دراسة مقارنة للحصول على أفضل ترتيب للنماذج التنبوئية الرياضية لتحليل السلاسل الزمنية مع تطبيق عملي

سهى طالب عبد الرحمن
قسم الرياضيات ، كلية التربية –ابن الهيثم، جامعة بغداد

استلم البحث في 6 حزيران 2010
قبل البحث في 27 أيلول 2010

الخلاصة

يهدف البحث إلى إجراء التنبؤ المبني على النماذج الموفقة لبيانات السلسلة الزمنية المبحوثة. إن اختيار الأنموذج يعتمد على كيفية استخدامه لأغراض التنبؤ، درجة الدقة المتوخاة من عملية التنبؤ، والوقت والإمكانيات المتاحة للمستخدم، والبيانات المتوافرة ذات القدر المعنوي تطبيقياً ، إلى أي مدى ينبغي أجراء عملية التنبؤ. تم اقتراح العديد من النماذج وهي (الانتقال العشوائي، والاتجاه الخطي ، وخط الاتجاه ألتربيعي، وخط الاتجاه الآسي ، وشكل الحرف - S ، والوسط المتحرك ، والتقنية الآسية الخطية لبراون، والتقنية الآسية الخطية لهوت، ونماذج بوكس –جنكنز المتمثلة بالانحدار الذاتي من الرتبة الأولى والوسط المتحرك من الرتبة الأولى والأنموذج المختلط من الرتبة الأولى أيضاً . الجانب التطبيقي أختص بالبحث عن الأنموذج الرياضي الأمثل الذي يقتضي استخدامه للتنبؤ بالكميات المنتجة ومبالغ المبيعات السنوية للسكاير العراقية اعتماد على سلسلة زمنية سابقة خلال المدة (2000– 2009) سنة. تم اعتماد مؤشرين، الأول حساب إحصاءه متوسط مربعات الخطأ والاخر إحصاءه متوسط الأخطاء. بينت نتائج تحليل السلسلتين بموجب مؤشر أهمية الأخطاء المتمثل بمتوسط مربعات البواقي بأن الانموذج الأمثل لبيانات الكميات المنتجة من السكاير والمبالغ المباعة بموجبها تلك الكميات هو أنموذج خط الاتجاه ألتربيعي و الاتجاه الخطي على التوالي .

وبموجب مؤشر قيم التحيز لمتوسط البواقي، ترشحا نموذج شكل الحرف – S عن بيانات السلسلتين بوصفه الانموذج الأمثل.

Comparative Study for Obtaining the Best Ordered of the Mathematical Forecast's Models in Time Series Analysis With Application

S. T. Abdul Rahman
Department of Mathematics, College of Education Ibn-Al-Haitham, University of Baghdad

Received in June,6, 2010
Accepted in Sept,27,2010

Abstract

The research aimed to forecasting based on fitting models to past observations in a given time series. The model that should choose depends on how would the forecast used, the degree of accuracy which had been required from the forecast, the amount of time and capital available to the user, the amount and type of data available to be meaningful application, and how far ahead that must be forecast. Several forecasting models suggested, such as (Random Walk, Linear trend, Quadratic trend, Exponential trend, S – Curve, Moving Average, Brown's Linear Exponential Smoothing, Holt's Linear Exponential Smoothing, ARMA (1,0), ARMA(0,1) and ARMA (1,1). Application section deals with searching for an optimal of mathematical model that ought to be used for prediction or forecasting the quantities and account of sells the Iraqi's cigarettes depends along the period in the past observations (2000 – 2009) yrs. Two criteria had been applied, the first statistic measured the magnitude of the errors (MSE) and the second statistic measure bias (ME). The two series (Quantity & Account) according to the magnitude of the residuals through applying (MSE) indicator showed that the best mathematical models were with Quadratic Trend and Linear trend respectively.
The two series (Quantity & Account) according to the Bias through applying (ME) indicator showed that the best mathematical model was with S-Curve.

Introduction

The time-series decomposition approach to forecasting is based on the principle of "breaking down" a time series into each of its components (¾) and then forecasting by predicting each component separately, except randomness, which it cannot predict (Makridakis et al.,1983) [1].
Classical decomposition breaks the time-series data into four components: trend, seasonality, cycle, and irregular movement.
Trend: Is the long-term behavior of the data.
Seasonality: Is the periodic fluctuation of constant lengths that usually repeats itself at fixed intervals.
Cycle: Is the wave-like patterns of ups and downs that are similar to seasonality. Cycles can take place over a long time period such as several years. The length of time between peaks is not fixed as it is with seasonal movement.
Irregular Movement: Is movement in the data that we cannot attribute to trend, seasonality, or cycle. This component is often known as noise or white noise and some time called residuals.
The Seasonal Decomposition Analysis allows you to perform a classical decomposition of time-series data using a multiplicative or additive model.
Multiplicative: The program uses the multiplicative ratio-to-moving-average method to make the seasonal adjustment. The program assumes that the components of the data (trend,

seasonal, and random effects) cause proportional changes in the series, and therefore expresses the result as a product of the components.
Additive: The program uses the additive difference from the moving-average method to make the seasonal adjustment. It simply adds or subtracts given amounts and expresses the result as a sum of the components.
The analysis provides several methods that can be used to adjust or transform the data to make it more appropriate for the analysis performing [2].

Objective

Selecting an optimal model in fitting the studied data through obtaining the best ordered of mathematical forecast's models in time series analysis.

Using the Forecasting Analysis

There is frequently a delay between the time you become aware of an upcoming event or need and when that event or need actually occurs. This delay is the primary reason for planning and forecasting.
Forecasting is based on fitting a model to past observations in a given time series. The model you choose depends on how you will use the forecast, the degree of accuracy you require from the forecast, the amount of time and capital available to you, the amount and type of data available to you, and how far ahead you must forecast.
Including several forecasting models, we can modify by changing the parameters. With each model we can transform the data or apply a constant rate of inflation. Depending on the model we choose, the estimation also allows us to optimize the parameters or include a constant term. If the data are seasonal, we can apply differencing and choose a seasonal adjustment. Your choise depends on how will you use the forecast, the degree of accuracy you require from the forecasting modules that we studied as follows [3]:
Random Walk: This model randomly forecasts the next observation based on the current observation and ghe mean and standard deviation for the difference of the values.
Linear Trend: Fits a straight line through the data and into the forecast periods. The linear trend on the function: $Z(t) = a + bt$, where (t) represents the time index. It estimates the coefficients by least squares. This method is useful when the actual values follow a straight line.
Quadratic Trend: This model fits a quadratic curve through the data and into the forecasting periods. Fits a quadratic curve (second-order polynomial) through the data and into the forecast periods. The calculation procedure bases the quadratic curve on the function: $Z(t) = a + bt + ct^2$, where (t) represents the time index.
The estimates of the coefficients obtained by least squares. This method is useful when the actual values follow a quadratic curve.
Exponential Trend: Fits an exponential curve through the data and into the forecast periods. The first steps takes the natural logarithm of the Z(t) function and then uses least squares estimates. The exponential curve is based on the function: $Z(t) = \exp(a + bt)$.
This method is useful for data that are increasing or decreasing at a non constant rate.
S-Curve: Fits an S-shaped curve through the data and into the forecast periods. The first steps take the natural logarithm of the Z(t) function and the reciprocal of (t), and then use least squares estimates. The exponential curve is based on the function: $Z(t) = \exp(a + b/t)$. This method is useful only if the actual values follow the S-shaped curve [4].
Moving Average: This model uses the moving average to smooth the data and to predict future values. The calculation procedure computes a new average for each period by adding the next observation to the calculation and dropping the oldest observation. This ensures that the number of points in each average is always the same.
Brown's Linear Exponential Smoothing: This model smoothes the data and predicts future values by applying a double-smoothing formula to the data using one parameter, alpha. In

other words, the calculation procedure smoothes the data twice uses the alpha parameter. This type of smoothing is useful for non seasonal data.

