Differencing a Time Series and Modifications to the Variogram

B. Abraham
&
N. Balakrishna
University of Waterloo

RR-01-02
Revised Version of RR-99-09

DIFFERENCING A TIME SERIES AND MODIFICATIONS TO THE VARIOGRAM

B. Abraham and N. Balakrishna[1]
University of Waterloo

Abstract

In the context of building an Autoregressive Integrated Moving Average (ARIMA) model, the specification of the degree of differencing (d) is very important. In this paper we consider the variogram and some modification of it to specify d. The behaviour of these is illustrated by a simulation study. We also use these to specify d in five well known time series.

Keywords: ARIMA process, Differencing, Simulation, Stationarity, Variogram

1. Introduction

The theory and the analysis of time series becomes easier under the assumption that an observed series is a realization from a stationary stochastic process. But in practice, a great deal of time series data is best described by non-stationary processes. While analysing discrete time series data, it is often assumed that an observed series is a realization of some stochastic process which becomes mean stationary (homogenous stationary in the sense used in Box and Jenkins (1976)) after sufficient number of differencing. In this paper we discuss a simple method of determining the degree of differencing.

Let d be the number of differencing required for stationarity and B denote the back shift operator such that $B^h z_t = z_{t-h}$. Suppose that $\{a_t\}$ is a white noise process with common mean zero and common variance σ^2. Then

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)(1 - B)^d z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t, \qquad (1.1)$$

$$\text{or} \quad \phi(B)(1 - B)^d z_t = \theta(B) a_t$$

defines an autoregressive integrated moving average model of order (p, d, q) (ARIMA (p, d, q)). That is, after differencing 'd' times, the process $\{z_t\}$ in (1.1) becomes a stationary ARMA (p, q). Our interest here is to determine the value of d for a given time series.

[1]Cochin University of Science and Technology, Cochin, 682 022 India. Work was done while visiting the University of Waterloo.

Box and Jenkins (1976) discuss a method for determining 'd' using the sample autocorrelation and partial autocorrelation plots. Cressie (1988) used the concept of generalized covariance to introduce a graphical procedure for determining d. For $h = 1, 2, ...,$ he calls $z_{t+h} - z_t$ as the I_1-process and discuss the variogram, $\text{Var}(z_{t+h} - z_t) = 2V(h)$ for a stationary process. In addition he introduces an I_d-process a quantity $\lambda' z$ where λ is a vector depending on h and and d. For example, when the process is I_2,

$$\lambda' z = hz_t - (h+1)z_{t+1} + z_{t+h+1}$$

where

$$\lambda' = [h \quad -(h+1) \quad 1]$$

For an I_3 process

$$\begin{aligned} \lambda' z &= -h(h+1)z_t + 2h(h+2)z_{t+1} \\ &\quad -(h+2)(h+1)z_{t+2} + 2z_{t+h+2}. \end{aligned}$$

where

$$\lambda' = [-h(h+1) \quad 2h(h+2) \quad -(h+2)(h+1) \quad 2]$$

He also constructs some unbiased estimators for $\text{Var}(\lambda' z)$. Based on these, some graphical procedures are introduced to determine the value of d (the degree of differencing).

In this paper we present a simple graphical procedure to find 'd' based on some modifications to the variogram. Section 2 describes the procedure, section 3 considers a simulation study to confirm the patterns. Section 4 looks at some known examples and see how the procedure suggested here arrives at the degree of differencing. Section 5 gives some concluding remarks.

2. Methodology

Suppose the process z_t is stationary; let us consider the length h difference

$$u_{1t}(h) = z_{t+h} - z_t = (1 - B^h)z_{t+h}. \tag{2.1}$$

Then $V_1(h) = \text{Var}(u_{1t}(h)) = 2(\gamma_0 - \gamma_h)$, where $\gamma_h = \text{Cov}(z_{t+h}, z_t)$.

If the process is only stationary after a difference is taken then $\gamma_0 = \text{Var}(z_t)$ will not exist. However, $V_1(h)$ will exist and will be a function of h for finite h.

