Forecasting Weekly Material Cost for a Daily Newspaper in Sri Lanka

JMTN Jayasinghe[1#] and RAB Abeygunawardana[2]
[1]*Department of Statistics and Computer Science, University of Kelaniya, Kelaniya, Sri Lanka*
[2]*Department of Statistics, University of Colombo, Colombo 3, Sri Lanka* [#]
thilinijayasi@gmail.com

Abstract— *In newspaper printing different types of costs occurred such as labour cost, material cost, and electricity expenditures. Among them material cost of a newspaper become a main key point in newspaper manufacturing cost. Pricing of the newspaper is depending on the material cost associated with the newspaper. Therefore, having a model to forecast will be beneficial for the company in budget planning and pricing. In this study the analysis was carried out by using weekly data of material cost (paper, plate and 4 colour inks separately) from January 2013 to July 2014 for a daily newspaper from leading newspaper company in Sri Lanka. Main objective of this research is to fit a suitable model for forecasting material cost. The data set was divided into two parts; one for model fitting and other for model validation. Univariate time series model was fitted to the data. Different Auto Regressive Integrated Moving Average (ARIMA) models were fitted for those data and best model for forecasting was identified by using minimum Mean Absolute Percentage Error (MAPE) value. ARIMA (2, 1, and 1) model with minimum MAPE 16.13% was identified as the best model for total material cost forecasting. ARIMA(1,1,2), ARIMA(0,1,1), ARIMA(1,1,2), ARIMA(3,1,2) , ARIMA(3,1,2) , ARIMA(1,1,2) for paper, plate, cyan ,magenta ,yellow and black ink costs respectively with minimum MAPE values 23.86 %, 7.9%, 8.33% , 8.5426 % , 8.29% and 8.9%. Models for material costs separately have a minimum MAPE value than taking the total material cost. It can be conclude that it is better to use fitted models separately for material cost for forecasting.*

Keywords*: Auto Regressive Moving Average (ARIMA), Mean Absolute Percentage Error (MAPE)*

I. INTRODUCTION

In newspaper printing different types of costs occurred such as labour cost, material cost and electricity expenditures. Among them material cost of a newspaper become a main key point in newspaper manufacturing process. Deciding price for a newspaper is becomes a difficult task that newspaper companies face. Pricing of the newspaper is depending on the material cost associated with the newspaper. Therefor having a model to forecast material cost will be beneficial for newspaper industry in budget planning and pricing for a newspaper. The main objectives of this research are to identify the patterns in weekly material cost of a daily newspaper and to fit a suitable model to forecast the material cost of a daily newspaper.

A. Literature review

A. It is difficult to find out studies publish on material cost forecasting of newspapers. Therefor researches done on price forecasting in several areas were considered in this case. According to report of digital magazine and newspaper publishing in Canada (2014) printing cost is highly contributed in newspaper production cost. Pereira E (2011) stated that "cost and demand – seems to be an efficient way to coordinate price decisions in concentrated markets and long run profitability". Thus the demand and material cost forecasting is very essential in product management. Lots of researchers suggested that time series models can be used in price forecasting. Green S (2011) suggested a time series analysis for stock price forecasting. Different ARIMA models were fitted for different companies. Different time intervals were taken and ARIMA models were fitted. Among them best models and best time intervals were identified for forecasting.

II. METHODOLOGY AND EXPERIMENTAL DESIGN

The analysis was carried out by using a statistical approach. Univariate Time Series model were fitted here because of the significance of forecasting future values of a time series go beyond a range of disciplines. Characteristics of time series are trend, cycle, seasonal, and random components in business and economic time series. Many strong methods were developed to identify these factors by estimating statistical models. Box–Jenkins Methodology is a developed method to fit ARIMA models for a time series. According to the researchers ARIMA models provide accurate results for out of samples than the other time series models such as exponential smoothing, Auto Regressive and Moving average models.

