes. The method used and the result obtained will give an indication about the future behaviour of ground water table in these districts and will also help the planners and ultimate users to take appropriate measures in utilization of ground water in most efficient manner.

Keywords ARIMA, Forecasting, Modeling and seasonality

1. Introduction

Our planet is apparently rich in water but about 97.5 per cent (Hamdy and Lacirignola, 2005) of its water resource is saline and unsuitable for drinking or irrigation or is of little direct use of people; only 2.5 percent is usable but sometimes not readily accessible. About 68.7 per cent of fresh water is piled up in high latitude and altitude as snow and glaciers and 29.9 per cent is remaining as ground water and soil moisture (Anon, 2010). India gets 4000 km^3 of water (0.70 per cent of the global or 9.36 per cent of the Asian precipitation) annually as precipitation but it renders home to about 16 percent of world's population on 2.45 per cent of the terrestrial surface (NCIWRD, 1999).

The distribution of this natural resource in India is spatially and temporally so uneven that hydrological extremes of flood and drought are annual events in some or other parts of the country. The same is also applicable to West Bengal, a state of India. When Jalpaiguri district received more than 3470 mm of rain, Purulia, Bankura and Birbum on an average receive 1780mm, 1290mm and

1370mm of rainfall respectively in 2013(IMD, 2014). The south-west monsoon virtually generates more than 80 percent of annual precipitation in West Bengal and that pours during the months of June-September. But a single cloud burst over a place can generate more than 60 percent of annual rain within a span of just 48 hours. Likewise to that of India, the management of spatially uneven and temporally skewed rain-water in West Bengal is the most serious challenge for the water-managers.

Water is indispensable for all economic and social development. However, with ever increasing population growth (13.84% in last decade in West Bengal, Rudra, 2013), industrial growth, water pollution and atmospheric changes, the per capita water availability is shrinking day by day (-14.69% in last decade in West Bengal, Rudra, 2013). The Bengal Delta, which was described as areas of 'excess' water in the colonial document, now suffers from acute dearth of water during lean months (Rudra, 2013). Increasing water demands vis-à-vis growing water scarcity are expected to reduce diversion of water for agriculture in the future. Efforts are needed to utilize the available limited water resources efficiently. Water resources not only provides water for irrigation but also provide water for a range of other productive uses such as livestock, fishing, home gardens, aquatic products and micro enterprises such as brick making. Hence water is crucial for sustainable development

of nation. So, there has been a tremendous pressure on these precious natural resources, mainly due to rapid industrialization and human population. In this direction, works of Roy (2013) and Adhikary (2012) are few to mention.

Keeping in mind, the importance and scarcity of available water resources, it has become imperative to study the behavior of ground water and its fluctuations; model and forecast the fluctuation in advance to help policy formulation for efficient ground water management.

2. Materials and Methods

Bankura, Purulia and Birbhum, three districts falling under Red and laterite Zone of West Bengal are characterized by low precipitation. By and large the districts are very dry and most of the portion are belonging to not so developed and under developed zones, millions of people are maintaining their hard lives. Poor or scarcity in availability of long term consistent data on ground water is the main problem in time series analysis on such parameters. The present study has been under taken with the help of ground water level data on seasonal basis (i.e January, May, August and November) for the period 2005 to 2013 in three districts viz. Bankura, Birbhum and Purulia under drier Red and Laterite zone of West Bengal (http://gis2.nic.in/cgwb/Gems data.aspx, last accessed on 1/12/2014).

Descriptive statistics are useful to describe the patterns and general behavior of a data set; it includes numerical and graphical statistical measure like minimum, maximum, average, standard error, skewness, kurtosis, simple growth rate.