Holt's Linear Exponential Smoothing: This model applies a double-smoothing formula to the data using two smoothing parameters:alpha and beta, alpha to smooth the estimate of the level, and beta to smooth the estimate of the trend. This type of smoothing is useful for non seasonal data, but provides more flexibility than Brown's Linear Exponential Smoothing.

ARMA Model of low ordered: This model estimates and forecasts using the methods prescribed by Box and Jenkins (1976) [5]. The AR MA (Auto Regressive Moving Average) analysis allows to model a discrete time series as a function of a constant autoregressive term (AR), and a moving-average (MA) term. In addition to that we had applied the mixed model ARMA. In three models low level of ordered were used. (i.e. AR(1), MA(1) and ARMA(1,1).

Indicators of selecting an optimal Model: The summarizes of the performance accuracy for selecting an optimal model in fitting the studied data it displays by the following criteria:

1- The mean squared error (MSE).
2- The mean error (ME).

Each of the statistics is based on the one-ahead forecast errors, which are the differences between the data value at time t and the forecast of that value made at time t-1. The first statistic measures the magnitude of the errors. A better model will give a smaller value. The second statistic measures bias. A better model will give a value close to (0.0).

Application Section: For studying the techniques of applying the different time series models that had been mentioned in the theoretical section and how can we choose an optimal one for making the forecast, table (1) includes the two series of quantities of product and the accounts of sells (in Iraqi dinar) of cigarettes along the period (2000 – 2009) yrs.

Source: Repot layout from the marketing office in the state company for tobacco and cigarettes (SCTC) dated in 4^{th} of May 2010 [6].

Summary Explanation About (SCTC):

The state company for tobacco cigarettes (SCTC) is one of the most important companies in Iraq. The cigarettes industry is considered one of the traditional industries in Iraq which was started at the beginning of the last century and the production was manually at the beginning in small private factories, those factories merge together in 1963 according to the nationalization resolution and named as "The Iraqi Tobacco State Enterprise". Since then the company plan to develop the quality of the products by inserting automated production lines and replacing the old lines. The company expanded to produce diverse products so a new factory for producing safety match constructed in 1979.

The company owns two production sites in Baghdad, the first is Baghdad factory in al – Karada – Nazmiya specialized in producing cigarettes and match, also there is a printing section used for printing the packing requirements, the second site is al-Nasir factory in Orfaly, the factory erected and put in trial operation to produce cigarettes in 1988.

The name of the company changed to (SCTC) in 1997.

Results of the Forecasting Models: [7]

The following results represents the forecasts Number of generated (12) yrs. along the periods (2010 – 2021) yrs. For the two series (Quantity & Account) of Iraqi cigarettes that should be produced by the (SCTC) through using the different studied forecasts mathematical models.

Forecast Summary:

Forecast model selected: (Random walk)

Table (2) and figure (1) illustrate that.

Forecast Summary: Linear trend

Forecast model (Quantity) = 582426.0 + -72980.4 t

Forecast model (Account) = 3.75409E7 + -4.59257E6 t

Tables (3), (4) and figure (2) illustrate that.

Forecast Summary: Quadratic trend

Forecast model: Quadratic trend = 867918.0 + -215726.0 t + 12976.9 t^2

Forecast model: Quadratic trend = 4.99669E7 + -1.08056E7 t + 564819.0 t^2
Tables (5), (6) and figure (3) illustrate that.
Forecast Summary: Exponential trend
Forecast model Exponential trend (Quantity) = exp(14.5277 + -0.851924 t).
Forecast model: Exponential trend (Account) = exp(19.2136 + -0.965363 t).
Tables (7), (8) and figure (4) illustrate that.
Forecast Summary: S-Curve
Forecast model (Quantity): S-curve trend = exp(7.71533 + 7.26136 /t).
Forecast model (Account): S-curve trend = exp(11.6546 + 7.68001 /t).
Tables (9), (10) and figure (5) illustrate that.
Forecast Summary: Moving Average
Forecast model (Quantity - Account): Simple moving average of 5 terms.
Table (11) and figure (6) illustrate that.
Forecast Summary: Brown's linear exp .
Forecast model (Quantity): Brown's linear exp. smoothing with alpha = 0.1.
Forecast model (Account): Brown's linear exp. smoothing with alpha = 0.991.
Table (12) and figure (7) illustrate that.
Forecast Summary: Holt's linear exp
Forecast model (Quantity): Holt's linear exp. smoothing with alpha = 0.991 and beta = 0.0279.
Forecast model (Account): Holt's linear exp. smoothing with alpha = 0.991 and beta = 0.0279.
Table (13) and figure (8) illustrate that.
Forecast Summary: ARMA (1, 0): with constant : Model (Quantity & Account)
Estimated white noise standard deviation = 199697.0 : Number of iterations: 4
Estimated white noise standard deviation = 9.0703E6 : Number of iterations: 4
Tables (14), (15) and figure (9) illustrate that.
Forecast Summary: ARMA (0, 1): with constant : Model (Quantity & Account) :
Estimated white noise standard deviation = 213419.0
Estimated white noise standard deviation = 9.88565E6
Tables (16), (17) and figure (10) illustrate that.
Forecast Summary: ARMA (1 , 1) : with constant : Model (Quantity & Account) :
Estimated white noise standard deviation = 203247.0
Estimated white noise standard deviation = 5.53863E6
Tables (18), (19) and figure (11) illustrate that.
Assessment of the Forecast's Models:

According to the studied indicators that assess the magnitude of the residuals (MSE) as well as the Bias criteria through (ME), table (20) showed the summary of the priority's significant of the studied models distributed by the two series (Quantity & Account).

The priority's significant [8] of the studied models distributed by the two series (Quantity & Account) according to the magnitude of the residuals and Bias criteria through applying MSE and ME Indicators respectively showed that the three best ordered mathematical forecasting models with the product of the Quantities series in the magnitude of the residuals by the indicator (MSE) were with Quadratic Trend, then followed by Linear Trend model and finally with ARMA Model of (1, 0). In addition to that the three best ordered mathematical forecasting models with the Account series showed that the Linear trend represents the first model, then followed by Exponential trend and finally followed by ARMA (1,1) .

Now, the priority's significant of the studied models distributed by the two series (Quantity & Account) according to the Bias criteria applying (ME) Indicators showed that the three best ordered mathematical forecasting models were with the product of the Quantities series are S - Curve, then followed by Moving Average model and finally with Brown's Linear Exponential Smoothing. In addition to that the three best ordered mathematical forecasting models were with the Account series showed that the S - Curve represents the first model , then followed by Moving Average and finally followed by Random Walk.