The data were taken from a popular newspaper company. Daily newspaper was selected and secondary data was used. The data were recorded the material cost individually by the company daily but the company order the materials in weekly basis. As the material cost the company considered paper cost, plate cost and cyan, magenta, yellow and black colour ink costs. Those data were obtained individually. Material cost from 06th January 2013 to 28th August 2014 was taken. For model fitting purpose data from 06th January 2013 to 25th July 2014 were taken and the rest of data were used for model validation .As weekly material cost sum of paper, plate and ink (cyan, magenta, yellow and black) costs were taken. For each case descriptive statistics were obtained. Then time series plot for material cost was obtained and checked whether seasonal variations and trend pattern exists or not. Box–Jenkins Methodology was applied to find out the ARIMA models. More than two ARIMA models were fitted for each cases and best model was identified by the minimum MAPE value of the residuals. In each case normality of the residuals were checked. The same approach was applied for the individual costs occurred and the analysis was carried out.

III. RESULTS

Total material cost for printing the newspaper was taken for analysis. The cost was in million rupees. Descriptive statistics were obtained and the results were shown in table 1.

Table 1: Descriptive statistics of weekly total material cost

Summary Statistics	Value
Mean	1.6664
Standard deviation	0.4763
Median	1.6214
Minimum	0.9068
Maximum	2.9280

Average weekly total material cost of the newspaper is 1.67 million and the dispersion around the mean is 0.48 million.

Time series plot of weekly material cost for the newspaper was obtained and figure 1 shows the result.

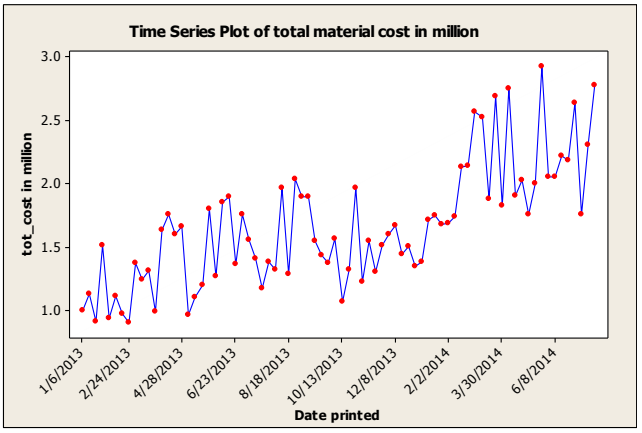

Figure 1: Time Series plot of weekly material cost

Time series contain an increasing trend. No seasonal pattern present in the series. This imply that time series is non-stationary.

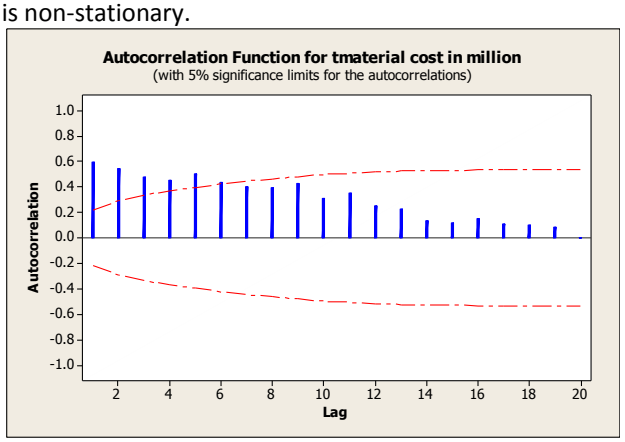

Figure 2: ACF of weekly material cost

ACF decays slowly. The series has a trend pattern. This implies that time series is non-stationary. Significant lags can be obtained up to lag 6.

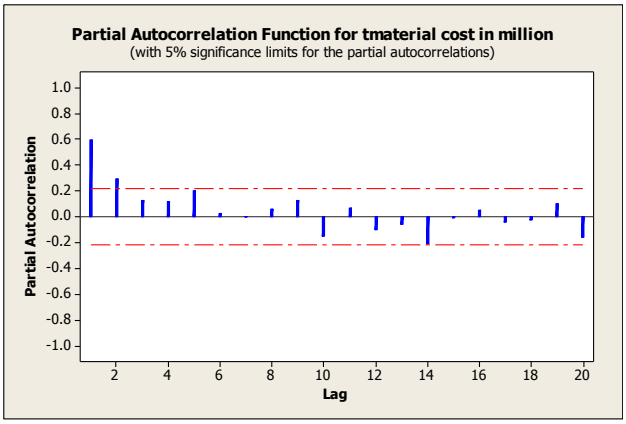

Figure 3: PACF of weekly material cost

First difference of the time series was taken to make the series stationary.

