EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 3, No. 3, 2010, 406-416
ISSN 1307-5543 – www.ejpam.com

SPECIAL ISSUE ON
GRANGER ECONOMETRICS AND STATISTICAL MODELING
DEDICATED TO THE MEMORY OF PROF. SIR CLIVE W.J. GRANGER

K-th Moving, Weighted and Exponential Moving Average for Time Series Forecasting Models

Chris P. Tsokos

Department of Mathematics and Statistics, University of South Florida, Tampa, FL, 33620

"What you leave behind is not what is engraved in stone monuments, but what is woven into the lives of others"—— Pericles

Abstract. The objective of the present study is to investigate the effectiveness of developing a forecasting model of a given nonstationary economic realization using a $k-th$ **moving average**, a $k-th$ **weighted moving average** and a $k-th$ **exponential weighted moving average process**. We create a new nonstationary time series from the original realization using the three different weighted methods. Using real economic data we formulate the best ARIMA model and compare short term forecasting results of the three proposed models with that of the classical ARIMA model.

2000 Mathematics Subject Classifications: 62M10, 91B84

Key Words and Phrases: time series, k-th moving average, k-th weighted moving average, k-th exponential weighted moving average

Preface

I personally first met Professor Granger when we were both invited in 1972 to the **International Symposium on "Mathematical Methods in Investment and Finance"**, in **Venice**, **Italy**. He lectured on "Empirical Studies of Capital Markets" and I on "Forecasting Models from Nonstationary Time Series: Short-Term Predictability of Stocks". I very much enjoyed the time spent together and the most interesting and inspiring discussion we had on various subjects related to the theme of the symposium. His truly outstanding contributions, especially in econometric nonlinear time series that were written with **simplicity**, **applicability** and **constructive interpretations**, will reflect his legacy.

Email address: `profcpt@cas.usf.edu`

1. Introduction

For the past forty years a significant amount of research efforts has been oriented in working with nonstationary time series especially for developing forecasting or predicting models for a large variety of problems that our global society encounter, such as health research, global warming, engineering, ecological, and educational, among others. We must recognize the important works of Clive W. J. Granger and his co-workers [1]-[8], for their significant contributions to the subject matter.

In the present study we review some basic concepts of time series that were recently investigated with Shou Hsing Shih, among others. [9]-[15]. Some other interesting references on the subject matter are [16]-[23]. Starting with a nonstationary time series of a given phenomenon, we analytically use a $k-th$ **moving average**, a $k-th$ **weighted moving average** and $k-th$ **exponential moving average**. We formulate the best possible ARIMA models for short term forecasting for each of the new structured time series using the usual optimal procedures for a real stock price data from the Fortune 500 list. A residual analysis comparison of these forecasting models is given. In addition, the classical best ARIMA model was also developed for the subject data, and the results were compared with the proposed models that we have introduced. In all cases the new models give better short term forecasting results than the classical ARIMA model.

2. Basic Review

The autoregressive process of order p combined with a moving average process of order q gives us the general ARIMA model of order (p,q). Since we usually work with nonstationary realizations, we use a difference filter of order d to reduce the nonstationary time series into a stationary format so that we can begin to develop our forecasting models. Thus, the common notation is given by ARIMA (p,d,q). For a given nonstationary time series $\{x_t\}$, we introduce the difference filter as

$$(1-B)^d$$

where $B^j x_t = x_{t-j}$, d being the degree of differencing of the series. Furthermore, we let $\pi(B)$ be the autoregressive operator that characterizes the behavior of the homogeneous nonstationary time series $\{x_t\}$. Thus,

$$\pi(B)(x_t - c) = \pi(B)x_t$$

for any constant c, which implies that

$$\pi(B) = \phi(B)(1-B)^d$$

for $d > 0$, where $\phi(B)$ is stationary autoregressive operator. Hence, with the appropriate d we reduce the nonstationary process into stationary. We can represent the autoregressive integrated moving average model ARIMA (p,d,q) as

$$\phi_p(B)(1-B)^d x_t = \theta_q(B)\varepsilon_t$$

or

$$\phi_p(B)(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)$$

and

$$\theta_q(B)=1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q,$$

where s and ϕ are the weighted operator of the model. This is the analytical format we follow to develop the actual ARIMA models for real data.

