International Research Journal of Finance and Economics
ISSN 1450-2887 Issue 83 (2012)

http://www.internationalresearchjournaloffinanceandeconomics.com

Use of Univariate Time Series Models for Forecasting Cement Productions in India

Purna Chandra Padhan
Assistant Professor, Economics
XLRI, School of Business and Human Resources
P. Box No. 222, C.H. Area (East), Jamshedpur-831001, Jharkhand, India
E-mail: pcpadhan@xlri.ac.in
Tel: +91-657- 398 3191(0), Fax: +91-657 222 7814

Abstract

In recent years, there has been a great deal of discussion on applications of various time series forecasting models and their performance in forecasting business activities. This paper discussed few of these univariate time series forecasting models and their application for forecasting cement productions in India. Applying monthly data spreading over April 1993 to March 2011, on productions of cement in lakh tons, the forecasting performance of various competing models is evaluated through forecast accuracy criteria such as mean absolute percent error (MAPE), mean squared deviation (MSD) etc. Among these models Seasonal Autoregressive Integrated Moving Average (SARIMA) model performs better than other competing models in forecasting cement productions. It provides lowest MAPE value. In fact, the model has also an advantage over other models as it explicates autoregressive and moving average process the data along with seasonality. Therefore, as a policy implication, SARIMA model can be used for forecasting cement productions in India.

Keywords: **Cement, forecasting models, smoothing, seasonality, SARIMA model**
JEL Classification Codes: C53, C27, L11, L61

1. Introduction

With the passage of time and development of human civilizations the societies have become extremely prone towards seeking for comfortable and luxurious life[1]. In a modern economy, although globalization and liberalization policy have improved the standard of living of many people, in India large chunk of people are still living below poverty line. They are not even able to meet their basic necessities. It's a day dream for them to think about such an elite life style. However, the other part of the story is very rosy. The so-called middle and upper class people have passion for comforts and thriving for luxurious life. Thus demand for such items has amplified substantially. Today society has been dominated by varieties of lucrative and fashionable goods and services. Having one's own house, car and other household accessories has become an aspiration for every single individual, which has resulted in booming various interrelated sectors, including the demand for real estate market. Another

[1] History witnessed a lot about the invasion and extension of empire by different kings, rulers across globe in order to increases their wealth, supremacy to lead a luxurious and comfortable life.

side is that there has been tremendous improvement in infrastructure facilities in terms of roads, railways, airports, bridges, canal, buildings, IT sector etc. in India. Cement, being a related product to this is quite affected by such development. Although it's not an end product itself, the demand/sales of cement depends upon many factors such as its price, substitute goods, growth of the economy, growth of subsidiaries industries, infrastructure, licensing policy, so on and so forth. All these factors have augmented the demand for cement over the years. Thus forecasting sales/demand/productions of cement has become a challenging task. Although several methods, both qualitative and quantitative techniques, are available in the literature, an understating of complete scenario behind cement sales is essential to do the job in a better way. In such situation often question arises that which forecasting techniques would one use to forecast cement demand in a better way? The present paper tries to answer this question partly by using different forecasting models to forecast the demand/sales/productions of cement in India.

The rests of the paper are designed as follows. In section II review of literature is given. In section III an overview of cement Indian industry is highlighted. Section IV discusses about various univariate time series forecasting models. The estimated results and analysis are discussed in section V. Conclusion is provided in section VI.

2. Review of Literature

There has been a great deal of discussion in economic literature about applications of various forecasting models for forecasting desired issues. Several time series forecasting techniques such as naive models to ARCH- GARCH and further advanced model have been applied to explain forecasting performance of preferred variables. However there are paucities of studies with specific reference to cement industry. Few noteworthy studies relating to cement industry and other related industries many be highlighted here. In a study, Pei Liu et.al (2008) establishes quarterly and monthly cement forecast for Taiwan with seasonal ARIMA and artificial neural network (ANN) models. With monthly data since January 2004 to March 2005, the study verifies the accuracy of each forecast model. In an attempt Ramani and Dholakia (1999) have suggested a computer based decision support system to forecast cement demand in India. For forecasting productivity, demand and sales of different products there are ample evidence of applications of various time series techniques. While Ghosh (2008) applied univariate time-series techniques such as Multiplicative Seasonal Autoregressive Integrated Moving Average (MSARIMA) and Holt-Winters Multiplicative Exponential Smoothing (ES) for seasonally unadjusted monthly data spanning from April 2000 to February 2007 to forecast the monthly peak demand of electricity in the northern region of India. The study found that the MSARIMA model outperforms the ES model in terms of lower root mean square error, mean absolute error and mean absolute percent error criteria. Mandal (2005) have applied ARIMA model for forecasting sugarcane productions in India with annual data from 1950-51 to 2002-03. Wankhade et. al. (2010) have considered ARIMA model for forecasting pigeon pea production in India with annual data from 1950-51 to 2007-08. Through ARIMA model Iqbal et.al (2005) forecasted area and production of wheat in Pakistan. Using uni-variate time series analysis, Rothman (1999), Johnes (1999), Proietti(2001), Gil-Alana(2001), have forecasted the employment rate in US and UK.