When dealing with more number of variables in regression analysis, particularly when all the variables are not equally important or efficient in describing the relationships, it becomes difficult to decide on which are the variables to be retained in the multiple regression equation. To know the relationship between ground water level and rain fall pattern, regression analysis along with step down is carried out by taking the ground level data as dependent variable and rain fall (RF_1), lagged rain fall (RF_2), time (t) and three seasonal dummy variables (D_1, D_2 and D_3 for January, May and August respectively) as a independent variable. The multiple regression model can be written as:

$$Y = \beta_0 + \beta_1 RF_1 + \beta_2 RF_2 + \beta_3 t + \beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + e$$

$D_1 = 1$, if the month is January

$\quad = 0$, otherwise

$D_2 = 1$, if the month is May

$\quad = 0$, otherwise

$D_3 = 1$, if the month is August

$\quad = 0$, otherwise

$$E\left(Y/D_1 = 1, D_2 = 0, \ D_3 = 0\right)$$

$$= \text{Average groundwater table during January}$$

Similarly for May and so on, where β_0 intercept, $\beta_1, \ldots \ \beta_6$ are regression coefficient corresponding to independent variables RF_1, RF_2, t, D_1, D_2 and D_3 respectively. In case of dummy variable, if we are interested to know the average ground water level in January, dummy variables for the remaining season will be zero.

For the forecasting purpose use of Box-Jenkins methodology, particularly with time series data, are ample in literaturerature; Sahu and Mishra (2013), Mishra *et al.* (2014) a few to mention. In this paper we have used both simple and seasonal ARIMA to forecast the ground water tables in these three districts. Seasonal data for the period 2005-2012 has been used for building forecast model, while data for the years 2013 are taken for model validation. Box et al. (1994) have generalized the ARIMA model to deal with seasonality and have defined a general multiplicative seasonal ARIMA model, commonly known as seasonal ARIMA model. The seasonal ARIMA model described as ARIMA (p, d, q) X (P, D, Q)$_s$, where (p, d, q) non-seasonal part of the model and (P, D, Q) seasonal part of the model with a seasonality S, can be written as

$$\varphi_p\left(B\right)\Phi_P\left(B^S\right)\nabla^d \nabla_S^D Z_t = \theta\left(B\right)\Theta_Q\left(B^S\right)a_t$$

where, $\varphi\left(B\right)$ and $\theta\left(B\right)$ are polynomials of order p and q, respectively; $\Phi\left(B^S\right)$ and $\Theta\left(B^S\right)$ are the polynomials in B^S of degrees P and Q respectively; p=order of non-seasonal autoregressive operator; d= number of regular differencing; q=order of the non-seasonal moving average; P = order of seasonal auto-regression; Q=order of seasonal moving average. The ordinary and seasonal difference components are designated by $\nabla^d = \left(1-B\right)^d$ and $\nabla_S^D = \left(1-B^S\right)^D$ of orders d and D; B is the backward shift operator; d is the number of regular differences; D is the number of seasonal difference; S = seasonality. Z_t denotes observed value at time t, where t = 1, 2, ... k; and a_t is the Gaussian white noise or estimated residual at time t.

Given a set of time series data, one can calculate the mean, variance, autocorrelation function (ACF) and partial autocorrelation function (PACF). The calculation enables one to look at the estimated ACF and PACF which gives an idea about the correlation between observations, indicating the sub-group of models to be entertained. This process is done by looking at the cut-offs in the ACF and PACF. At the identification stage, one would try to match the estimated ACF and PACF with the theoretical ACF and PACF as a guide for tentative model selection, but the final decision is

made once the model is estimated and diagnosed.

Among the competitive models, best models are selected based on minimum value of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Mean Error (ME), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE) and maximum value of R^2 and of course the significance of the coefficients of the models. Best fitted models are put under diagnostic checks through auto correlation function (ACF) and partial autocorrelation function (PACF) of the residual.

$$AIC = 2k - 2\ln(L) \quad BIC = -2*\ln(L) + k*\ln(n)$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right) \quad MPE = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{X_i - \hat{X}_i}{X_i}\right)*100$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \hat{X}_i\right| \quad MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_i - \hat{X}_i}{X_i}\right|*100$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \hat{X}_i)^2}{n}} \quad R^2 = \frac{\sum_{i=1}^{n}\left(\hat{X}_i - \overline{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

where, $X_i, \overline{X}, \hat{X}_i$ are the value of the ith observation, mean and estimated value of the ith observation of the variable X and k is the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model.