A popular criterion in evaluating the best fit model for a given time series is Akaike's, information criteria (AIC) [24]. The smallest value of AIC is considered the best fit model that results in the smallest average mean square error.

3. The $k-th$ Moving Average Time Series Model

In time series analysis, the use for the $k-th$ simple moving average is to usually smooth a given time series. It can also be used in discovering short-term, long-term trends and seasonal components of a given phenomenon. The $k-th$ simple moving average process of a time series $\{x_t\}$ is defined by

$$y_t=\frac{1}{k}\sum_{j=0}^{k-1}x_{t-k+1+j} \tag{1}$$

where $t=k,k+1,\ldots,n$.

It is clear that as k increases, the number of observations $\{y_t\}$decreases and the series $\{x_t\}$ gets closer and closer to the mean of the series $\{x_t\}$ as k increases. In addition, when $k=n$, the series $\{y_t\}$ reduces to a single observation, and is equal to the true mean μ. On the other hand, if we select a fairly small k, we can smooth the edges of the series without losing much of the general information.

We proceed to develop the new model by transforming the original time series $\{x_t\}$ into the new time series $\{y_t\}$ using (1). We begin the process of reducing the new time series, usually nonstationary, to stationary time series by selecting the appropriate differencing filter. We then proceed with the model building procedure to develop the best fit time series model, using the criteria of AIC to make our selection.

Once we have developed the forecasting model for the new time series $\{y_t\}$, we proceed to forecast values of $\{y_t\}$ and proceed to apply the back-shift operator to obtain estimates of the original phenomenon $\{x_t\}$, that is,

$$\hat{x}_t=k\hat{y}_t-x_{t-1}-x_{t-2}-\ldots-x_{t-k+1}. \tag{2}$$

We can summarize the process of developing the subject model as follows:

1. Transforming the original time series $\{x_t\}$ into $\{x_t\}$ by using (1).
2. Check for stationarity of the series $\{x_t\}$by determining the order of differencing d, where $d=0,1,2,\ldots$ according to the KPSS test, until we achieve stationarity.
3. Deciding the order m of process. For our case, we let $m=5$ where $p+q=m$.

4. After (d,m) has been selected, listing all possible set of (p,q) for $p+q \leq m$.

5. For each set of (p,q), estimating the parameters of each model, that is,

$$\phi_1, \phi_2, ...\phi_p, \theta_1, \theta_2, ..., \theta_q.$$

6. Compute the AIC for each model, and choose one with the smallest AIC.

7. Solve the estimates of original time series by using (2).

We shall illustrate the subject model using the following application.

We consider the actual daily closing price of a stock A, from the New York Stock Exchange for 500 days. The actual data is shown by Figure 1 below.

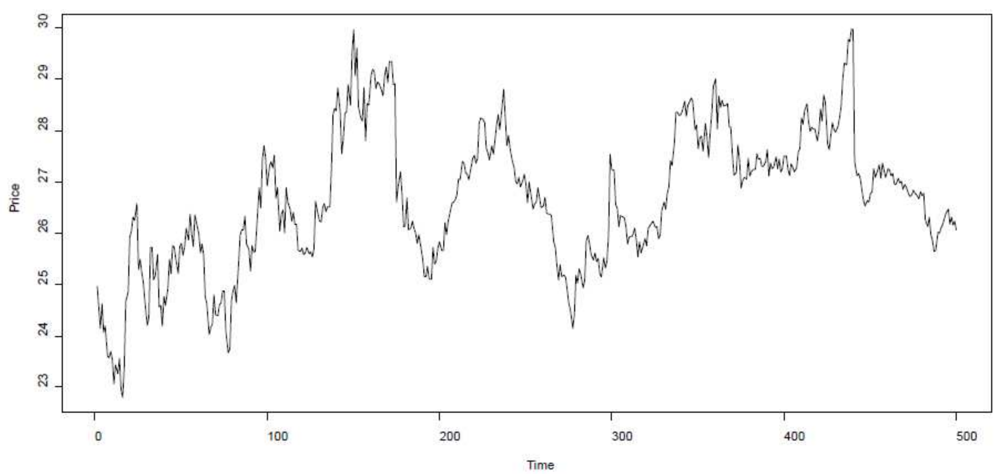

Figure 1: Daily Closing Price for Stock A.