3. An Overview of Indian Cement Industry

Cement is one of the most significant commercial products, used mainly for constructions works. The industry heavily depends upon three sectors namely, coal, power and transport. India is the largest producer of cement in the world only next to China. It has a total capacity of more than 12.5 million tons annually and the industry has been performing pretty well. Although, the global demand for cement is rising 4.1% annually, the recent US subprime crisis and global economy meltdown affected the production, sales, and foreign trade partially. With government initiative and growth of the economy along with growing infrastructure projects, housing facilities, road networks, educational projects, sports activities, booming real estate market, increased populations etc., the Indian cement Industry is flourishing and growing at a prolific pace. It is also expected to grow at a faster pace in near future due to several pending and upcoming projects in various functional areas. In terms of numbers, the industry comprised several big and small companies. Approximately 137 large and 365 mini cement plants are there across India which is again the second largest to China. Out of these 20 companies are the leading companies in terms of value and volume of production. Few of them are ACC Cement Ltd., Gujarat Ambuja Ltd., Aditya Birla, Ultratech Cement, Shree Cement etc. which accounts almost 70 % of its total cement productions.

While opening the pages of history, one can find that origin of cement industries in India goes dated back to 1889. The industry was started in Kolkata with a small domestic company who started manufacturing argillaceous. The remarkable growth of cement productions had begun since early 1914 with the then Indian Cement Company (ICC) Ltd whose plant was located in Porbunder, Gujarat started manufacturing cement production. The ICC had an installed capacity of 1000 and productions capacity of 10, 000 tons annually. Immediately after its set up, Government started regulating volume and value of productions as well as distributions. The real breakthrough in terms of productions, manufacturing and installed capacity has begun since World War I whereby demand for Indian cement had increased immensely due to war reconstructions. In recent years the industry witnessed excellent growth owing to growing subsidiaries sectors. According to RNCO's research report on Aug 2008, on ' Indian Cement Industry Forecast to 2012' the countries production of cement on annual basis will grow at 12% (CAGR) during 2011-12 to reach 303 million metric tons. The industry employs more than 120,000 people as per the Cement Manufacturers Association Report 2008 is concerned. These reports predict that, Indian Cement Industry has achieved an installed capacity of 242 million tons and is targeted to reach 300 million tons by 2011-12 and 600 million by 2020. It has 97 per cent of the installed capacity through dry process. Indian cement industry is efficient and eco-friendly and in terms of technology front it has largely adopted state-of-art manufacturing technologies.

Over the years, production and sales of cement shows tremendous growth although sales are seasonal. Cement sales are generally high during summer and low during rainy season, as most of the construction work materializes during summer. The industry is also picking up very well as government have taken several steps for infrastructure development, boom in real estate markets in both rural and urban areas and growth of other subsidiary areas.

4. Forecasting Models

Forecasting implies what will happen the desired event in future, which involves lots of uncertainties. It is a process of studying and analyzing past and present to have a saying over future. Business firms and industries are more concerned about their future while taking various strategic and operational decisions to continue business and earn maximum profit through efficient utilization of resources. They are more concerned about the sales, productivity, human resources. Sales forecasting is one of the most common phenomena observed in industry, as it assists other subsidiary department of the industry such as finance, human resources, marketing, supply chain etc. Sales forecasting, either volume or value, receives top priorities. Although forecasted values are obtained through several qualitative and quantitative methods, each method has its own pros and cons. The selection of these models depends upon the knowledge, availability of data and context of forecasting. A firm/an industry, like cement, can make use of multifarious forecasting models of their interest. This paper applies few of these univariate time series forecasting model to forecast future demand of cement industry. Since time series analysis generally forecast future values on the basis of past data, the following models have been applied.

4.1. Naive Models

The simplest model, which considers immediate past values could be the future forecasted value, is naive model. To incorporate seasonality and trend in the data sets, naive model can also be specified with seasonality and trend in the model[2]. Large number of data point is not required while forecasting, rather only one data is good enough. Full weightage is given to immediate past values to represent in future. This method does not capture past information precisely. It might not be a suitable method for forecasting any time series data which has different time series features in data.

Naive Model: $\hat{Y}_{t+1} = Y_t$ (1)

4.2. Simple and Moving Average Models

The most commonly and widely used method for forecasting any historical data is simple average. The forecasted values for the next period will be nothing but sum of all items divided by number of observations is called average value. Minimum data point required for this method is two. Because of its simplicity many industry professionals use this method. However, it has few advantage and disadvantage. If we look at disadvantage, two most important are that average value is highly dominated by extreme values and it gives equal weightage to each and every observations.

Simple Average $\hat{Y}_{t+1} = \frac{1}{t}\sum_{i=1}^{t} Y_i$ (2)

The other form of average which could be suitable is moving average model. In the moving average process, forecasted values obtained by adding newest and dropping oldest observations recursively of order k where k is the number of moving average process. Minimum data point required is of order k where k is any finite positive number. Except random if the data has other three properties of time series such as seasonality, trend or cyclicality, then moving average model might be suitable for forecasting. Mathematically, the model may be expressed as

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-k+1}}{k} \quad (3)$$

Double moving average process is nothing but moving average of first moving average model. The main idea of double moving average is to smooth the data in a much better way. The double moving average model can be specified as

$$M_t = \hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-k+1}}{k}. \quad (4)$$

$$M_t' = \hat{Y}_{t+1} = \frac{M_t + M_{t-1} + \ldots + M_{t-k+1}}{k} \quad (5)$$

The adjustment parameter is $\alpha_t = M_t + (M_t - M_t')$
$= 2M_t - M_t'$

and $\beta_t = 2/(k-1)(M_t - M_t')$

The p period ahead forecast values is $\hat{Y}_{t+p} = \alpha_t + \beta_t P$ (6)

4.3. Exponential Smoothing Models

Exponential smoothing model is a procedure which continuously revises the forecasted values in the light of more recent experience. This model is extensively used in forecasting literature. In simple exponential smoothing model recent observations gets higher weightage and distance past gets lower weightage. The beauty of the model is that weights decline geometrically for which it's named as

[2] Naive Trend Model: $\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1})$, Naive Seasonality Model: $\hat{Y}_{t+1} = Y_{t-k+1}$ and Naive Trend with seasonality model is $\hat{Y}_{t+1} = Y_{t-k+1} + \frac{Y_t - Y_{t-k}}{k}$ where k being the number of seasonal factor

exponential smoothing model. Except trend the model is suitable for other three different forms of time series properties. Since total weight is equal to 1, the weight of each and every observation is restricted within 0 to 1. If it is desired that prediction be stable and random variations smoothed, a small value of weighted parameter namely 'alpha' is required, whereas if a rapid response to a real change in the pattern of observations is desired, then larger value of 'alpha' is appropriate. One method of estimating the weight is lowest mean squared error (MSE) value.