3. Results and Discussion

The ground water level in the selected districts seem to be depleting in May month which may be due removal of water for irrigation purpose, more water absorption by plants particularly deep rooted plants and other water evolved activates. Since 2006, among the selected districts, on an average Purulia is found to be more fluctuating district with a standard error of 0.277 meter (Table. 1). Within the districts, ground water level shows more variation in August, November, November with a standard error value of 0.371, 0.229, 0.282 meters respectively in case of Purulia, Birbum and Bankura. Purulia and Bankura districts show the negative growth rates of -2.21 and -0.86 percent respectively. Thereby indicating the decrease in ground water tables which may be due to implementation of ground water recycling/recharging programs like constructions of water cannels, rain water harvesting techniques, improved irrigation methods in recent years.

Table 1. *Perse* description of ground water table in Bankura, Purulia and Birbhum

Bankura					
	January	May	August	November	Average
Mean	4.715	6.198	2.391	3.404	4.177
Standard Error	0.195	0.146	0.279	0.282	0.170
Kurtosis	3.948	-0.238	4.076	3.740	4.434
Skewness	1.869	1.067	1.933	1.409	2.109
Minimum	4.212	5.774	1.663	2.209	3.816
Maximum	6.090	6.945	4.367	5.328	5.398
SGR %	0.490	-1.968	0.326	-3.758	-0.860
Purulia					
Mean	4.760	6.976	2.309	3.667	4.428
Standard Error	0.218	0.195	0.371	0.364	0.227
Kurtosis	1.748	-0.916	3.551	1.680	2.104
Skewness	1.174	0.694	1.795	1.168	1.288
Minimum	4.059	6.346	1.178	2.444	3.751
Maximum	6.155	7.953	4.890	5.985	5.913
SGR %	-2.100	0.014	-4.127	-4.394	-2.212
Birbum					
Mean	4.808	6.381	2.636	3.681	4.377
Standard Error	0.134	0.208	0.255	0.229	0.128
Kurtosis	-1.137	-1.536	-0.978	-0.800	-1.290
Skewness	-0.373	0.005	0.509	0.615	0.181
Minimum	4.140	5.500	1.628	2.839	3.833
Maximum	5.318	7.276	3.894	4.883	4.923
SGR %	1.597	3.091	0.798	-1.080	1.398

The result of stepwise regression is framed in table. 2. The results indicates that in all the three districts ground water level improved significantly as rainfall increases in the respective month. But the significant positive coefficient for remaining all the independent variable indicates that the ground water level is depleting years to come. These regression equation provide the average ground water level for different seasons under given rainfall condition.

From the results of SARIMA modeling (Table.3), one can find that none of the series is seasonally stationary in nature; first order seasonal differencing is necessary for all the series to make these stationary. The fitted model ranges from ARIMA $(0,0,0)(0,1,0)_4$ to ARIMA $(3,1,3)(3,1,3)_4$. Among the competitive models which satisfies the maximum criteria (higher R^2, lowest AIC, BIC, ME, RMSE, MAE, MAPE and MPE) and having significant parameters are retained for further use. The model ARIMA$(1,0,0)$X$(1,1,1)_4$, ARIMA $(2,0,0)$ $(1,1,1)$ and ARIMA$(1,0,1)(2,1,0)$ are found best suited for Bankura, Purulia and Birbhum respectively.

These models are again validated not only by comparing observed and predicted values for the 2013 (four season, Table.4), Figure. 2 shows that observed values are closely matching with those of the expected values during the period

of validation. Now, the above best fitted models are used to get the forecasting values for 2014 (four seasons) and same is presented in table. 4. It can be seen from the table that for Bankura and Purulia the selected SARIMA models have tendency of over estimating while for Birbhum it has got tendency of underestimating. What is alarming from predicted values for 2014 is that in both the districts of Bankur and Purulia ground water table will mostly go further downwards in majority of the seasons while in Birbhum this may improve a bit in all the seasons. Using the above estimated values for ground water levels in these three districts ground water management strategies could be framed for the 2015.