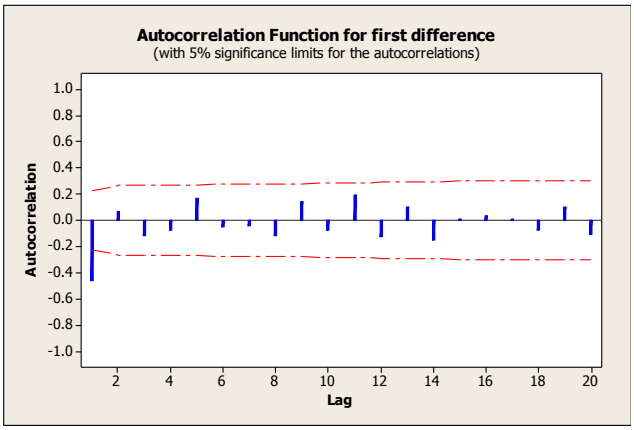

Figure 4: ACF of first difference weekly material cost

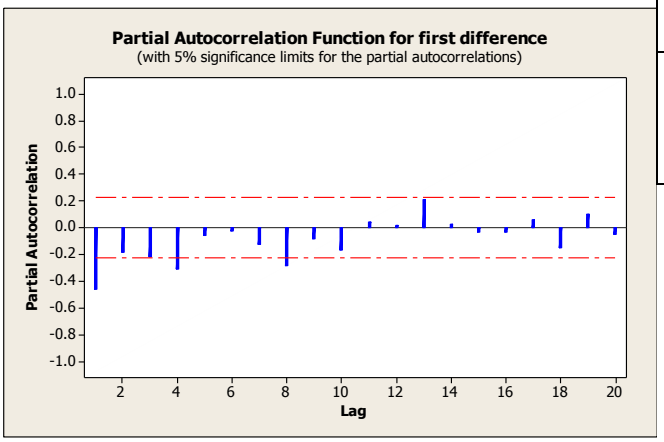

Figure 5: PACF of first difference weekly material cost

ACF decay quickly .No trend pattern in the first difference. ACF cut off at lag 1. Thus the series is stationary.

PACF of the first difference of the weekly material cost shown in figure 5. PACF cut off at lag 1.

According to cut off lags ARIMA (1, 1, 1) model was identified as the first step. AR (1) model was insignificant. There for ARIMA(2,1,1), ARIMA(1,1,2) ,ARIMA(1,1,0) and ARIMA(0,1,1) models were fitted. Summarised results were mention in table 2 for the other models.

Table 2: Comparison of ARIMA models of weekly total material cost

Model	**Significance of coefficients**	**Box-Pierce (Ljung-Box)**
ARIMA(1,1,1)	AR(1) term is insignificant Model not good	Lags 12,24,36,48 are insignificant No Auto correlation
ARIMA(2,1,1)	All AR and MA terms are significant. Constant term is insignificant Model is good	Lags 12,24,36,48 are insignificant No Auto correlation
ARIMA(1,1,2)	All terms are significant Model is good	Lags 12,24,36,48 are insignificant No Auto correlation
ARIMA(0,1,1)	All terms are significant Model is good	Lags 12,24,36,48 are insignificant No Auto correlation
ARIMA(1,1,0)	AR term and constant terms are significant Model is good	Lags 12,24,36,48 are insignificant No Auto correlation

According to the table 2 all coefficients are significant in those models except AR (1) in ARIMA (1, 1, 1) and no auto correlations between residuals. All the terms in other models were significant and therefor lowest values of Mean Absolute percentage error (MAPE), Mean Absolute Deviation (MAD) and Mean standard deviation (MSD) values were used to identify the best ARIMA model.

Table 3: Measurement of accuracy of fitted models of weekly material cost

Model	**MAPE %**	**MAD**	**MSD**
ARIMA(2,1,1)	16.136	0.264	0.104
ARIMA(1,1,2)	16.798	0.274	0.114
ARIMA(1,1,0)	16.889	0.273	0.105
ARIMA(0,1,1)	16.903	0.276	0.121

ARIMA (2, 1, 1) model has the minimum MAPE value. Hence Box-Jenkins model to forecast weekly plate cost of the newspaper was chosen as ARIMA (2, 1, and

1).Results obtained for ARIMA (2, 1, and 1) models were mention below.