For example if the process is

$$(1-B)z_t = (1-\theta B)a_t, \tag{2.2}$$

then we can show that (see Box and Kramer (1992))

$$V_1(h) = \text{Var}(u_{1t}(h)) = [(h-1)(1-\theta)^2 + (1+\theta^2)]\sigma^2, \tag{2.3}$$

a linear function of h. Box and Lucerno (1997) plots $V_1(h)/V_1(1) = 1 + \frac{(h-1)(1-\theta)^2}{1+\theta^2}$ versus h and calls it the variogram. We adopts this terminology. It is a linear function of h with slope $(1-\theta)^2/(1+\theta^2)$. When $h = 1$, $V_1(h)/V_1(1) = 1$. In particular if $z_t = a_t$ (white noise) then $V_1(h)/V_1(1) = 1$ for all h. On the other hand if $(1-\phi B)z_t = a_t$ (i.e. z_t an autoregressive process of order 1) then $V_1(h)/V_1(1) = (1-\phi^h)/(1-\phi)$ and this flattens out (converges to a constant) as h becomes large.

If the process is only 'stationary' after a second difference is taken then $\text{Var}(u_{1t}(h))$ does not exist. However, the variance of

$$u_{2t}(h) = (1-B)(1-B^{h-1})z_{t+h} \tag{2.4}$$

exists. For example if the process is

$$(1-B)^2 z_t = (1-\theta_1 B - \theta_2 B^2)a_t, \tag{2.5}$$

then

$$V_2(h) = \text{Var}(u_{2t}(h)) = \begin{cases} (1+\theta_1^2+\theta_2^2)\sigma^2 & \text{if } h = 2 \\ \{1+(1-\theta_1)^2+\theta_2^2+(\theta_1+\theta_2)^2 & \\ +(h-3)(1-\theta_1-\theta_2)^2\}\sigma^2 & \text{if } h \geq 3, \end{cases} \tag{2.6}$$

which is a linear function of h for $h > 2$.

It should also be noted that if the process $\{z_t\}$ is stationary moving average (i.e. $(0,0,q)$ process) then $\text{Var}(u_{1t}(h))$ and $\text{Var}(u_{2t}(h))$ are constants. In addition if $\{(1-B)z_t\}$ is stationary MA (i.e. $(0,1,q)$) then $\text{Var}(u_{2t}(h))$ is constant. Similarly if $\{(1-B)^2 z_t\}$ is stationary MA then $\text{Var}(u_{2t}(h))$ is linear in h; but $V_3(h) = \text{Var}(u_{3t}(h))$ where

$$u_{3t}(h) = (1-B)^2(z_{t+h} - z_{t+2}) = (1-B)^2(1-B^{h-2})z_{t+h} \tag{2.7}$$

is constant. For a general ARIMA $(p,1,q)$ process with

$$\theta(B)/\phi(B) = \psi(B) = \sum_{j=0}^{\infty} \psi_j B^j, \quad \psi_0 = 1,$$

so that $(1-B)z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$, $\sum_{j=0}^{\infty} |\psi_j| < \infty$ we can show that

$$V_1(h) = \left[\sum_{j=0}^{h-1} \left(\sum_{i=0}^{j} \psi_i \right)^2 + \sum_{j=0}^{\infty} \left(\sum_{i=1}^{h} \psi_{i+j} \right)^2 \right] \sigma^2, \quad h \geq 1 \tag{2.8}$$

and

$$V_2(h) = \left[\sum_{i=0}^{h-2} \psi_i^2 + \sum_{j=0}^{\infty} (\psi_{h+j-i} - \psi_j)^2 \right] \sigma^2, \quad h \geq 2. \tag{2.9}$$