The Holt's and Winter's methods are the two extended version of simple exponential smoothing models. The Holt's method is suitable with data having small local trend. If data has non-seasonal pattern but revolves around local linear trend, the Hol'ts linear exponential smoothing method is suitable. To estimate future values it requires initial estimate of a current level and slope of trend. On the other hand, the Winter's method is suitable when data has both seasonal and trend along with random pattern. This is a broader smoothing model, which requires an estimation of the initial smoothed values, a trend estimate and seasonality. The models may be expressed as,

Simple exponential smoothing model $\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$ (7)

Holt's linear exponential smoothing

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (8)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (9)$$

$$\hat{Y}_{t+p} = L_t + pT_t \quad (10)$$

Winter's multiplicative seasonal method

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (11)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)(T_{t-1}) \quad (12)$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma)S_{t-s} \quad (13)$$

$$\hat{Y}_{t+p} = (L_t + pT_t)S_{t-s+p} \quad (14)$$

Where, L_t is new smoothed value or current level estimate, a is smoothing constant for level, Y_t is new observation or actual value in period t, β is smoothing constant for trend estimate, T_t is trend estimate, γ is smoothing constant for seasonality estimate, s is seasonal estimate, p is number of forecasting period ahead, s is length of seasonality, $Y\hat{}_{t+p}$= forecast for p periods into future.

4.4. Seasonal Decomposition Models

When time series data has all the four patterns such as random, trend, seasonal, and cyclical, then seasonal decomposition method could be used for decomposing them into four different patterns. In this process each component is estimate separately and then forecasted separately. The final forecasting is estimated by combining all of these four parts. It is possible to determine what portion of sales changes represents an overall change in demand and what portion is due to seasonality in demand. There are two different versions of seasonal decomposition models namely additive and multiplicative models[3]. The multiplicative model is used when variability of time series being analyzed is increasing, while the additive model is used when the variability of time series over a period of time is constant. The multiplicative seasonal decomposition model may be expressed as

$$\hat{Y}_{t+1} = \hat{T}_{t+1} \times \hat{S}_{t+1} \times \hat{C}_{t+1} \times \hat{I}_{t+1} \quad (15)$$

4.5.Curve Fitting

Another interesting class of model, which entirely depends upon time being independent variable and data under consideration is dependent variable, is called as curve fitting. This is estimated through

[3] Additive Seasonal Decomposition model may be expressed as $\hat{Y}_{t+1} = \hat{T}_{t+1} + \hat{S}_{t+1} + \hat{C}_{t+1} + \hat{I}_{t+1}$

regression model. Such models are both linear and non-linear. They may be called as linear trend, quadratic trend, and exponential trend and can be expressed as

Linear Trend Model: $\hat{Y}t = \hat{a} + \hat{b}1T$ (16)

Quadratic trend model: $\hat{Y}t = \hat{a} + \hat{b}_1 T + \hat{b}_2 T^2$ (17)

Exponential trend model: $\hat{Y}_t = \hat{a}.e^{\hat{b}1T}$ (18)

4.6. ARIMA and SARIMA Regression Models

The autoregressive integrated moving average, ARIMA (p,d,q), model popularly known as Box-Jenkins (1970) forecasting model is a regression based model in which forecasted values are obtained by regressing past values of the variable itself and current and past values of error term with different lag length. The ARIMA model combines autoregressive and moving average process with integration of order d, which should be stationary. The application of the model requires four steps, such as identification, estimation, diagnostic checking and forecasting. As the model is very powerful and useful with stationary data, the stationary of the series can be tested with correlogram or unit root tests. After verifying stationary properties, it is essential to find out highest order of autoregressive process 'p' through partial autocorrelations function and highest order of moving average process 'q' through autocorrelation function. The same can be verified with alternative model selection criteria such as Akaike Information Criteria, Schwarz Bayesian Criteria, Adjusted R^2, and Final Prediction Error and so on. Once the order of p and q selected the specified regression model is estimated with maximum likelihood estimation procedure. The next step is to go for diagnostic checking whether the estimated results are correct or not. It's obtained through above mentioned model selection criteria's. Also it is verified with ACF and PACF obtained from the residual of the specified ARIMA model. If the residual is free from all classical assumption of the regression model and stationary then the model is correct. Additionally, diagnostic checking has been performed with $\chi 2$ and Ljung-Box Q statistics. Lastly obtain the forecasted values by estimating the regression model with different forecasting period ahead recursively.