Conclusions

1- The two series (Quantity & Account) according to the magnitude of the residuals through applying (MSE) indicator showed that the best mathematical models were with Quadratic Trend and Linear trend respectively.
2- The two series (Quantity & Account) according to the Bias through applying (ME) indicator showed that the best mathematical model was with S - Curve.

References

1. Makridakis, S. and Wheelwright, S.C. (1983), Iterative Forecasting, Holden-Day, San Francisco, California.
2. Kendall,M. and Ord, J.K. (1990) Time Series, Oxford University Press, New York, Liu, H. and Wu, T. 2005, Sample Size Calculation and Power Analysis of Time-Averaged Difference, Journal of Modern Applied Statistical Methods, 4 :(2)434-445.
3. Liu, H. and Wu, T. (2005), Sample Size Calculation and Power Analysis of Time-Averaged Difference, Journal of Modern Applied Statistical Methods, 4 :(2) 434-445.
4. Voit, E.O. and Schwacke, L. (1998), Scalability properties of the S–distribution, Biometrical Journal, 40: 665-684.
5. Box, G.E.P. and Jenkins, G.M. (1976), Time Series Analysis - Forecasting and Control. Holden-Day.: San Francisco, California.
6. Repot layout from the marketing office in the state company for tobacco and cigarettes (SCTC) dated in 4th of May (2010).
7. STATGRAPHICS Package, Version 4 for data analyze and obtaining the Forecast's Models.
8. Pandit, S.M. and Wu, S.M. (1983), Time Series and System Analysis with Applications, John Wiley and Sons, New York.

Table (1) Quantities (per carton) of products and the accounts of sells (in Iraqi dinar) of cigarettes along the period (2000 – 2009) yrs

Yrs.	Quantity of product	Account of sells
2000	482904	28538821
2001	668030	37637366
2002	591736	36533018
2003	28846	18615193
2004	23439	846526
2005	5001	267000
2006	5934	209000
2007	2763	110710
2008	1272	46202
2009	412	13740

Table (2) Estimations indicators for quantity & account series through forecasts equation by random walk model

Statistic	Estimation indicators for Quantity	Estimation Period for Account
MSE	3.97023E10	8.01257E13
ME	-53610.2	-3.16945E6

Table (3) Estimations indicators for quantity & account series through forecasts equation by linear trend model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	3.29905E10	5.34098E10
ME	5.23869E-11	29403.8

Table (4) Linear trend model summary estimation period for quantity& account series

Series	Parameter	Estimate	Stnd. Error	t	P-value
Quantity	Constant	582426.0	124079.0	4.694	0.001553
	Slope	-72980.4	19997.1	-3.64955	0.006499
Account	Constant	3.75409E7	6.22638E6	6.02933	0.000313
	Slope	-4.59257E6	1.00347E6	-4.57668	0.001810

Table (5) Estimations indicators for quantity & account series through forecasts equation by quadratic trend model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	2.50013E10	7.08782E13
ME	2.32831E-11	-4.09782E-9

Table (6) Quadratic trend model summary estimation period for quantity & account series

Series	Parameter	Estimate	Stnd.	t	P-value
Quantity	Constant	867918.0	185971.0	4.66696	0.002296
	Slope	-215726.0	77669.2	-2.7775	0.027397
	Quadratic	12976.9	6881.2	1.88585	0.101292
Account	Constant	4.99669E7	9.90193E6	5.04618	0.001486
	Slope	-1.08056E7	4.13546E6	-2.6129	0.034766
	Quadratic	564819.0	366387.0	1.54159	0.167075

Table (7) Estimations indicators for quantity & account series through forecasts equation by exponential trend model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	5.34098E10	9.38565E10
ME	29403.8	-143038.0

Table (8) Exponential trend model summary estimation period for quantity & account series

Series	Parameter	Estimate	Stnd.	t	P-value
Quantity	Constant	14.5277	0.4919	29.534	0.000000
	Slope	-0.851924	0.0792768	-10.7462	0.000005
Account	Constant	19.2136	0.583153	32.9478	0.000000
	Slope	-0.965363	0.0939836	-10.2716	0.000007

Table (9) Estimations indicators for quantity & account series through forecasts equation by S-curve model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	1.00129E12	6.42219E15
ME	-154567.0	-1.36357E7

Table (10) S - curve model summary estimation period for quantity & account series

Series	Parameter	Estimate	Stnd. Error	t	P-value
Quantity	Constant	7.71533	0.878172	8.78568	0.000022
	Slope	7.26136	2.23073	3.25516	0.011611
Account	Constant	11.6546	1.08319	10.7596	0.000005
	Slope	7.68001	2.7515	2.79121	0.023515

Table (11) Estimations indicators for quantity & account series through forecasts equation by moving average model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	4.16481E10	2.13957E14
ME	-151778.0	-1.16334E7

Table (12) Estimations indicators for quantity & account series through forecasts equation by Brown's linear exp. model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	9.38565E10	9.45386E13
ME	-143038.0	904061.0

Table (13) Estimations indicators for quantity & account series through forecasts equation by Holt's linear exp. model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	4.41094E10	8.69601E13
ME	-31394.8	-1.91515E6

Table (14) Estimations indicators for quantity & account series through forecasts equation by ARMA (1, 0) model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	3.87051E10	7.99483E13
ME	-23986.7	-1.24257E6

Table (15) ARMA(1, 0) Model summary estimation period for quantity & account

Series	Parameter	Estimate	Stnd. Error	t	P-value
Quantity	AR(1)	0.817487	0.218792	3.73637	0.005735
	Mean	125560.0	327587.0	0.383287	0.711492
	Constant	22916.3			
Account	AR(1)	0.921902	0.16936	5.44344	0.000613
	Mean	-2.60871E6	4.54078E7	-0.0574507	0.955595
	Constant	-203734.0			

Table (16) Estimations indicators for quantity & account series through forecasts equation by ARMA (0, 1) model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	4.18195E10	9.14996E13
ME	-16482.0	-242968.0

Table (17) ARMA(0, 1) Model summary estimation period for quantity & account

Series	Parameter	Estimate	Stnd. Error	t	P-value
Quantity	MA(1)	-0.780833	0.190579	-4.09716	0.003451
	Mean	185539.0	115066.0	1.61246	0.145526
	Constant	185539.0			
Account	MA(1)	-0.931157	0.139036	-6.69723	0.000153
	Mean	1.17288E7	5.75964E6	2.03637	0.076101
	Constant	1.17288E7			

Table (18) Estimations indicators for quantity & account series through forecasts equation by ARMA (1,1) model

Statistic	Estimation Period for Quantity	Estimation Period for Account
MSE	4.0617E10	2.97683E13
ME	-13639.9	1.37688E6