If we set $\phi(B) \equiv 1$ then for an ARIMA $(0,1,q)$ process the variances are given by

$$V_1(h) = \left[1 + \sum_{j=1}^{q-1} \left\{ 1 - \sum_{i=1}^{j} \theta_i \right\}^2 + \sum_{j=1}^{q} \left(\sum_{i=j}^{q} \theta_i \right)^2 + \left\{ 1 - \sum_{i=1}^{q} \theta_i \right\}^2 (h-q) \right] \sigma^2 \text{ if } h \geq q \tag{2.10}$$

and

$$V_2(h) = \begin{cases} \left[1 + \theta_1^2 + 2(\theta_2^2 + \cdots + \theta_{q-2}^2) + (1-\theta_{q-1})^2 + (\theta_1 - \theta_q)^2 + \theta_{q-1}^2 + \theta_q^2 \right] \sigma^2 \text{ if } h = q \\ \left[1 + 2(\theta_1^2 + \theta_2^2 + \cdots + \theta_{q-1}^2) + (1+\theta_q)^2 + \theta_q^2 \right] \sigma^2 \text{ if } h = q+1 \\ 2 \left[1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2 \right] \sigma^2 \text{ if } h > q+1. \end{cases} \tag{2.11}$$

Thus $V_1(h) = [C_1 + (h-q)C_2]\sigma^2$ for $h \geq q$ which is linear in h and

$$V_2(h) = C_3 \sigma^2, \quad h > q+1$$

where C_1, C_2, C_3 do not involve h.

On the other hand if the process is ARIMA $(p,2,q)$ then $Var(u_{1t}(h))$ does not exist. But we can compute

$$V_2(h) = \left[\left(\sum_{j=0}^{h-3} \psi_j \right)^2 + \sum_{j=0}^{\infty} \left(\sum_{i=0}^{h-2} \psi_{j+i} \right)^2 \right] \sigma^2$$

and

$$V_3(h) = \left[\sum_{j=0}^{h-3} \psi_j^2 + \sum_{j=0}^{\infty} \{ \psi_{h+j-2} + \psi_j \}^2 \right] \sigma^2.$$

In this case if we take $\phi(B) \equiv 1$ then the corresponding variances for the ARIMA $(0,2,q)$ are given by

$$V_2(h) = \left[1 + \sum_{j=1}^{q-1} \left\{1 - \sum_{i=1}^{j} \theta_i\right\}^2 + \sum_{j=1}^{q} \left(\sum_{i=j}^{q} \theta_i\right)^2 + \left\{1 - \sum_{i=1}^{q} \theta_i\right\}^2 (h-q-1)\right] \sigma^2 \text{ if } h \geq q+1$$

and

$$V_3(h) = \begin{cases} \left[1 + \theta_1^2 + \cdots + \theta_{q-3}^2 + (1+\theta_{q-2})^2 + (\theta_1 - \theta_{q-1})^2 + (\theta_2 - \theta_q)^2 + \theta_3 + \cdots + \theta_q^2\right] \sigma^2 \text{ if } h = q \\ \left[1 + \theta_1^2 + \cdots + \theta_{q-2}^2 + (1+\theta_{q-1})^2 + (\theta_1 - \theta_q)^2 + \theta_2^2 + \cdots + \theta_q^2\right] \sigma^2 \text{ if } h = q+1 \\ \left[1 + \theta_1^2 + \cdots + \theta_{q-1}^2 + (1+\theta_q)^2 + \theta_1^2 + \cdots + \theta_q^2\right] \sigma^2 \text{ if } h = q+2 \\ 2(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2 \text{ if } h \geq q+3. \end{cases}$$

Hence $V_2(h)$ is linear in h for $h \geq q+1$ and $V_3(h)$ is constant for $h > q+2$.

In practice, since it is unusual to see $d > 2$, we will not pursue this further. Expressions for $V_i(h)$, $i = 1,2,3$ for some processes are given in the Appendix (see Table A.1). From these the patterns given in Table 2.1 can be specified.