Given $\{x_t\}$, $n = 500$ observations, we generate a new 3-day moving average time series $\{y_t\}$, using (1). Following the nine step procedure, we have identified the best model that characterizes the behavior of the new time series $\{y_t\}$, to be a combination of a second order autoregressive process with a third order moving average process that required a first order differencing filter. That is, ARIMA $(2,1,3)$.

The resulting model is given by

$$(1-.8961B-.0605B^2)(1-B)y_t = (1+.0056B-.0056B^2-B^3)\varepsilon_t.$$

Expanding the autoregressive operator and the first differencing filter, we have

$$(1-1.8961B+.8536B^2+.0605B^3)y_t = (1+.0056B-.0056B^2-B^3)\varepsilon_t.$$

The stationary model that estimated the new time series, $\{y_t\}$, can be written as

$$\hat{y}_t = 1.8961y_{t-1} - .8356y_{t-2} - .0605y_{t-3} + .0056\varepsilon_{t-1} - .0056\varepsilon_{t-2} - \varepsilon_{t-3}.$$

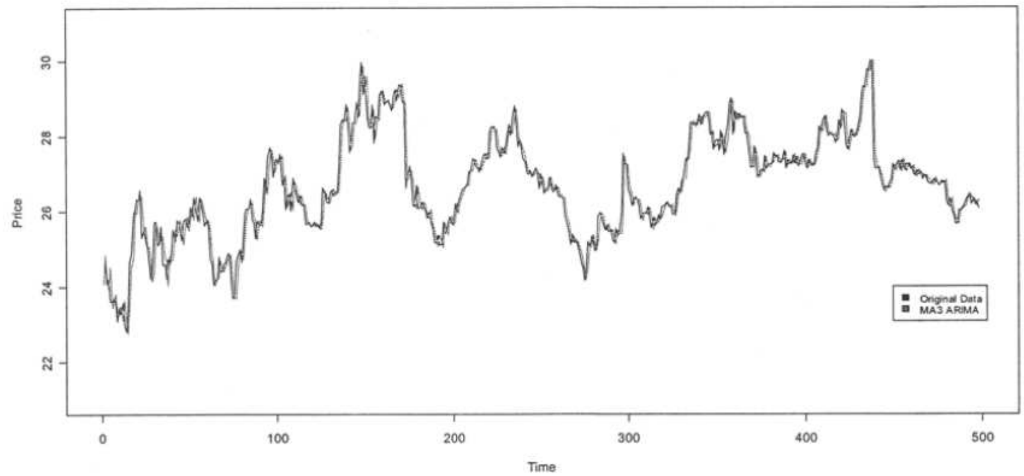

Figure 2: Comparison of 3-day Moving Average Model and the New Time Series.

Thus, for an estimating value of y_t we can obtain an estimate of the original time series x_t using (2). A graphical comparison of the original time series with that of the 3-day moving average model is shown by Figure 2. The proposed estimated model is quite close to the original time series. The residual analysis supports the visual comparison. That is, the average sample residual estimate is $\bar{r} = 0.010$, with a standard error of the residuals of 0.017.

Thus, having estimates of $\{y_t\}$ we can obtain estimates of the original$\{x_t\}$ using the defined functional forms given by expression (2). These results will also be compared with the classical ARIMA forecasting model.

4. The $k-th$ Weighted Moving Average Time Series Model

The $k-th$ weighted moving average process is also a good smoothing procedure. The structure of it is slightly different from the simple moving average process. It puts more weight on the most recent observation, and the weight consistently decreases up to the initial observation. Also, it captures the original realized series better than the moving average process. The subject model supports the fact that the recent observations should weigh more than the initial ones. The $k-th$ moving average process of a given time series $\{x_t\}$ is defined as follows,

$$z_t = \frac{1}{(1+k)k/2} \sum_{j=0}^{k-1} (j+1) x_{t-k+1+j} \tag{3}$$

Where $t = k, k+1, \ldots, n$. Similar to the moving average process, as k increases, the number of realization of the series $\{z_t\}$ decreases, as $k \to n$. From (3) the new time series $\{z_t\}$ becomes

$$z_t = \frac{1}{(1+n)n/2} \sum_{j=0}^{k-1} j x_j. \tag{4}$$

For a small k, we can smooth the edges of the time series, and the new realization $\{z_t\}$ is closer to the actual series$\{x_t\}$.