Autoregressive moving average mode $Y_t = \alpha + \varphi_1 Y_{t-1} + \ldots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$ (19)

Forecast of ARIMA $\hat{Y}_{t+1} = \hat{\varphi}_1 Y_t + \hat{\varphi}_2 Y_{t-1} + \hat{\theta}_1 t + \hat{\theta}_2 \hat{\varepsilon}_{t-1} + \hat{\delta}$ (20)

Where a is the intercept term, ϕ is the parameter of autoregressive process, θ is the parameters of moving average process,

The seasonal autoregressive integrated moving average, SARIMA (p, d, q)(P,D,Q)s model is nothing but an extension of ARIMA model with seasonal data. Here the abbreviations stands for order of; p for regular autoregressive term, d for regular difference, q for regular moving average process, P for seasonal autoregressive term, D for seasonal difference at seasonal lag, Q for seasonal moving average term and s is for seasonal order. Thus SARIMA model encompasses both regular autoregressive and moving average terms that account for the correlations at lower lags and seasonal auto regression and moving average terms that account for the correlation at the seasonal lags. The estimation procedure of SARIMA model is as similar with ARIMA model. Again here the order of p, d, q etc. are selected with same AIC, SBC, FPE criteria.

4.7. Evaluation Criteria

Forecast error (e_t) is calculated as difference between actual and forecasted values. The lowest forecast error provides best model. Thus, the evaluation criteria for selecting appropriate forecasting model with a set of models are mean absolute deviation (MAD), mean squared error (MSE) and mean absolute percent error (MAPE). They are defined as,

$$MAD = \frac{1}{n}\sum_{t=1}^{n} | e_t |$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(e2_t)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{Y_t}$$

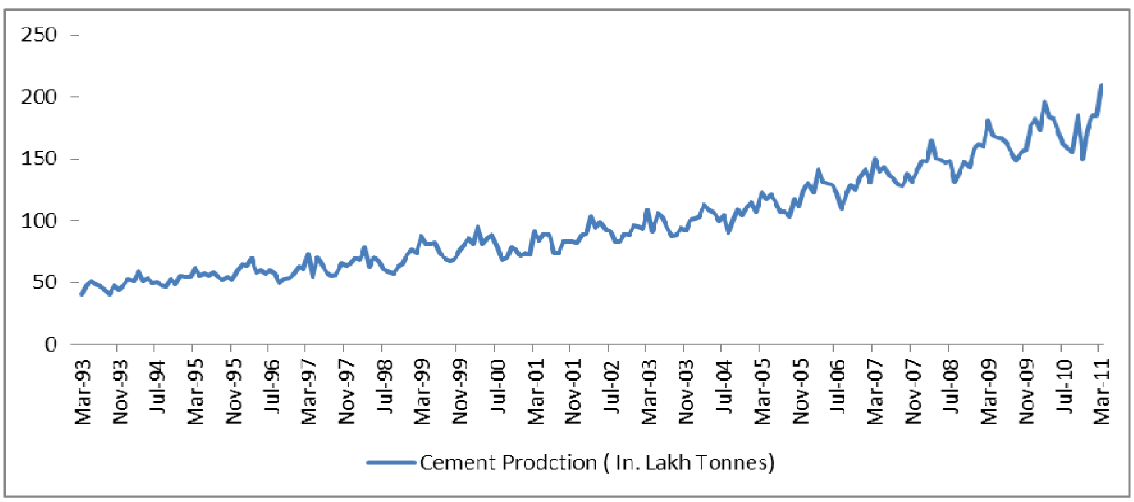
Where $e_t = Y_t - \hat{Y}_t$. MAPE is the best measure as forecast error is expressed in percentage term.

5. Estimated Result Analysis

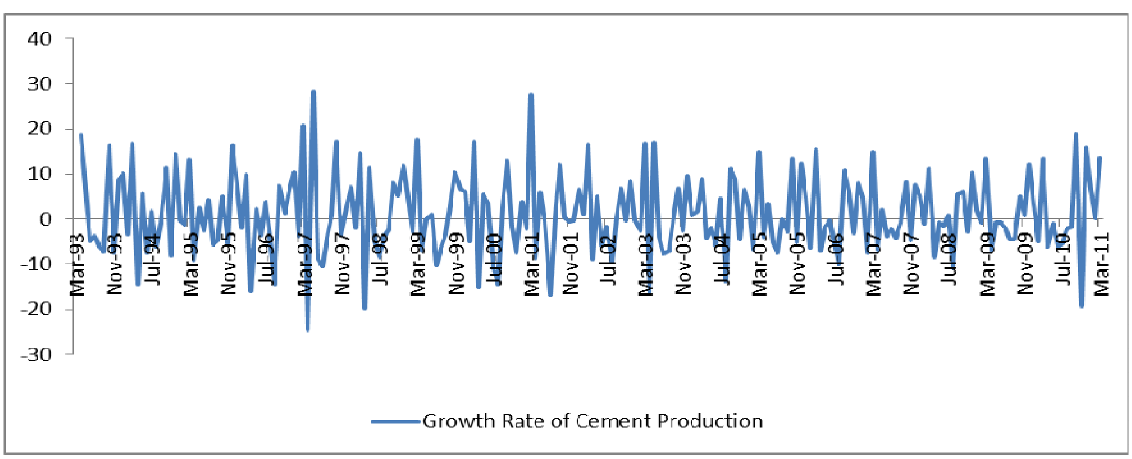
We have selected cement industry for forecasting cement productions. The essence of selecting this industry is that market for cement is growing very fast due to boom in real estate sector, infrastructure development and above all economic growth. The monthly data from March 1993 to March 2011 is collected from Center for Monitoring Indian Economy (CMIE), Prowess database Mumbai. The unit of measurement of cement productions is lakh tons.

While Graph 1 represents the time series plot of cement productions in lakh tons, graph 2 shows the monthly growth rate of it. From the graph it is evident that cement productions shows increasing trend along with seasonal pattern. Further seasonality and trending pattern is verified with the autocorrelations function of data sets. The growth rate of cement productions is quite volatile over a period of time showing upswings and downswings in it. Sometimes growth rate is positive and very high while other times it's negative.