groundwater table depth. While the average value for ground water table shows negative growth rates for Purulia and Bankua but positive for Birbhum district; inter season differences are also there among the districts. Fluctuations in the depths of ground water table can very well be explained with linear multiple regression equation using dummy variables for different seasons. Like other examples, the SARIMA models can very well be explored for modeling and forecasting ground water table in different places. Forecast values shows somewhat different picture than the average behavior of the series. Forecasting value shows that groundwater depth in Birbhum will reduce but for other two districts, this will increase. Thus there is a need for formulation of plans and its execution to arrest the growth in groundwater table depths in Purulia and Bankura.
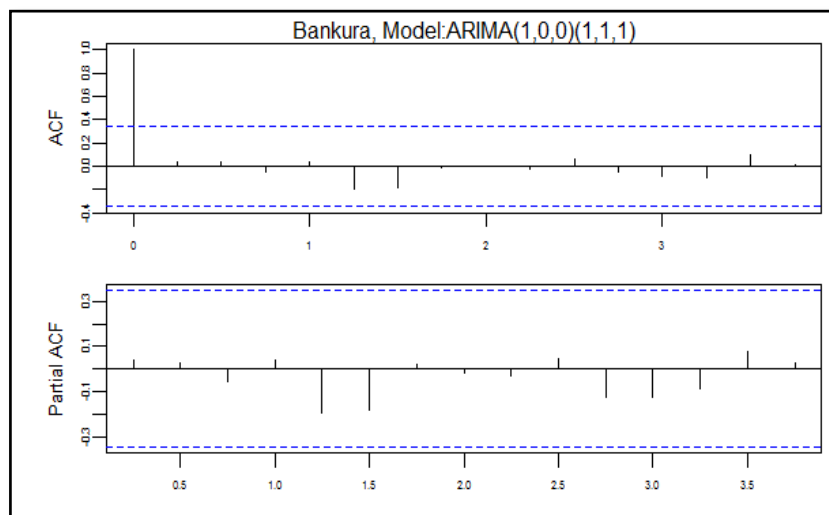
4. Conclusions

From the study one can find contrasting features in

Table 2. Step wise regression model

District	Model	R^2	Significance
Bankura	$Y = 3.141 - 0.0077RF_1 + 1.26D_1 + 3.46D_2 + 1.01D_3 + 0.028t$ (0.002) (0.299) (0.382) (0.746) (0.01)	0.88	0.00
Purulia	$Y = 3.92 - 0.0051RF_1 + 0.97D_1 + 3.52D_2$ (0.001) (0.425) (0.389)	0.82	0.00
Birbhum	$Y = 3.37 - 0.0048RF_1 + 1.04D_1 + 2.95D_2 + 0.026t$ (0.0008) (0.241) (0.220) (0.009)	0.90	0.00