Thus, we proceed using (3) to create the new time series $\{z_t\}$, and we begin the process of reducing the nonstationary time series to stationary and develop the best fit forecasting model using the same criteria as in the previous proposed model.

Now, we begin the nine step procedure we discussed for the $k-th$ moving average model to obtain an estimate of the present model.

Using the developed model we can forecast values of $\{z_t\}$ and proceed to apply the back-shift operator to obtain the estimates of the original realization$\{x_t\}$, that is,

$$\hat{x}_t = \frac{[(1+k)k/2]\hat{z}_t - (k-1)x_{t-1} - (k-2)x_{t-2} - ... - x_{t-k+1}}{k}. \tag{5}$$

The usefulness of the $k-th$ weighted moving average model will be illustrated for comparison purposes with the same closing price of stock A from the New York Stock Exchange.

We shall use the same data as in the precious illustration to develop the proposed model. Following the recommended procedure, the new nonlinear time series $\{z_t\}$, best fits an ARIMA $(1,1,3)$. That is,

$$(1+.9037B)(1-B)z_t = (1+1.5084B-.08348B^2+.2456B^3)\varepsilon_t.$$

Expanding the autoregressive operator and the differencing filter, we have

$$(1-.0927B-.9037B^2)z_t = (1+1.5048B+.8348B^2+.2456B^3)\varepsilon_t$$

or

$$\hat{z}_t = .0927z_{t-1} + .9073z_{t-2} + \varepsilon_i + 1.5084\varepsilon_{t-1} + .8348\varepsilon_{t-2} + .2456\varepsilon_{t-3}. \tag{6}$$

Thus, the estimated time series model of the new time series $\{z_t\}$ is given by

$$\hat{z}_t = .0927z_{t-1} + .9073z_{t-2} + \varepsilon_i + 1.5084\varepsilon_{t-1} + .8348\varepsilon_{t-2} + .2456\varepsilon_{t-3}. \tag{7}$$

Using equation (7) we obtain estimates of z_t and proceed to use these estimates in expression (5) to obtain estimates of the original time series of the closing price of stock A. A graph of the actual data of $\{z_t\}$ with that estimated by the proposed model $\{\hat{x}_t\}$ is given below by Figure 3. Again, the present model fits quite well with the original observations. The average sample residual is $\bar{r} = 0.0087$ with a sample residual standard error of 0.018. Thus, this proposed model is slightly better than the previous model.

5. The $k-th$ Exponential Weighted Moving Average Time Series Model

The k-days exponential weighted moving average process, in addition to what the previous two models offer, instead of decreasing weight consistently as the weighted moving average method does, it decreases the weight exponentially. That is, we put much more emphasis on the most recent observations and we are not much concerned with the older observations.

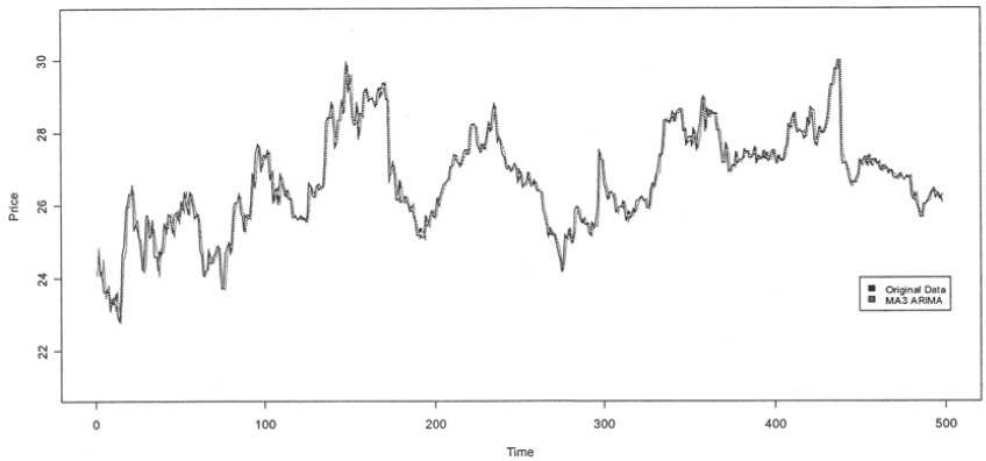

Figure 3: The Weighted 3-day Moving Average Model and New Time Series.