Graph l: Productions of Cement (in Lakh tones)



Graph 2: Growth Rate of Productions of Cement



Further, descriptive statistics of cement productions is displayed in table 1 which exhibits some preliminary understanding about nature of the data. It shows that during 2010 total production of cement was 2070.86 lakh tons which was highest. As indicated from graph, total productions have shown increasing trend with increasing variability between highest and lowest values within a year. The production is generally highest during peak season i.e. March-April and lowest during offseason i.e. Aug-Sept, as most of the construction work takes place during peak season and not during rainy and harvesting season. In terms of distributions, the series is also not normally distributed as the values of skewness and kurtosis is far from 0 and 3 respectively for the entire period.

Table 1: Descriptive Statistics

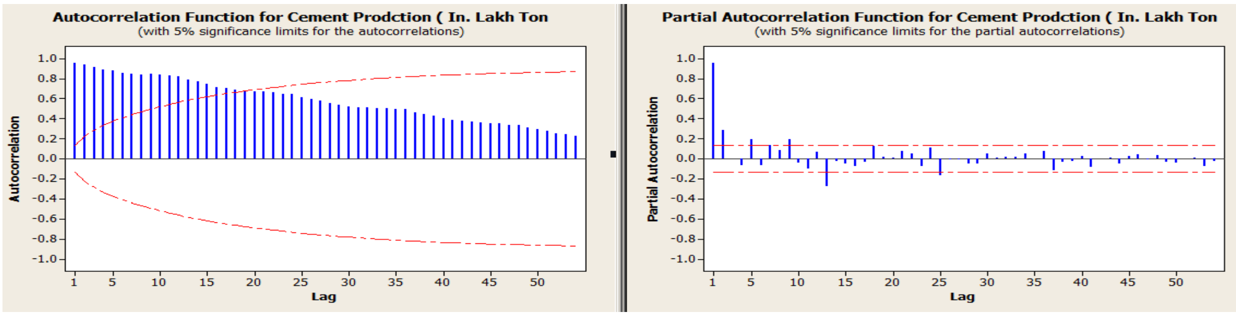
Des Stat.	Mean	Med	Std. Dev.	Variance	Kurtosis	Skewness	Min.	Max.	Sum	N
1993	46.15	47.03	3.02	9.13	-0.07	-0.27	40.67	51.11	461.45	10
1994	51.27	50.60	3.48	12.08	1.04	0.90	46.87	59.08	615.25	12
1995	55.65	55.04	2.89	8.35	0.03	0.59	51.56	61.18	667.83	12
1996	58.42	57.90	5.54	30.65	0.36	0.31	49.00	69.41	701.08	12
1997	62.29	62.90	5.80	33.66	-0.62	0.36	54.90	72.80	747.50	12
1998	66.03	65.94	6.09	37.04	-0.18	0.46	57.36	78.10	792.36	12
1999	76.34	76.00	6.37	40.54	-0.82	-0.02	66.55	87.37	916.09	12
2000	79.89	79.84	8.15	66.48	-0.25	0.17	67.45	95.32	958.70	12
2001	81.46	82.73	6.70	44.94	-1.22	-0.10	72.12	92.00	977.57	12
2002	91.36	90.13	6.12	37.45	-0.04	0.47	82.78	103.45	1096.35	12
2003	96.10	94.20	6.89	47.49	-0.62	0.66	87.50	108.77	1153.18	12
2004	104.15	104.25	6.07	36.85	1.66	-0.92	90.00	112.60	1249.80	12
2005	113.89	114.75	7.04	49.56	-1.31	0.01	103.45	124.67	1366.69	12
2006	127.05	128.72	8.04	64.62	1.49	-0.49	109.42	141.48	1524.64	12
2007	136.81	137.64	6.64	44.07	-0.28	0.3 1	127.55	149.93	1641.69	12
2008	147.80	147.74	8.13	66.16	1.37	0.08	131.80	163.87	1773.60	12
2009	163.23	161.91	9.21	84.75	0.14	0.50	148.30	181.11	1958.70	12
2010	172.57	173.00	14.04	197.13	-0.82	-0.15	149.28	196.00	2070.86	12
2011	192.70	184.51	14.41	207.68	-0.23	1.73	184.25	209.34	578.10	3
Total	97.93	88.43	40.31	1624.73	-0.63	0.62	40.67	209.34	21251.4	217

Here Med implies Median, Std. Dev. implies Standard Deviation.

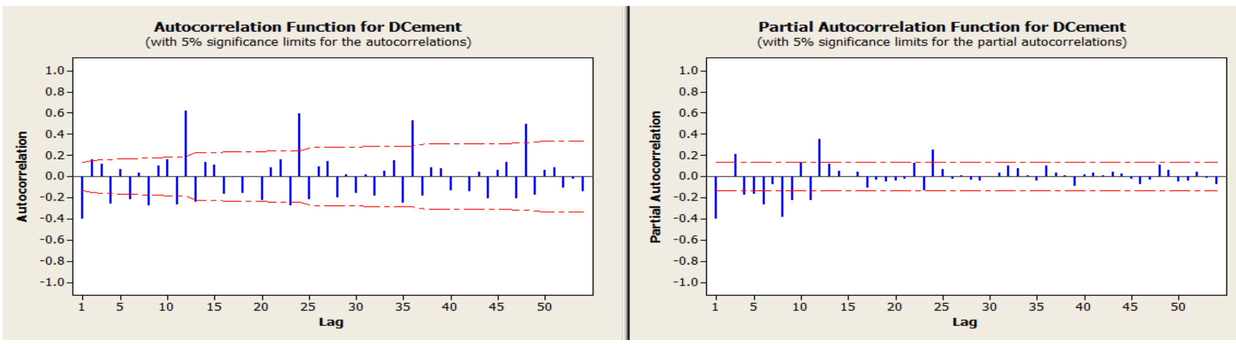
The autocorrelation and partial autocorrelation function are used to identify the nature of data whether they follow any systematic pattern or not of the time series analysis. It also helps us to identify whether series is stationary or not[4] . The ACF and PACF graph for cement productions at level and fist differences are represented in graph 3a and 3b. Autocorrelation coefficient dies out slowly and statistically significant at seasonal interval for the series at 5% significance level. This indicates that the series is non-stationary and follows trend and seasonal pattern. While with first difference, the ACF of cement sales dies out quickly although at certain seasonal interval they are statistically significant at 5% significance level which indicates that the series is stationary. Similarly, a new series is generated from the original series by considering seasonal difference of order s=12 (since it's a monthly data) and then ACF and PACF are obtained for the newly deseasonalised series. The result shown in graph 3c indicates that new series is still non-stationary as ACF dies out slowly, but its first difference is stationary as represented in graph 3d.