Table (19) ARMA(1,1) Model summary estimation period for quantity & account

Series	Parameter	Estimate	Stnd. Error	t	P-value
Quantity	AR(1)	0.62306	0.364764	1.70812	0.131369
	MA(1)	-0.42783	0.405937	-1.05393	0.326927
	Mean	159938.0	224627.0	0.712015	0.499492
	Constant	60287.0			
Account	AR(1)	0.531499	0.0624412	8.51198	0.000061
	MA(1)	-1.46393	0.205273	-7.13161	0.000188
	Mean	1.3451E6	1.16361E6	1.15598	0.285616
	Constant	630183.0			

Table (20) Summary of the priority's significant of the studied models distributed by the two series (quantity & account) according to the magnitude of the residuals and Bias criteria by MSE and ME Indicators

Studied Math. Model	Quantity		Account	
	magnitude of the residuals	Bias criteria	magnitude of the residuals	Bias criteria
Random Walk	4	4	6	3
Linear Trend	2	10	1	9
Quadratic Trend	1	9	4	8
Exponential Trend	9	11	2	7
S-Curve	11	1	11	1
Moving Average	6	2	10	2
Brown's Linear Exponential Smoothing	10	3	9	10
Holt's Linear Exponential Smoothing	8	5	7	4
ARMA Model of (1, 0) ordered	3	6	5	5
ARMA Model of (0, 1) ordered	7	7	8	6
ARMA Model of (1, 1) ordered	5	8	3	11

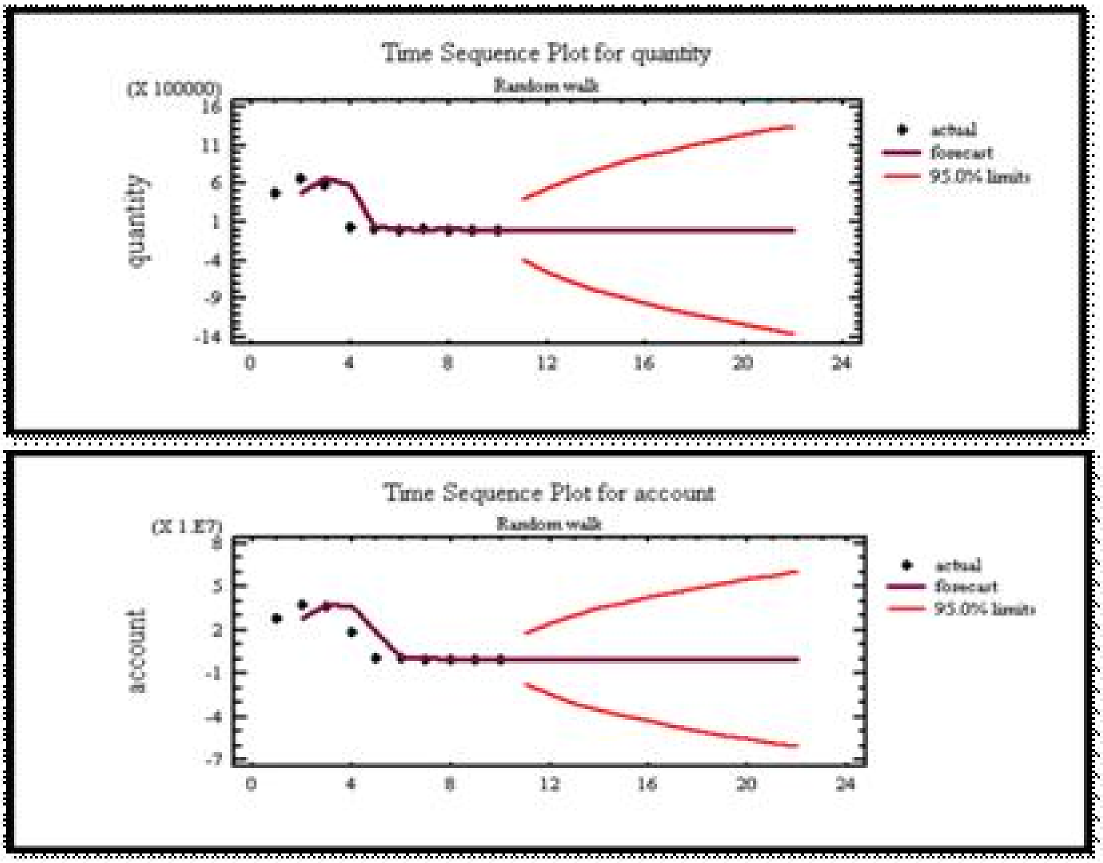

Fig. (1) Time Sequence Plot (quantity & account) with predicted periods by applying random walk model

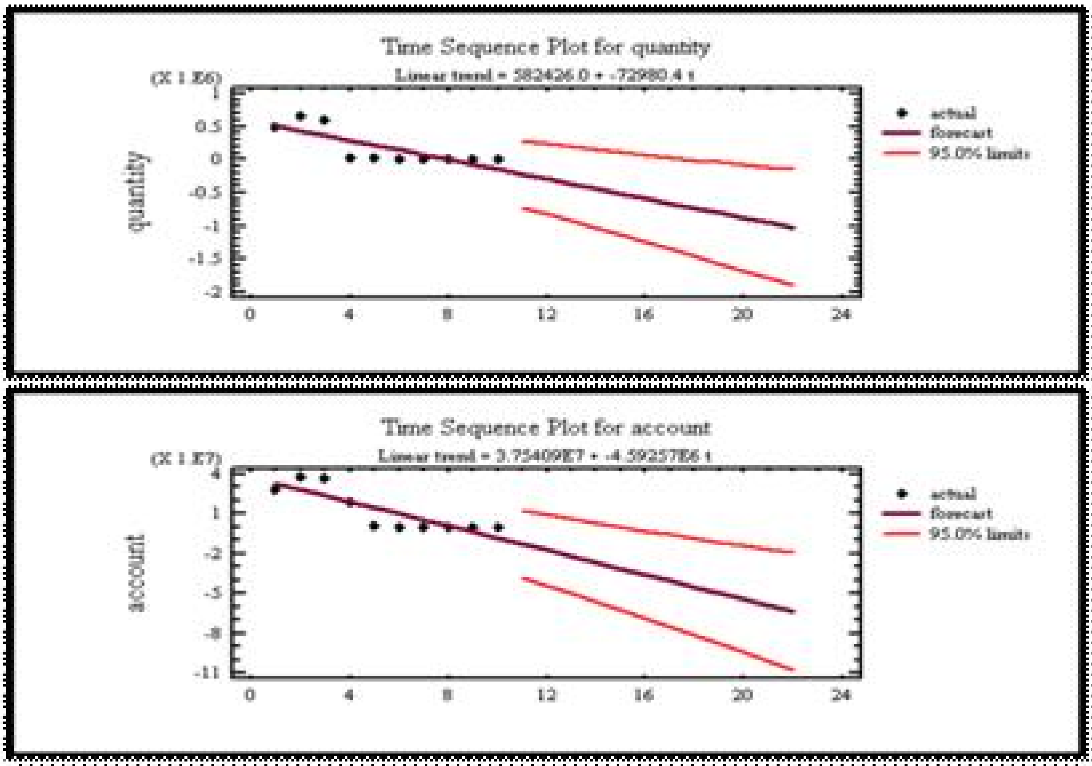

Fig. (2) Time sequence plot (quantity & account) with predicted periods by applying linear trend model