The $k-th$ exponential weighted moving average process of a given time series $\{x_t\}$ is defined by

$$v_t = \frac{1}{\sum_{j=0}^{k-1}(1-\alpha)^j}\sum_{j=0}^{k-1}(1-\alpha)^{k-j-1}x_{t-k+1+j} \tag{8}$$

where $t = k, k+1, \ldots, n$, and the smoothing factor α is defined as $\alpha = \frac{2}{k+1}$. If we let $k = n$, we have $\alpha = \frac{2}{n+1}$. Moreover, $\sum_{j=0}^{k-1}(1-\alpha)^j$ reaches its maximum when $k = 3$, and it gets closer and closer to 1 as k increases. As k increases, the number of observations of the new time series $\{v_t\}$ decreases, and it eventually reduces to a single observation when $k = n$. As $k \to n$, the time series $\{v_t\}$ becomes

$$v_t = \frac{1}{\sum_{j=0}^{n-1}(1-\alpha)^j}\sum_{j=0}^{n-1}(1-\alpha)^{n-j-1}x_{j+1}. \tag{9}$$

It is clear that the exponential weighted moving average process weights heavily on the most recent observation, and decreases the weight exponentially as time decreases. Also, if we choose a fairly small k, we can smooth the edges of the time series, as $\{v_t\}$ would be fairly close to the original time series.

As before, we proceed to develop this model by transforming the original time series $\{x_t\}$ into the new time series$\{v_t\}$ by applying (8). After obtaining the new time series, usually nonstationary, we follow the same procedure as previously stated to obtain the best possible model for$\{v_t\}$, by following the nine step procedure as in the first model. Having developed an estimate of$\{\hat{v}_t\}$we apply the back-shift operator to obtain estimates of the original phenomenon $\{x_t\}$, that is,

$$\hat{x}_t = \frac{\hat{v}_t}{\sum_{j=0}^{k-1}(1-\alpha)^j} - (1-\alpha)x_{t-1} - (1-\alpha)^2 x_{t-2} - \ldots - (1-\alpha)^{k-1}x_{t-k-1}. \tag{10}$$

Again for comparison purposes we shall illustrate the development of the subject model using the same data for stock A.

The best model of the newly generated time series $\{v_t\}$ is given by ARIMA $(3,1,2)$, that is, a third order autoregressive process combined with a second order moving average process with a first differencing filter. The resulting model is given by

$$(1-.4766B-.9054B^2)(1-B)v_t=(1-.04362B-.0728B^2-.9071B^3)\varepsilon_t. \qquad (11)$$

Expanding the autoregressive operator in (4) and the differencing filter, we have

$$(1-1.4766B-.4279B^2+.9045B^3)v_t=(1-.04362B-.0728B^2-.9071B^3)\varepsilon_t.$$

or

$$\hat{v}_t=1.4766v_{t-1}+.4279v_{t-2}-.9045v_{t-3}+\varepsilon_i-.4362\varepsilon_{t-1}-.0728\varepsilon_{t-2}-.9071\varepsilon_{t-3}.$$

The form of the actual estimates of $\hat{v}_t$ is given by

$$\hat{v}_t=1.4766v_{t-1}+.4279v_{t-2}-.9045v_{t-3}+\varepsilon_i-.4362\varepsilon_{t-1}-.0728\varepsilon_{t-2}-.9071\varepsilon_{t-3}. \quad (12)$$

Using the above equation (12), we can obtain estimates of v_t and proceed to use those estimates in expression (10) to obtain estimates of the original form of the data. A graphical comparison of the original stock A with the corresponding estimates is given by Figure 4 below. Again, the proposed model is quite close to the original realization of daily closing price