Type	Coef	SE Coef	T	P
AR 1	-1.5313	0.1188	-12.88	0.000
AR 2	-0.6287	0.0980	-6.42	0.000
MA 1	-0.9236	0.0947	-9.75	0.000
Constant	0.06272	0.07591	0.83	0.411

The full model:

$X_t = 0.06272 - 1.5313\ X_{t-1} - 0.6287\ X_{t-2} - 0.9236\ \varepsilon_{t-1} + \varepsilon_t$

Where $X_t = Y_t - Y_{t-1}$

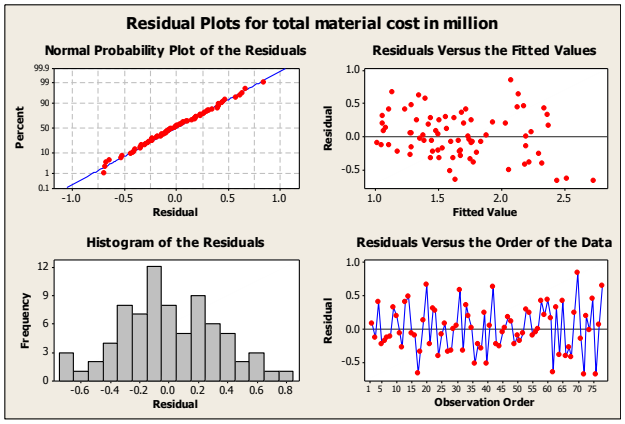

Figure 6: Residual plots of fitted model of weekly material cost

Histogram of residuals has a bell shaped appearance and normal probability plot of the residuals has approximately straight line. They suggested that residuals are normally distributed.

Results for individual material cost (Paper, plate and four colour inks) mentioned below. Models were fitted for each material cost.

Table 4: Descriptive statistics of weekly material cost individually

	Mean	Standard deviation
Plate cost(Rs)	1.1299	0.2708
Paper cost(Rs)	13.595	5.486
Cyan ink cost(Rs)	7.109	1.375
Magenta ink cost(Rs)	6.918	1.338
Yellow ink cost(Rs)	6.931	1.437
Black ink cost(Rs)	5.919	1.337

Summary of descriptive statistics for individual material costs were mentioned in table 4.

Summary for model identification for material cost individually mentioned in table 5.

Table 5: Properties of time series weekly material cost individually

	Time series properties	Stationary of the series
Plate cost(Rs)	Trend exist	Non stationary
Paper cost(Rs)	Trend exist	Non stationary
Cyan ink cost(Rs)	Trend exist	Non stationary
Magenta ink cost(Rs)	Trend exist	Non stationary
Yellow ink cost(Rs)	Trend exist	Non stationary
Black ink cost(Rs)	Trend exist	Non stationary

Since all the time series are non-stationary at the level first difference of the data were obtained and results mention in the table 6.

Table 6: Properties of first differenced time series weekly material cost individually

	Stationary of the first difference	Cut off lags of ACF and PACF
Plate cost(Rs)	Stationary	ACF-1 PACF-2
Paper cost(Rs)	Stationary	ACF-1 PACF-1
Cyan ink cost(Rs)	Stationary	ACF- 1 PACF-1
Magenta ink cost(Rs)	Stationary	ACF- 1 PACF-1
Yellow ink cost(Rs)	Stationary	ACF- 1 PACF-1
Black ink cost(Rs)	Stationary	ACF-1 PACF-1

All the five time series are stationary at the first difference. Their cuts off points are different for each other.

ARIMA (1,1,2) model for plate cost, ARIMA (1,1,1) model for paper cost ,cyan, magenta, yellow and black ink costs were identified as the basic models.

Table 7: Summary of fitted models for weekly material cost individually

	Model	MAPE value %	Normality of residuals
Paper cost	ARIMA(1,1,2)	23.886	Approximately normally distributed
Plate cost	ARIMA(0,1,1)	7.9043	Normally distributed
Cyan ink cost	ARIMA(1,1,2)	8.33	Normally distributed
Magenta ink cost	ARIMA(3,1,2)	8.5426	Normally distributed
Yellow ink cost	ARIMA(3,1,2)	8.2889	Normally distributed
Black ink cost	ARIMA(1,1,2)	8.905	Normally distributed

Parameter estimation and diagnostic checking for each material cost individually followed the same procedure and following results were obtained. As the identified models were not adequate extra models were fitted. Best fitted models and adequacy of each model is summarized in table 7.