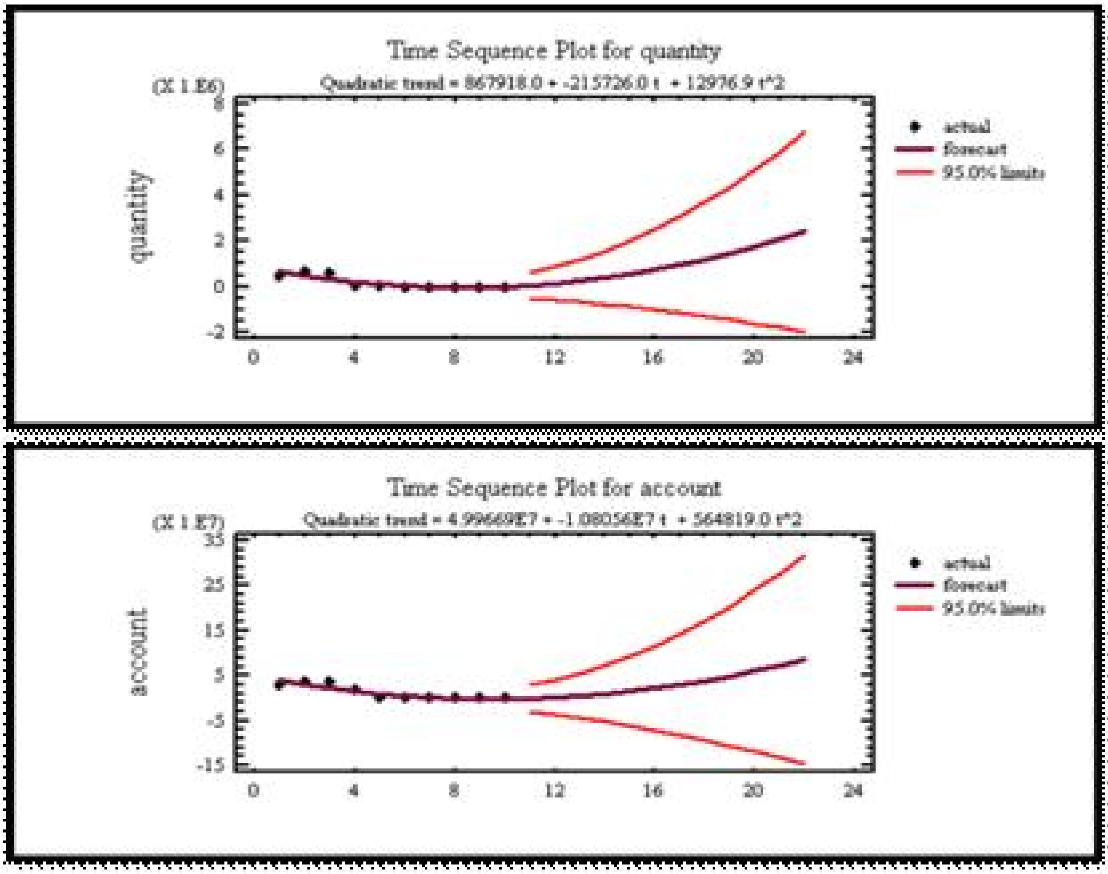

Fig. (3) Time sequence plot (quantity & account) with predicted periods by applying quadratic trend model

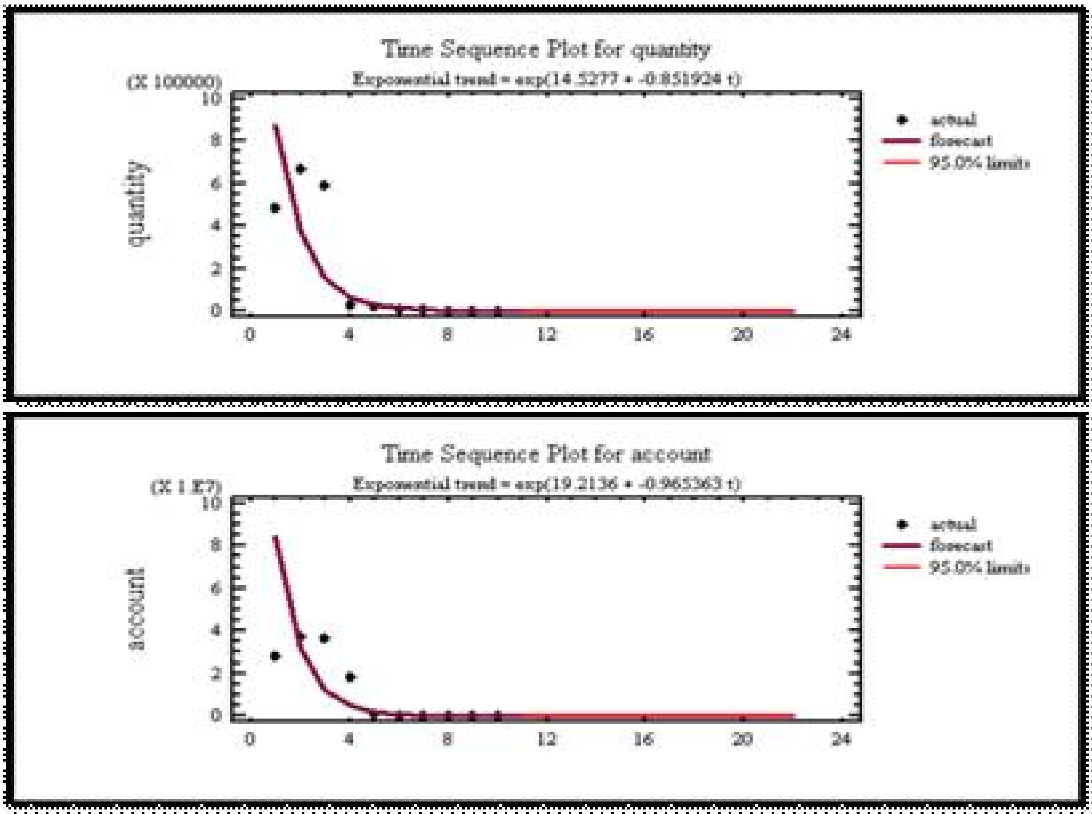

Fig. (4)Time sequence plot (quantity & account) with predicted periods by applying exponential trend model