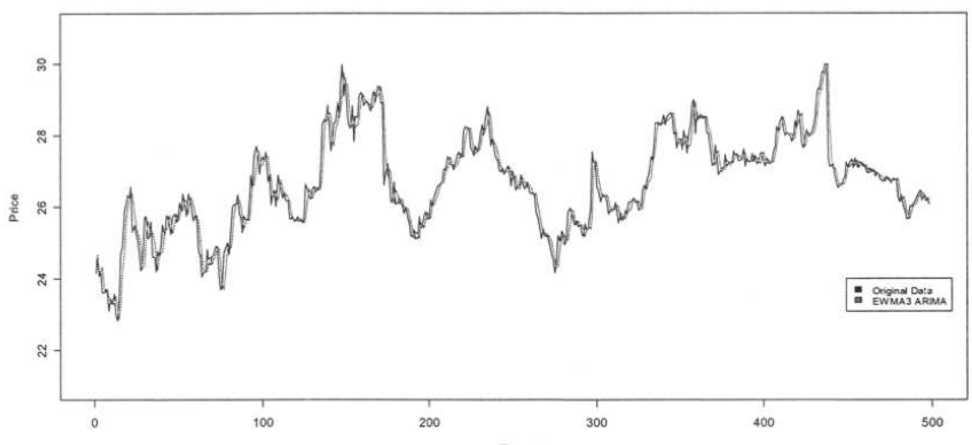

Figure 4: Comparison of the 3-day Exponential Weighted Moving Average Model and Original Time Series.

of Stock A. The residual analysis results is $\bar{r}=0.008$ with residual standard error of 0.017. This model seems to perform as well as the 3-day weighted moving average and slightly better than the 3-day moving average forecasting model. A comparison with the classical ARIMA model will be also given.

6. Classical ARIMA Model

The best fit of the stock A closing price is ARIMA $(0, 1, 2)$, a second order moving average with a first difference filter. It can be written as

$$(1 - B)x_t = (1 - 0.033B - 0.11B^2)\varepsilon_t. \tag{13}$$

Expanding the autoregressive operator and the difference filter, we have

$$x_t - x_{t-1} = (1 - 0.033B - 0.11B^2)\varepsilon_t. \tag{14}$$

The model for one day ahead forecasting time series of the closing price of stock A is given by

$$x_t = x_{t-1} - 0.033\varepsilon_{t-1} - 0.11\varepsilon_{t-2}. \tag{15}$$

A graph of the forecasting value obtained using (15) with the original time series is given below by Figure 5. Visually we cannot distinguish between the two graphs. However, the

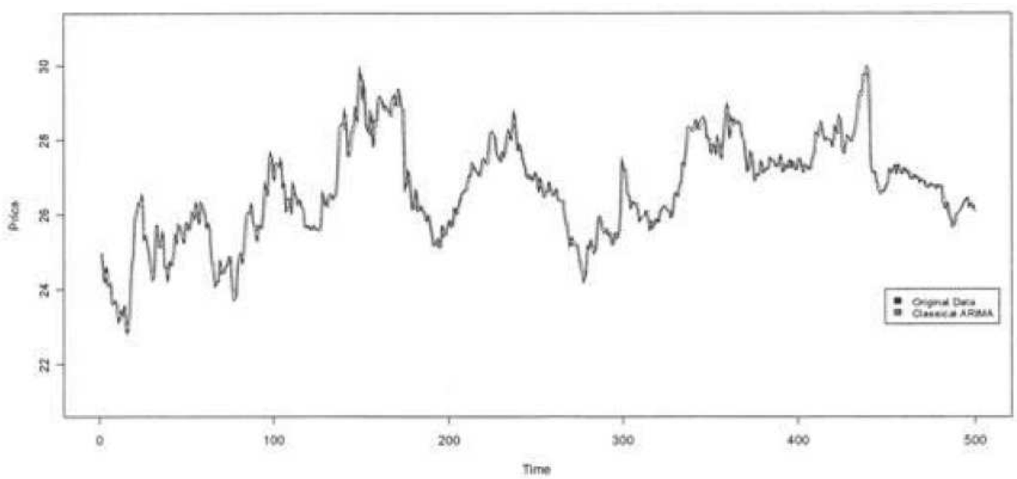

Figure 5: Comparisons of Classical ARIMA Model and the Original Time Series.

average residual $\bar{r} = 0.57$ with a residual standard error of 0.35, is greater than the other three proposed methods. Given in Table 1 is a summary of the residual analysis in comparison of the proposed three models with the classical ARIMA process. Thus, we can conclude that

Table 1: Comparison of Mean and Standard Error of Residuals.