Figures in the parentheses are the respective standard error of the parameter

Table 3. Best fitted SARIMA models

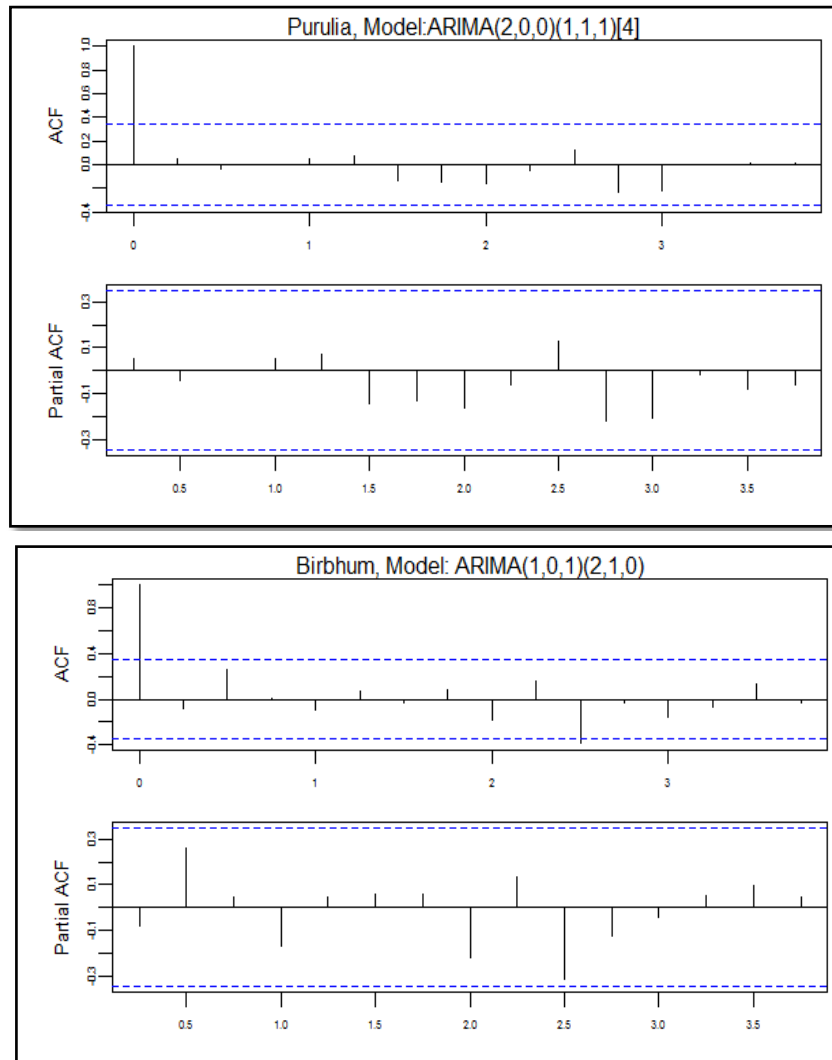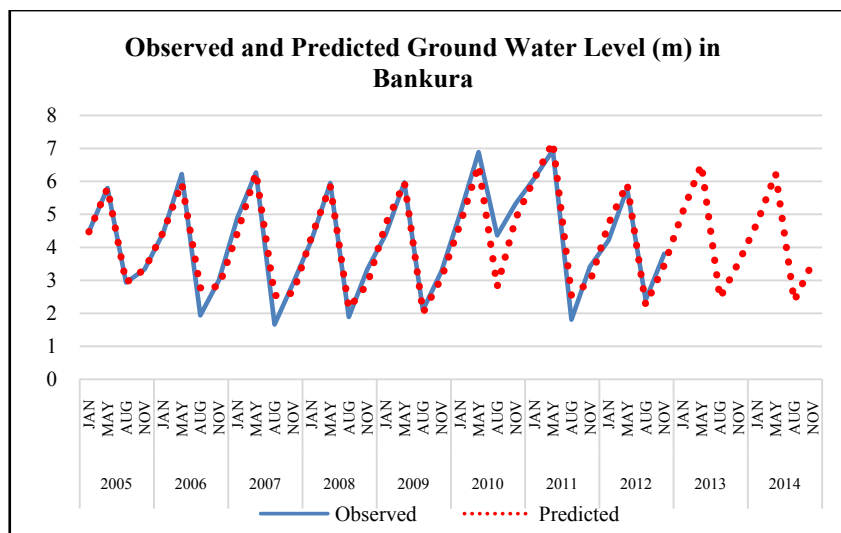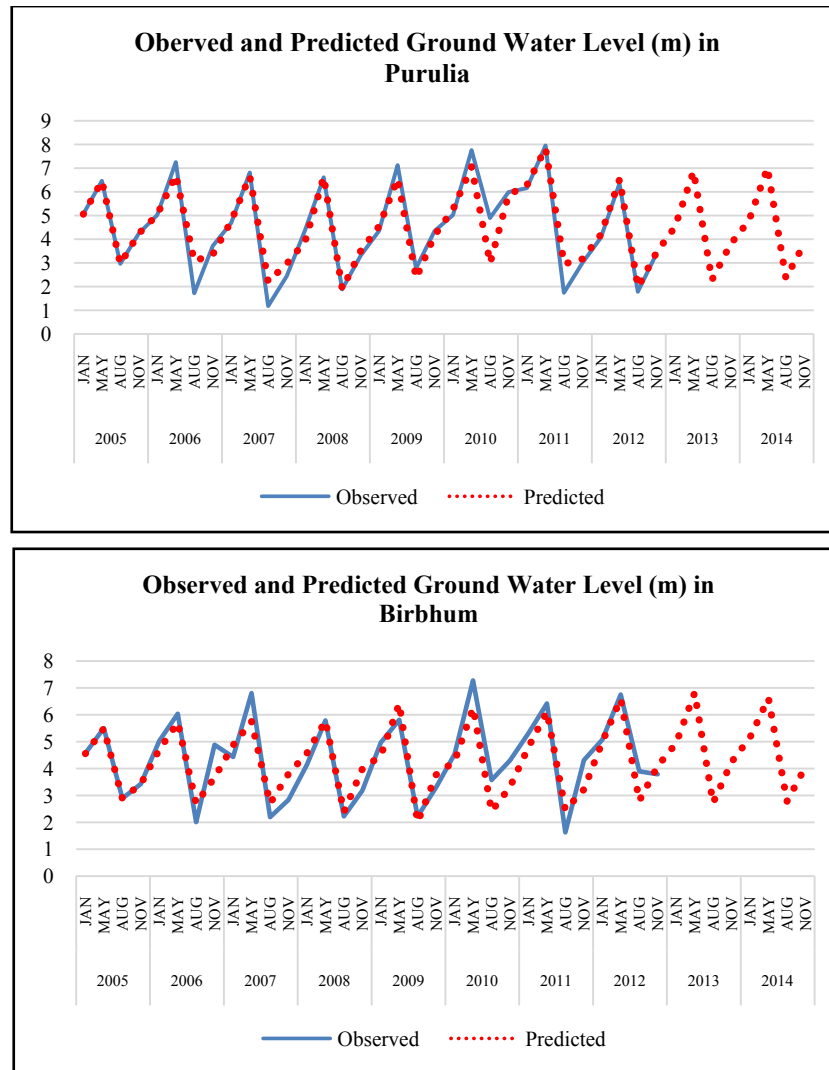| District	Best Fitted Model	AIC	BIC	ME	RMSE	MAE	MPE	MAPE	R^2
Bankura	ARIMA(1,0,0)(1,1,1)	56.38	61.71	0.055	0.448	0.307	-1.119	9.974	0.92
Purilia	ARIMA(2,0,0)(1,1,1)	70.82	77.48	-0.010	0.568	0.356	-6.160	12.978	0.92
Birbhum	ARIMA(1,0,1)(2,1,0)	70.27	76.93	0.096	0.616	0.492	-1.087	13.796	0.83