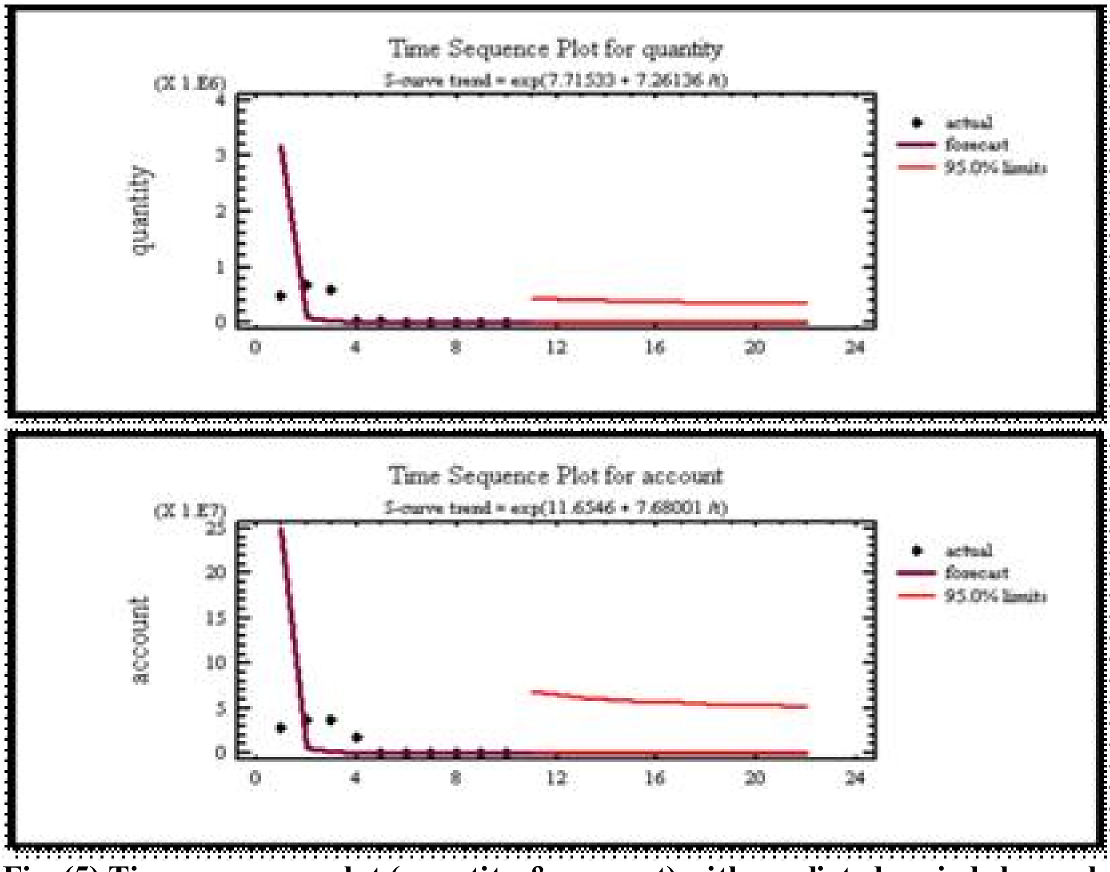

Fig. (5) Time sequence plot (quantity & account) with predicted periods by applying s - curve model

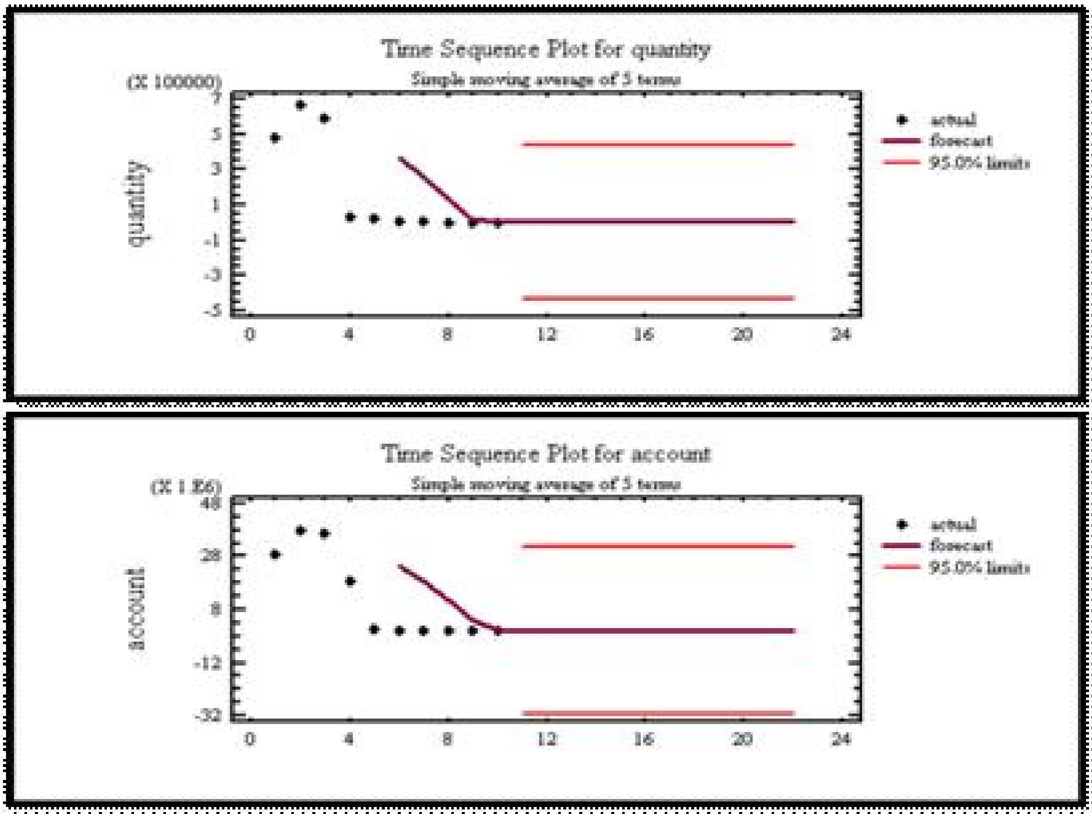

Fig. (6)Time sequence plot (quantity & account) with predicted periods by applying moving average model