	Mean Residual $\bar{r}$	Residual Standard Error
3-day Moving ARIMA	0.010	0.017
3-day Weighted ARIMA	0.008	0.018
3-day Exponential Weighted ARIMA	0.008	0.017
Classical ARIMA	0.570	0.350

the 3-day average, 3-day weighted average and the 3-day exponential weighted average will perform better than the classical ARIMA model.

References

[1] C Granger. “The typical spectral shape of an economic variable”. *Econometrica* 34: 150-161, 1966.

[2] C Granger. “Investigating causal relations by econometric models and cross-spectral methods”. *Econometrica* 37: 424-438, 1969

[3] C Granger and J Bates. “The combination of forecasts”. *Operational Research Quarterly* 20: 451-468, 1969.

[4] C Granger and M Hatanaka. *Spectral Analysis of Economic Time Series*. Princeton University Press, Princeton, NJ, 1964.

[5] C Granger and R Joyeux. “An introduction to long-memory time series models and fractional differencing”. *Journal of Time Series Analysis* 1: 15-30, 1980.

[6] C Granger and P Newbold. “Spurious regressions in econometric”. *Journal of Econometrics* 2: 111-120, 1974

[7] C Granger and P Newbold. *Forecasting Economic Time Series.* Academic Press; second edition: 1986.

[8] R Engle and C Granger. “Co-integration and error-correlation: Representation estimation and testing”. *Econometrica* 55:251-276, 1987

[9] C Tsokos. “Forecasting Models for Nonstationary time series-short-term predictability”*Mathematical Methods in Investment and Finance*, North Holland Publishing Company, 1971.

[10] S Shish and C Tsokos. “A weighted moving average process for forecasting”. *Journal of Modern Applied Statistical Methods*, Vol 16, No. 2, 2008.

[11] S Shish and C Tsokos. “Analytical models for economic forecasting”. *Proceedings of the 5th International Conference on Dynamic System*, 2008.

[12] S Shish and C Tsokos. “New nonstationary time series models with economic applications”. *Proceedings of the 5th International Conference on Dynamic System*, 2008.

[13] S Shish and C Tsokos. “A temperature forecasting model for the emissions and the atmosphere”. *International Journal Neural,Paralle;& Scientific Computations*, Vol 16, 2008.

[14] S Shish and C Tsokos. “Prediction models for carbon dioxide emissions and the atmosphere”. *International Journal Neural,Paralle;& Scientific Computations*, Vol 16.No.1, 2008.

[15] C Tsokos. “Forecasting models from nonstationary time series-short term predictability of stocks”, *Mathematical Methods in Investment and Finance*, North Holland Publishing Co., 520-63, 1973.

[16] G Box, G Jenkins, and G Reinsel. *Time Series Forecasting and Control*, Holden-Day, San Fransisco, 1970.

[17] G Box, G Jenkins, and G Reinsel. *Time Series Analysis: Forecasting and Control*, 3rd ed., Prentice Hall, Englewood Cliffs, NJ., 89-99,224-247, 1994.

[18] T Bollerslev. “Generalized autoregressive conditional heteroskedasticity”, *Journal of Econometrics*, 31, 307-327, 1986.

[19] P Brockwell and R Davis. *Introduction to Times Series and Forecasting,* Springer, New York, Sections 3.3 and 8.3, 1996.

[20] R Brown. *Smoothing, Forecasting, and Prediction of Discrete Time Series*, Prentice Hall, New Jersey, 1962.

[21] D Crane and J Eratty. “*A two stage forecasting model: exponential smoothing and multiple regression*”, Management Science, Vol 13, No.8, 1967.

[22] W Ebders. *Applied Econometrics Time Series*, John-Wiley & Sons, 139-149, 1995.

[23] A Fox. “Outline in time series”, *Journal of Royal Stat. Society*, Series B, Vol.34, No.3, 1972.

[24] H Akaike. “A new look at the statistical model identifications”, *IEEE Transactions on Automatic Control*, AC-19, 716-723, 1974.