Figure 1.   Residual ACF and PACF graphs

Figure 2. Observed and forecasted value

Table 4. Model validation and forecasting values of ground water level in meters

Bankura			
Months	2013		2014
	Observed	Predicted	Forecasted
January	4.67	5.08	4.74
May	5.97	6.51	6.23
August	2.41	2.46	2.37
November	2.21	3.54	3.55
Purulia			
January	4.10	4.52	4.78
May	6.47	6.87	7.02
August	1.86	2.28	2.37
November	2.57	3.79	3.83
Birbhum			
January	5.22	4.94	5.16
May	7.03	6.82	6.62
August	3.10	2.73	2.79
November	3.09	4.25	4.07

REFERENCES

[1] Adhikary, S K., Rahman, Md. M. and Gupta, A. D. 2012. A stochastic modelling technique for predicting groundwater table fluctuations with time series analysis. International Journal of Applied Sciences and Engineering Research. 2: 238:49.

[2] Anon. 1999. Integrated Water resource Development / A Plan for Action. NCIWRD, Govt. of India New Delhi.

[3] Anon. 2010. Water and related statistics. Information system organization water planning and project wing, central water commission, New Delhi. P. No.1.

[4] Anon. 2014. IMD, Pune. http://www.imd.gov.in/section/hydro/distrainfall. Last assessed on 12th December 2014.

[5] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. 1994. Time series analysis: forecasting and control. Upper Saddle River, NJ: Prentice-Hall.

[6]   Hamdy, A. and Lacirignola, C. 2005. Coping with water scaricity in the Mediterranean: what, why and how? Mediterranean agronomic institute, Bari-Italy. P.279.

[7]   Mishra P., Sahu P. K. and Uday J. P. S.2014. ARIMA Modeling Technique in Analyzing and Forecasting Fertilizer Statistics in India. Trend in Biosciences, 7(2):170-176.

[8]   Roy, A. 2014. Availability, scaricity and potentiality of groundwater resources in Purulia district of West Bengal: an appraisal. International Journal of Scientific Footprints: 2 (1): 78-93.

[9]   Rudra, K. 2013. The Status of water resources in West Bengal. http://www.indiawaterportal.org/articles. Last accessed on 12th December 2014.

[10]   Sahu PK, Mishra P 2013. Modelling and forecasting production behavior and import-export of total spices. In two most populous countries of the world. J Agric Res 51 : 81—97.