Table 2.1 Patterns of V_i(h), i=1, 2, 3

Case	Process	Patterns	
1.	Stationary Process Based on ARIMA(1,0,1) (1,0,0), (0,0,1)	$V_1(h)$	A constant for all $h \geq k$ where k depends on the particular process or it stabilizes to a constant as h increases.
		$V_2(h)$	A constant after some initial values.
		$V_3(h)$	A constant after some initial values.
2.	Nonstationary Process Based on ARIMA(0,1,1),(1,1,0) (1,1,1)	$V_1(h)$	A linear function of h for $h \geq k$ where k depends on the particular process or it stabilizes to a linear function.
		$V_2(h)$	Same pattern as $V_1(h)$ in Case 1.
		$V_3(h)$	Same pattern as $V_1(h)$ in Case 1.
3.	Nonstationary Process Based on ARIMA(2,2,2) (2,2,0), (0,2,2)	$V_1(h)$	Shows nonlinear tendency.
		$V_2(h)$	Same as $V_1(h)$ in Case 2.
		$V_3(h)$	Same as $V_1(h)$ in Case 1.

For estimation of $\text{Var}(u_{it}(h))$, we consider the sample variances of $u_{it}(h)$, $i = 1,2,3$

respectively. Thus we have,

$$S_i^2(h) = \sum_{t=1}^{n-h} \{u_{it}(h) - \overline{u}_i(h)\}^2/(n-h), \quad i = 1, 2, 3.$$

Using the results in the Appendix it can be shown that

(i) when $d = 1$, for finite h as $n \to \infty$

$$E(S_i^2(h)) \to V_i(h) \quad i = 1, 2, 3.$$

(ii) when $d = 2$, for finite h as $n \to \infty$

$$E(S_1^2(h)) \to \infty$$
$$E(S_i^2(h)) \to V_i(h) \quad i = 2, 3.$$

For convenience we consider the quantities

$$R_i(h) = S_i^2(h)/S_i^2(i), \ i = 1, 2, 3 \ .$$

Then we can adopt the following procedure for specifying d. For a given time series, plot $R_i(h)$ versus h $i = 1, 2, 3$. For $R_1(h)$ vs. h we consider $h = 1$ to 15 for $R_2(h)$ vs. h we take $h = 2, ..., 15$ and for $R_3(h)$ vs. h we take $h = 3$ to 15. If $R_1(h), R_2(h), R_3(h)$ are approximately constants or converging to constants then $d = 0$. If $R_1(h)$ is linear in h or linear in h after some initial periods, and the other two are constants (approximately) then $d = 1$. If $R_1(h)$ seems to increase nonlinearly, but $R_2(h)$ is linear in h (or linear after some initial period) or approaching a constant and $R_3(h)$ is approximately constant then $d = 2$. Thus this simple procedure can serve as a useful tool in deciding the degree of differencing. It can be added easily to any software dealing with time series.

3. Simulation Study

Assuming that $\text{Var}(a_t) = \sigma^2 = 1$, we generate 150 observations from a prescribed ARIMA process with specified parameter values and discard the first 50 to avoid transients. From the remaining 100 observations we compute $R_i(h)$ for $h = i, i+1, ..., 15$, $i = 1, 2, 3$. Then

the whole procedure is repeated 1000 times and averages of these quantities over the repetitions are computed. We consider several processes. Plots of some of these are shown in Figures 3.1-3.6. The emerging patterns are summarized in Table 3.1. For example Figure 3.3 corresponds to a $(0,1,1)$ process with $\theta = .5$. As expected the average $R_1(h)$ is increasing linearly in h, $R_2(h)$ is a constant for $h = 3,4,...$ and $R_3(h)$ is a constant for $h = 5,6,7,...$. Figure 3.5 describes the results for a $(0,2,2)$ process. In this case $R_1(h)$ appears to increase in a nonlinear fashion, $R_2(h)$ is linear in h for $h = 3,4,...,$ and $R_3(h)$ is a constant for $h = 5,6,7,...$. We generated many plots corresponding to different processes and only a few are included here to save space. In general the sample variances mimic the behaviour of the theoretical variances in Table A.1 on the average and these will help in determining the degree of differencing.