Table 8: ARIMA Models for each material cost

Cost	Model
Paper	$X_t = 0.6461\ X_{t-1} + 1.5029\ \varepsilon_{t-1} - 0.5595\ \varepsilon_{t-2} + \varepsilon_t$
Plate	$X_t = 0.007229 + 0.9270\ \varepsilon_{t-1} + \varepsilon_t$
Cyan ink	$X_t = 0.023188 + 0.7636\ X_{t-1} + 1.3330\varepsilon_{t-1} - 0.3513\ \varepsilon_{t-2} + \varepsilon_t$
Magenta ink	$X_t = -0.7774\ X_{t-1} - 1.1182\ X_{t-2} - 0.3407\ X_{t-3} - 0.3229\varepsilon_{t-1} - 0.9644\ \varepsilon_{t-2} + \varepsilon_t$
Yellow ink	$X_t = -0.8722\ X_{t-1} - 1.1798\ X_{t-2} - 0.3607\ X_{t-3} - 0.4868\varepsilon_{t-1} - 0.9991\ \varepsilon_{t-2} + \varepsilon_t$
Black ink	$X_t = 0.016776 + 0.6547\ X_{t-1} + 1.1753\varepsilon_{t-1} - 0.1972\ \varepsilon_{t-2} + \varepsilon_t$

Residuals of all the models were normally distributed. MAPE values of fitted models were less than 10% for each material except paper cost. Models for paper cost, plate cost, and four colours inks are mentioned in table 8. Here $X_t = Y_t - Y_{t-1}$, where Y_t is cost at week t.

IV) DISCUSSION AND CONCLUSION

Material costs were forecasted by using ARIMA model. Here more than one ARIMA model was fit for each case and best model was identified. Here MAPE value was used only for identify the best model. Here in some cases fitted models show some high MAPE value because the data which was used from a real time situation and the assumptions which we made in model fitting may not be satisfy. ARIMA (2, 1, 1) model was proposed for total material cost forecasting of the newspaper with minimum MAPE value from observed models. Residuals error percentages were not very high values. All of them are less than 20%. Actual and fitted values were not varying with big amount. Therefor the model is efficient in weekly material cost forecasting of the newspaper. Each material cost were also modelled individually .Those models MAPE values were less than 10% except paper cost. ARIMA (1,1,2) for paper cost with MAPE 23.886%, ARIMA(0,1,1) models for plate cost with MAPE 7.9043% , ARIMA (1,1,2) model for cyan ink cost with MAPE value 8.33% ,ARIMA (3,1,2) magenta and yellow ink cost with minimum MAPE value 8.54% and,8.29% respectively and ARIMA (1,1,2) for black ink cost with minimum MAPE value 8.905% are the fitted models for each material cost. Residuals of all the models were normally distributed. Therefor those models can be used to forecast material cost individually except the paper cost.

By comparing to the obtained results it can be conclude that models which were fitted individually for material cost forecasting is effective than the model for total material cost forecasting. These models are effective for any daily newspaper in Sri Lanka for forecasting the material cost.

ACKNOWLEDGMENT

I would like to express my heartiest gratitude for Wijaya Newspapers LTD for providing me the data. I also like to express thanks to the Department of Statistics and computer science, University of Kelaniya, Sri Lanka for giving me the facility to carry out this project

REFERENCES

Green, S., 2011. *Time Series Analysis of Stock Prices Using the Box-Jenkins Approach,* STATESBORO, GEORGIA: Georgia Southern University.

Nordcity, 2014. *Digital Magazine and newspaper publishing in Canada,* s.l.: Price Water House Coopers.

Pereire, 2011. *Using time series analysis to understand price setting,* s.l.: University of Campinas.

BIOGRAPHY OF AUTHORS

J.M.T.N.Jayasinghehas completed her B.Sc. (special) Degree in Statistics with a first class honors. She is currently working as a temporary demonstrator at Department of Statistics and Computer Science, University of Kelaniya. Her interested areas are Time Series Analysis, Operational Research and Regression models.

Mr.Rushan A.B.Abeygunawardana is a senior Lecturer (Grade II) at Department of Statistics, University of Colombo. His research interested areas are Small area estimation, Time Series Analysis and Quality Controlling.