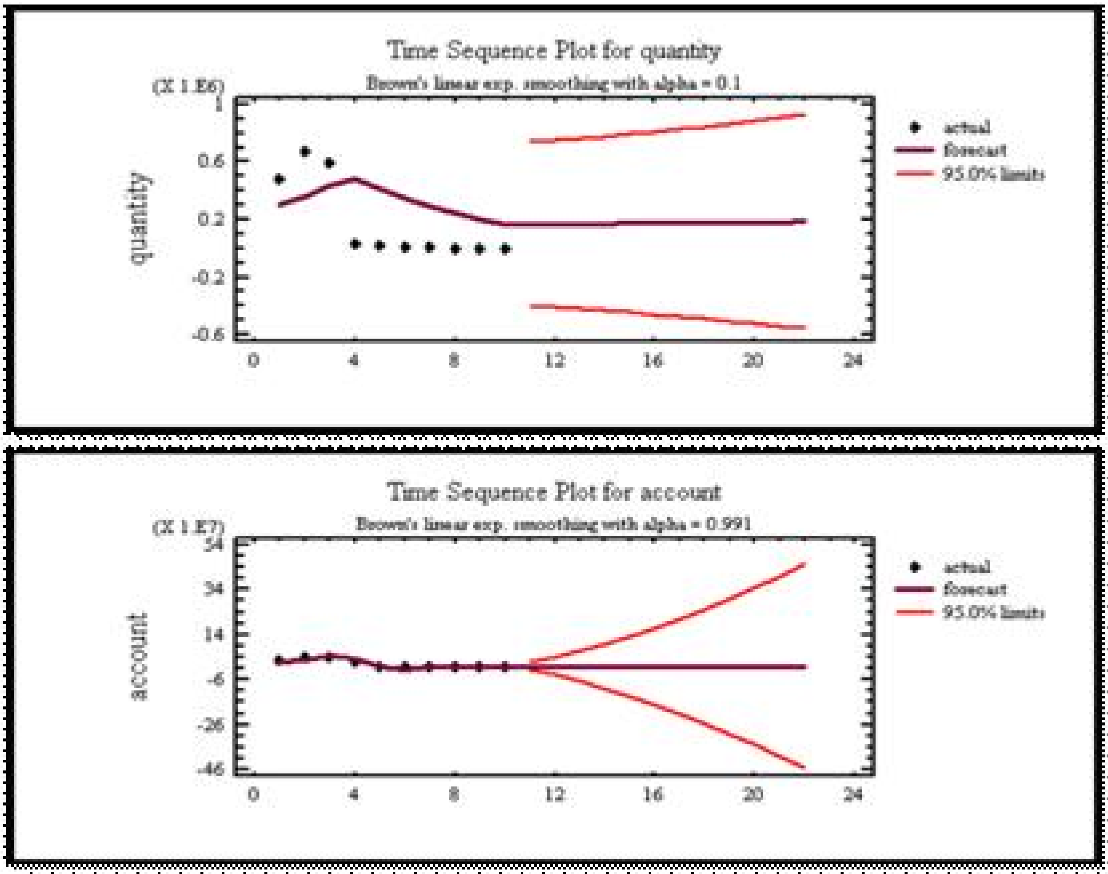

Fig. (7) Time sequence plot (quantity & account) with predicted periods by applying Brown's linear exp. Model

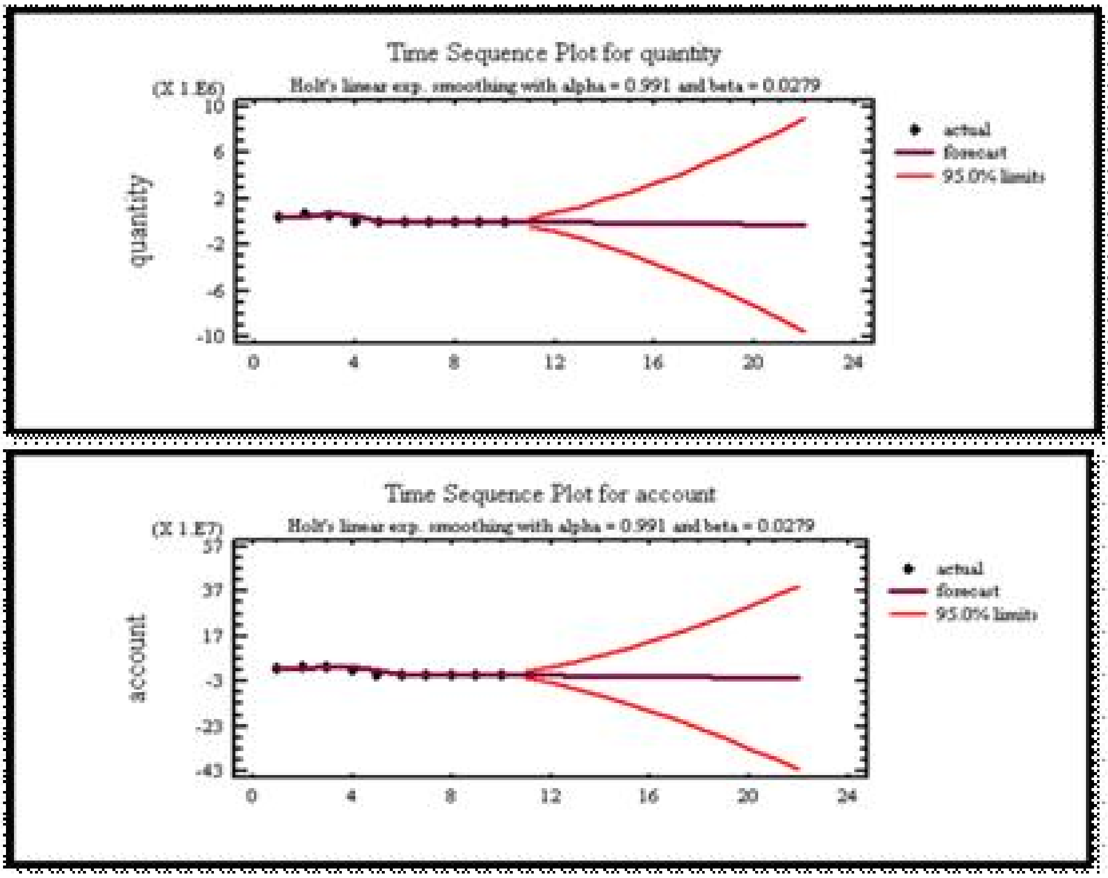

Fig. (8) Time sequence plot (quantity & account) with predicted periods by applying Holt's linear exp. model

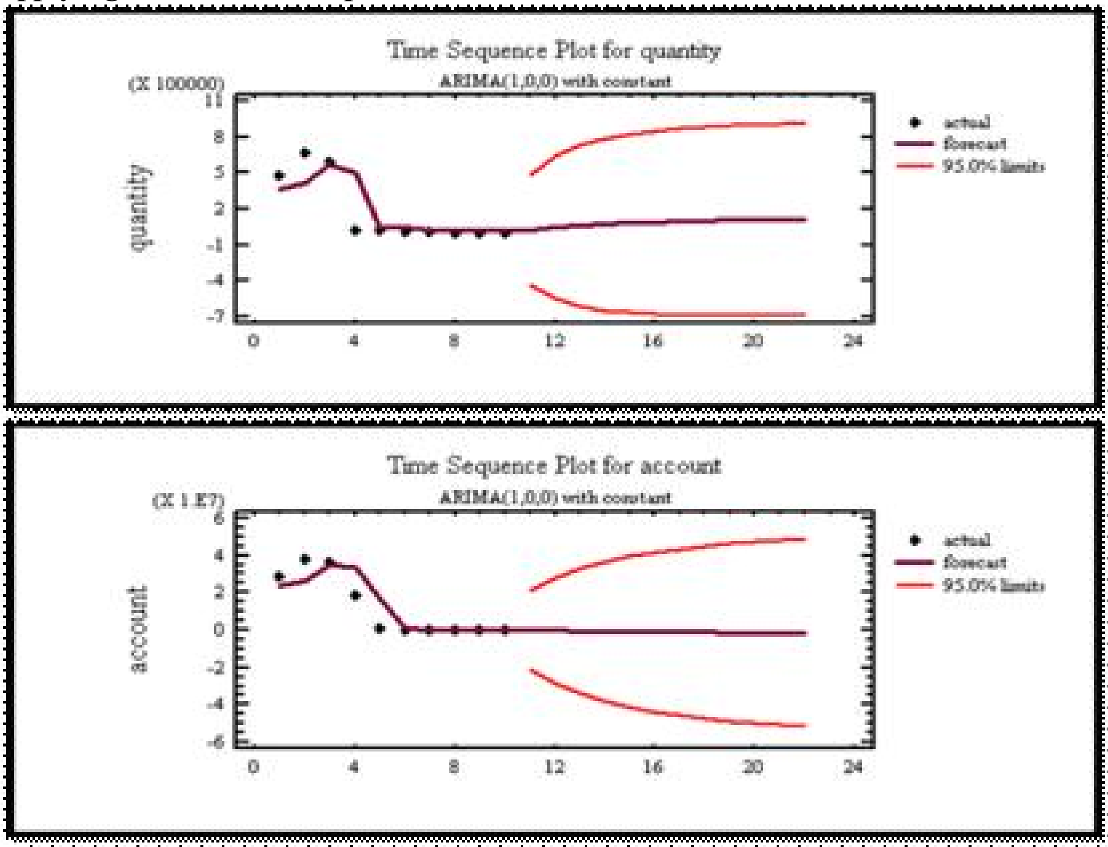

Fig. (9) Time sequence plot (quantity & account) with predicted periods by applying ARMA (1, 0) model