[4] The Autocorrelation measures the correlation between some value of a series (e.g., Y) and the value of that series at immediate lag. The Partial Autocorrelation measures the *additional correlation* between some value of a series (e.g., Y) and the value of that series at some lag, which is not accounted for by the next shorter lag

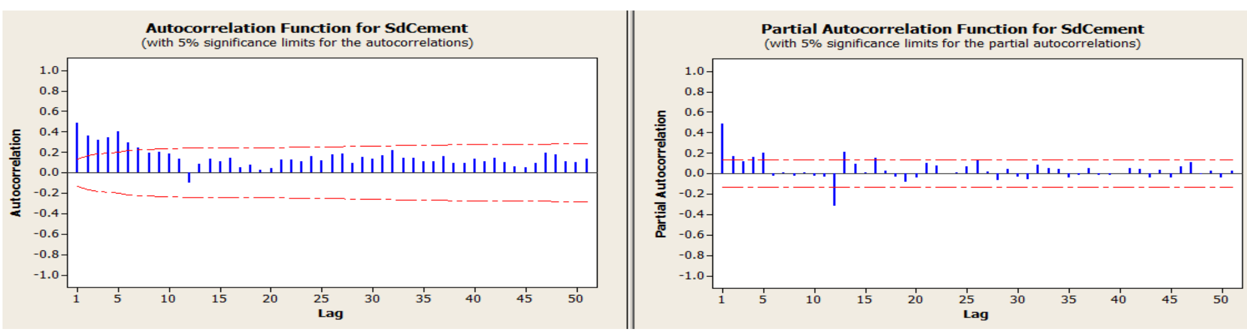
Graph 3a: ACF and PACF of Cement Productions



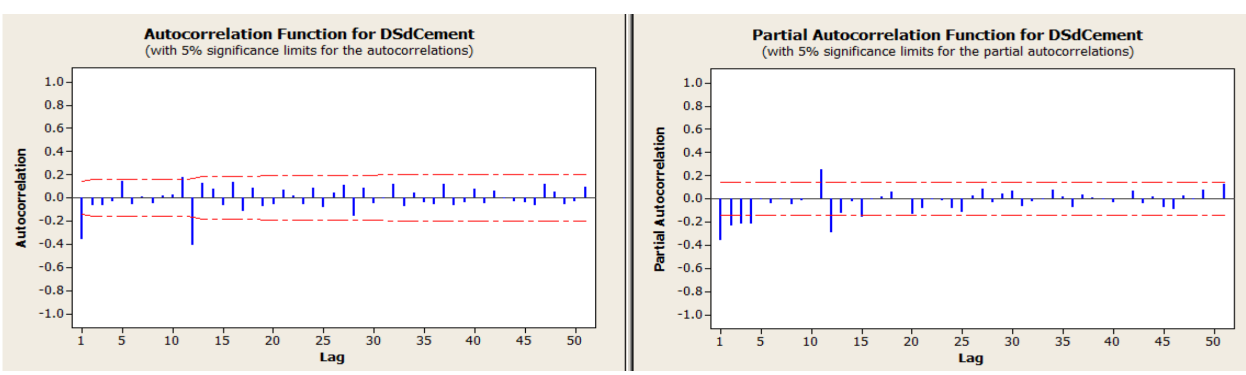
Graph 3b: ACF and PACF of Cement Productions with First Differences



Graph 3c: ACF and PACF of Cement Productions Seasonal Differences (lag 12) (Level)



Graph 3d: ACF and PACF of Cement Productions Seasonal Differences (First Difference)



Further, stationary and non-stationary property of the data set is verified with Phillips-Perron (1988) tests reported in table 2. The regression model is expressed both including a constant and, constants with trend term as well. It reveals the same result. The series is stationary at level when the model consists of trend. It is stationary at first difference too.

Table 2: Unit Root Tests (Phillips-Perron Tests)

Variable	Level		First Difference	
	C	CT	C	CT
Cement	0.60(20) (0.996)	5.232(0) * (0.000)	-30.224(9) * (0.000)	-35.402(10)* (0.000)
SCement	8.662(6)* (0.000)	-9.828(5) * (0.000)	-87.907(80) * (0.000)	-87.743(84) * (0.000)

*Implies 1% significance level, SCement means Cement sales with 12 period lag differences. C for constant Term and T for Trend.