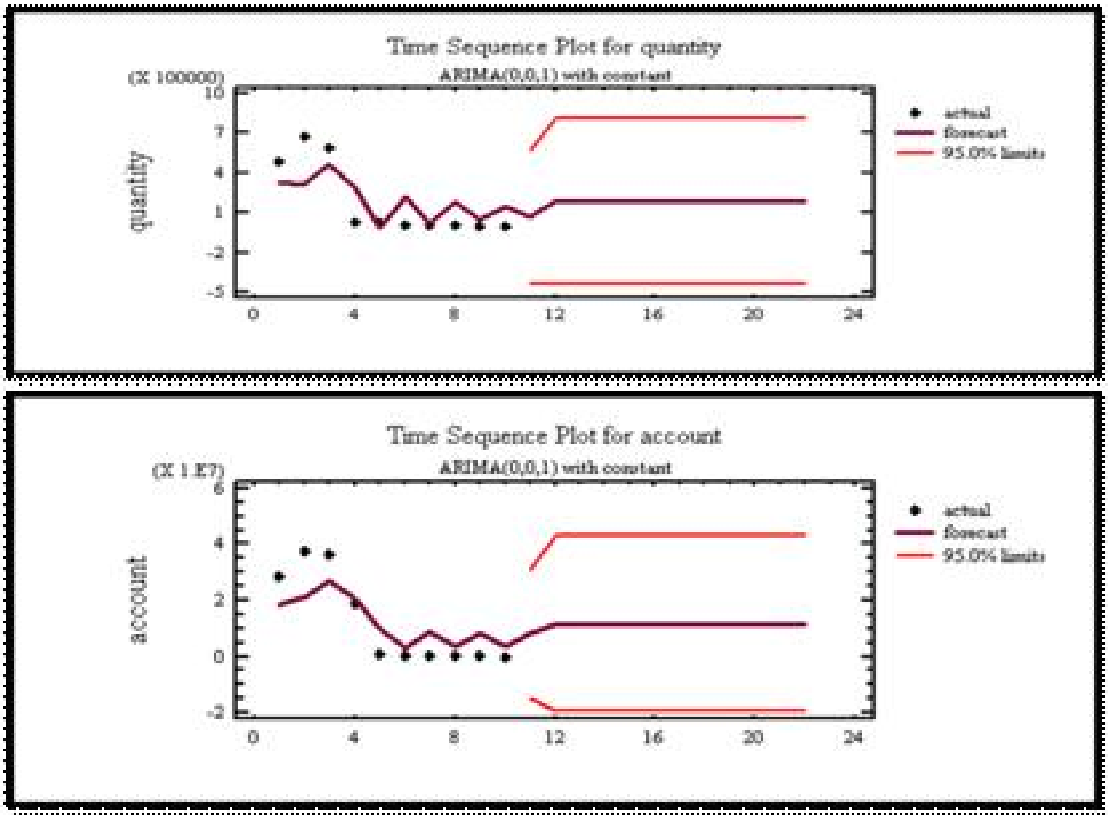

Fig. (10) Time sequence plot (quantity & account) with predicted periods by applying ARMA (0,1) model

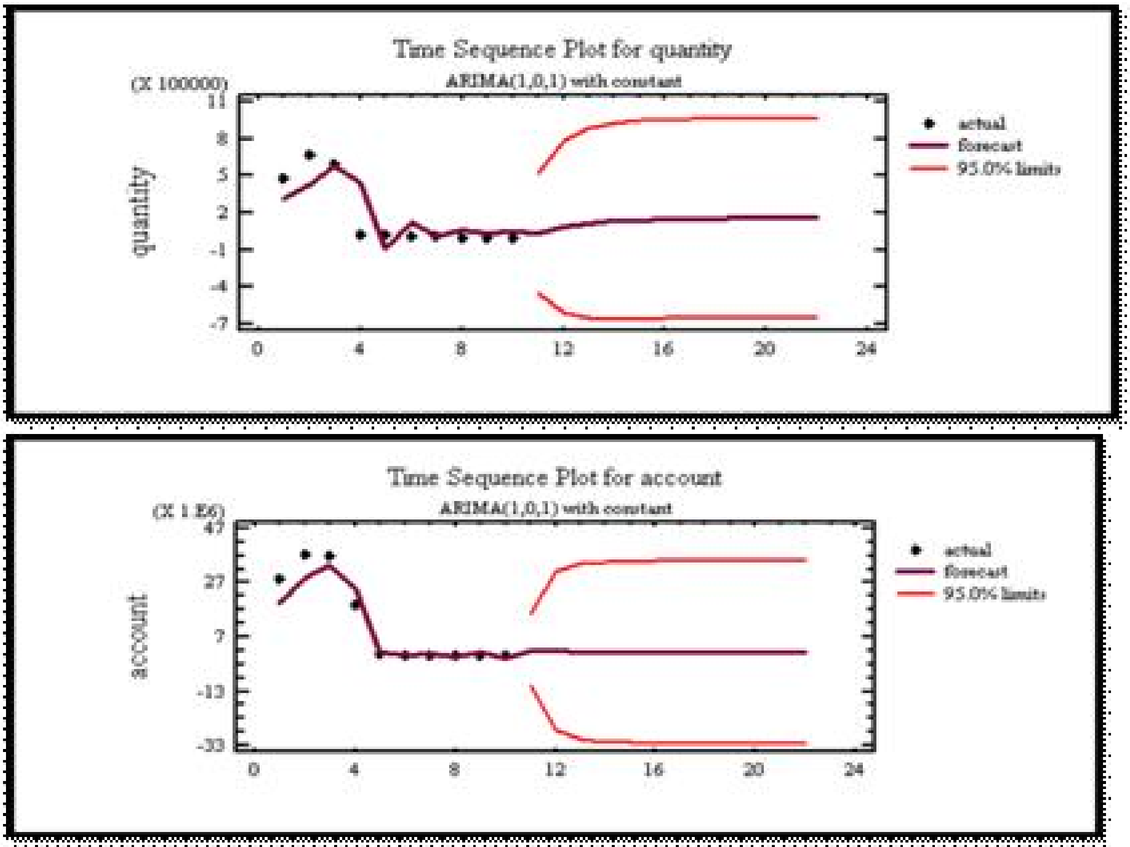

Fig. (11) Time sequence plot (quantity & account) with predicted periods by applying ARMA (1,1) model