The forecasted values are reported in table 3a and 3b[5]. Among the class of Naive model, simple Naive model provides the best forecasted values with lowest MAPE of 6.80. It is also supported by lowest mean absolute deviation (MAD) and mean squared deviation (MSD). The forecasted value for next period will be the current value. Similarly among the class of averages, three period moving averages provide better forecasted values with lowest error equivalent to 7.15 of MAPE, 6.65 of MAD and 72.59 of MSD. The MA (3) model provides forecasted values at best 3 forecasting period ahead. The forecasted values for next three periods of April, May and June 2011 are 192.70, 196.93 and 209.34 respectively. To select the best model among various trend lines, the quadratic trend line provides better forecasted values with lowest MAPE of 6.77.

Seasonal decomposition model generally applied when the data has all three different patterns of time series such as seasonality, trend and random pattern. If the variability between peak and trough point is increasing then multiplicative seasonal decomposition model can be applied while for constant variability additive model may be used. The model may be estimated with seasonal only and/or with seasonality and trend. Considering the alternative combinations of seasonal model specification, the multiplicative seasonal decomposition model including seasonal and trend term provides better forecasted values as it offers lowest MAPE, MAD and MSD values of 6.80, 6.54 and 76.13 respectively.

Similarly for exponential smoothing models, the selection of smoothing parameter is very critical. The smoothing values *'alpha'* is selected again based on lowest MAPE, MAD and MSD values. For forecasting with simple exponential model the value of smoothing parameter alpha is selected as 0.7 which provides better forecasted values. With this model forecasted value for April 2011 is 201.44 Lakh tons[6]. Similarly, Holt’s method is estimated with smoothing parameter alpha and beta is equivalent to 0.7 and 0.1 which provides lowest error. Similarly Winter's exponential smoothing model selects the smoothing parameter for level as 0.7, for trend it's 0.1 and for seasonality its 0.1 based on lowest MAPE value. The corresponding forecasted values are reported in table 3b. Among these classes of models simple exponential smoothing model provides the best forecasted values.

The necessary steps for applying ARIMA model such a specification, estimation, diagnostic checking and forecasting has been systematically followed[7]. Since the data is non-stationary at level and stationary at first difference, the model is estimated with the order of ARIMA (1, 1, 3). The forecasted values for next two years ahead are obtained which provides MAPE of 7.94 and MAD of 5.55. This is the optimal combinations of ARIMA model. Next SARIMA model is applied. Since data has seasonal pattern and deseasonalised data is non-stationary at level but stationary at first difference, optimal SARIMA (1, 0, 3) (3, 1, 1) 12 model is selected based on lowest MAPE values and is free

5 Due to space problem the table has been splited into two tables.
6 With other values of smoothing parameters, the MAPE, MAD and MSD is not reported due to space consumption. The results may be obtained from the author upon request.
7 For detail of the steps of ARIMA model one can refer any time series textbook or standard refereed article

from all sorts of problem of diagnostic checking[8]. Among all of these estimated forecasting models, SARIMA (1, 0, 3) (3. 1, 1) 12 model provides the best forecasted values as it has lowest MAPE value of 3.41. The corresponding MAD and MSD is 3.17 and 19.46 respectively which is lowest among these classes of model. Thus SARIMA forecasting model can be applied for forecasting cement productions in India. As per the estimated results the forecasted values for cement productions for the month of June, July and August 2011 will be 187.32, 184.33 and 175.19 lakh tons.

Table 3a: Forecasting Model Specifications

Forecast Period	Naive Model	Naive Trend	Naive Seasonal	Naive TS	Average	MA(3)	DMA (3)	Linear Trend	Exp. Trend	Quadratic Trend
Apr-11	209.34	234.2	183.97	185.08	98.2	192.70	208.0	164.99	180.86	183.70
May-11			182.19	184.30		196.93		165.61	182.02	184.83
Jun-11			171.07	155.89		209.34		166.23	183.19	185.97
Jul-11			161.85	147.59				166.84	184.37	187.12
Aug-11			158.27	144.78				167.46	185.55	188.26
Sep-11			155.44	142.25				168.07	186.74	189.41
Oct-11			184.90	171.95				168.69	187.94	190.57
Nov-11			149.28	133.87				169.30	189.14	191.73
Dec-11			172.99	160.55				169.92	190.35	192.90
Jan-12			184.25	169.83				170.54	191.57	194.07
Feb-12			184.51	169.16				171.15	192.80	195.25
Mar-12			209.34	193.96				171.77	194.04	196.43
Apr-12				191.90				172.38	195.28	197.61
May-12								173.00	196.53	198.80
Jun-12								173.61	197.80	200.00
Jul-12								174.23	199.06	201.20
Aug-12								174.85	200.34	202.40
Sep-12								175.46	201.62	203.61
Oct-12								176.08	202.92	204.82
Nov-12								176.69	204.22	206.04
Dec-12								177.31	205.53	207.26
Jan-13								177.93	206.85	208.49
Feb-13								178.54	208.17	209.72
Mar-13								179.16	209.51	210.96
MAPE	6.80	11.31	8.10	7.63	25.71	7.15	12.50	10.69	6.78	6.77
MAD	6.54	10.74	8.25	7.80	30.71	6.65	11.92	9.32	6.43	6.36
MSD	76.13	210.3	96.03	87.00	1608.7	72.59	216.9	137.02	70.87	68.94

Here MA stands for moving average; DMA stands for double moving average, TS stands for Trend and Seasonal

Table 3b: Forecasting Model Specifications

Forecast Period	SD MTS	SD MS	SD ATS	SD AS	Exponential Smoothing	Holts Method	Winter Method	ARIMA Model	SARIMA Model
Apr-11	162.64	96.79	163.68	96.89	201.44	205.27	200.13	196.18	193.84
May-11	165.94	98.39	165.89	98.48	201.44	208.53	206.11	198.41	191.54
Jun-11	166.41	98.30	166.46	98.44	201.44	211.79	209.79	191.67	187.32
Jul-11	168.56	99.21	167.56	98.92	201.44	215.05	215.83	188.58	184.33
Aug-11	165.07	96.79	166.15	96.89		218.31	212.16	187.33	175.19
Sep-11	168.41	98.39	168.35	98.48		221.56	218.31	187.02	174.16
Oct-11	168.87	98.30	168.92	98.44		224.82	222.03	187.18	188.31
Nov-11	171.05	99.21	170.02	98.92		228.08	228.24	187.58	177.34
Dec-11	167.50	96.79	168.61	96.89		231.34	224.19	188.10	197.41

8 Here the details of the estimated results may be obtained upon request.

Table 3b: Forecasting Model Specifications - Continued

Jan-12	170.88	98.39	170.81	98.48		234.60	230.52	188.69	203.37
Feb-12	171.34	98.30	171.38	98.44		237.85	234.26	189.30	199.58
Mar-12	173.54	99.21	172.48	98.92		241.11	240.65	189.93	222.11
Apr-12	169.93	96.79	171.07	96.89		244.37	236.21	190.57	209.47
May-12	173.34	98.39	173.27	98.48		247.63	242.72	191.21	207.28
Jun-12	173.80	98.30	173.84	98.44		250.88	246.50	191.85	202.23
Jul-12	176.03	99.21	174.94	98.92		254.14	253.06	192.50	197.26
Aug-12	172.35	96.79	173.53	96.89		257.40	248.24	193.15	190.92
Sep-12	175.81	98.39	175.73	98.48		260.66	254.92	193.79	187.28
Oct-12	176.27	98.30	176.31	98.44		263.92	258.74	194.44	205.17
Nov-12	178.51	99.21	177.41	98.92		267.17	265.47	195.08	188.87
Dec-12	174.78	96.79	176.00	96.89		270.43	260.27	195.73	211.08
Jan-13	178.28	98.39	178.20	98.48		273.69	267.13	196.38	218.58
Feb-13	178.73	98.30	178.77	98.44		276.95	270.98	197.02	214.66
Mar-13	181.00	99.21	179.87	98.92		280.21	277.88	197.67	237.90
MAPE	10.64	39.99	10.64	40.00	6.40	7.07	6.47	5.94	3.41
MAD	9.29	33.80	9.28	33.87	6.15	6.65	6.27	5.55	3.17
MSD	135.32	1608.54	135.63	1608.84	66.00	71.65	65.82	53.57	19.46

The abbreviations, SD MTS stands for seasonal decomposition multiplicative model with trend and seasonal only; SDMS stands for seasonal decomposition multiplicative model with seasonal only; SD TS stands for seasonal decomposition additive model with trend and seasonal only; SD AS stands for seasonal decomposition additive model with seasonal only.

6. Conclusion

Various univariate time series forecasting models for forecasting cement productions in India have been applied in this paper. It compares the out-of-sample forecast accuracy of different models using mean absolute percent error, mean absolute deviation, and mean squared error. The forecasting performances of these selected models are evaluated with these forecast accuracy criteria. Monthly data on cement productions in lakh tons since March 1993 to March 2011 are applied to forecast since April 2011. Since historical data on cement productions mostly follows seasonal pattern, the SARIMA model performs better for forecasting than other competing models as it provides lowest MAPE values. Thus as a policy implications SARIMA model may be considered for forecasting monthly cement productions in India.

References

[1] Box, G., and G. Jenkins, 1970. *"Time Series Analysis: Forecasting and Control".* San Francisco: Holden-Day.

[2] Ghosh, S., 2008. "Univariate Time-Series Forecasting of Monthly Peak Demand of Electricity in Northern India". *International Journal of Indian Culture and Business Management,* 1(4); 466 - 474 .

[3] Gil-Alana, L. A.. 2001. "A Fractionally Integrated Exponential Model for UK Unemployment". *Journal of Forecasting,* 20. 329-340.

[4] Iqbal, N., K. Bakhsh, A. Maqbool and A.S. Ahmad., 2005. "Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan". *Journal of Agriculture And Social Sciences,* 1(2); 120-22.

[5] Johnes, G.,1999. "Forecasting Unemployment". *Applied Financial Economics,* 6, 605-607.

[6] Mandal, B. N., 2005. "Forecasting Sugarcane Productions in India with ARIMA Model". Interstat. Www.Interstat .Statjournals.Net

[7] Pei, L., S. H. Chen, H. H. Yang, C.T. Hung, and M. R. Tsai., 2008. "Application of Artificial Neural Network and SARIMA in Portland Cement Supply Chain to Forecast Demand". Natural

Computation 2008. ICNC '08. Fourth International Conference On,18- 20 Oct. 2008, Volume: 3 Pp. 97-101.

[8] Phillips, P.C.B., and P. Perron.. 1988. "Testing for a Unit Root in Time Series Regression". *Biometrica,* 75;335-346.

[9] Proiettti, T.. 2001. "Forecasting the US Unemployment Rate". Working Paper, Dipartimneto Di Scienze Statistiche, Universita Di Udine, Udine, Italy.

[10] Ramani, K. V., and B. H. Dholakia., 1999. "A Decision Support System to Forecast Cement Demands, Information Technology for Development". 8(3).

[11] RNCOS2008. India Cement Industry Forecast 2012,

[12] Rothman, P., 1998. "Forecasting Asymmetric Unemployment Rates". Review of Economics and Statistics, 80, 164-168.

[13] Wankhade , R., S. Mahalle, S. Gajbhiye, and V.M. Bodade, 2010. "Use of the ARIMA Model for Forecasting Pigeon Pea Production in India". International Review of Business and Finance; 2(1); 97-102.