Table 3.1 Patterns of $R_i(h) = S_i^2(h)/S_i^2(i)$

Case	Process	Patterns	
1.	Stationary No differencing Fig. 3.1, 3.2	$R_1(h)$	Either a constant from $h = 1$ onwards or after some initial periods. In some cases it approaches a constant as h increases.
		$R_2(h)$	Constant after some periods.
		$R_3(h)$	Constant after some periods.
2.	Nonstationary $d = 1$ Fig. 3.3, 3.4	$R_1(h)$	Linear function of h for $h \geq$ some fixed value or it approaches a linear function after some periods.
		$R_2(h)$	Same pattern as $R_1(h)$ in Case 1.
		$R_3(h)$	Same pattern as $R_2(h)$ in Case 1.
3.	Nonstationary $d = 2$ Fig. 3.5, 3.6	$R_1(h)$	Non linear.
		$R_2(h)$	Same pattern as $R_1(h)$ in Case 1.
		$R_3(h)$	Same pattern as $R_2(h)$ in Case 1.

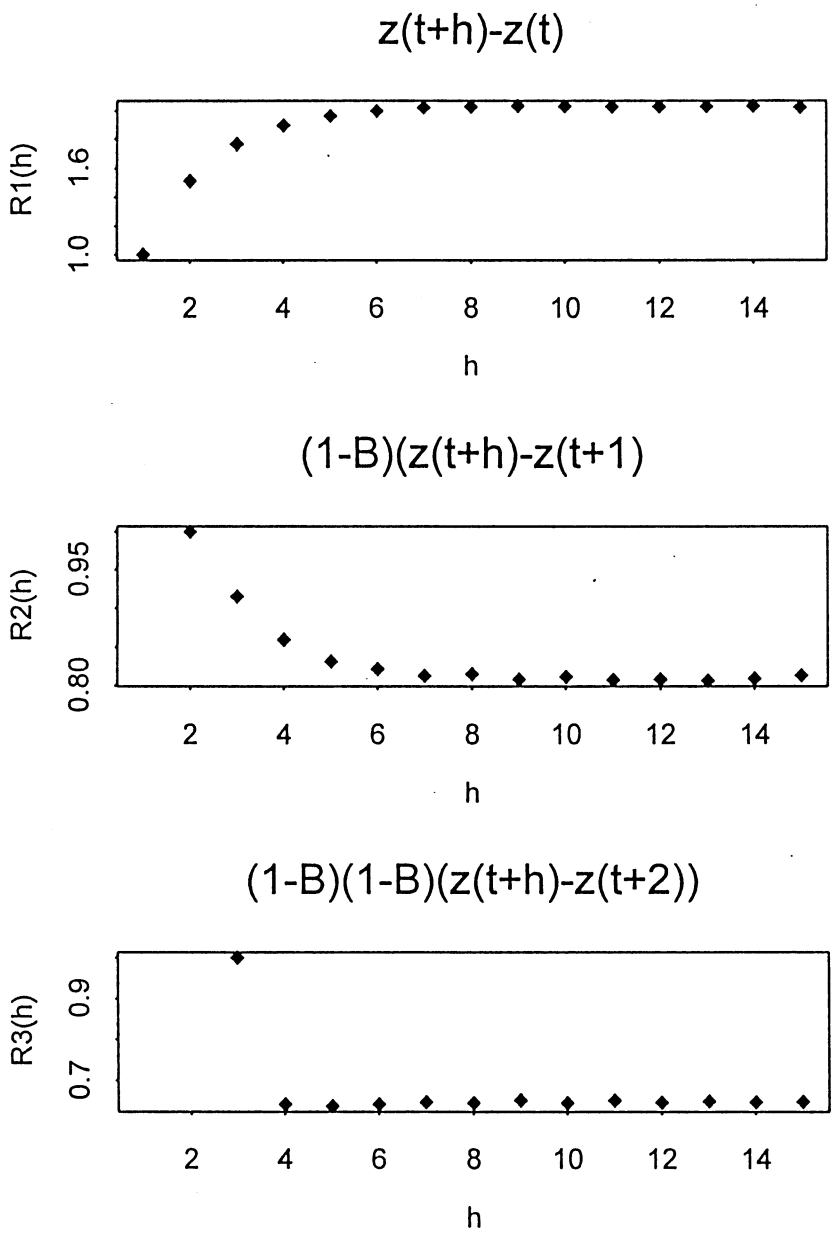


Figure 3.1: ARIMA(1,0,0), Phi=.5

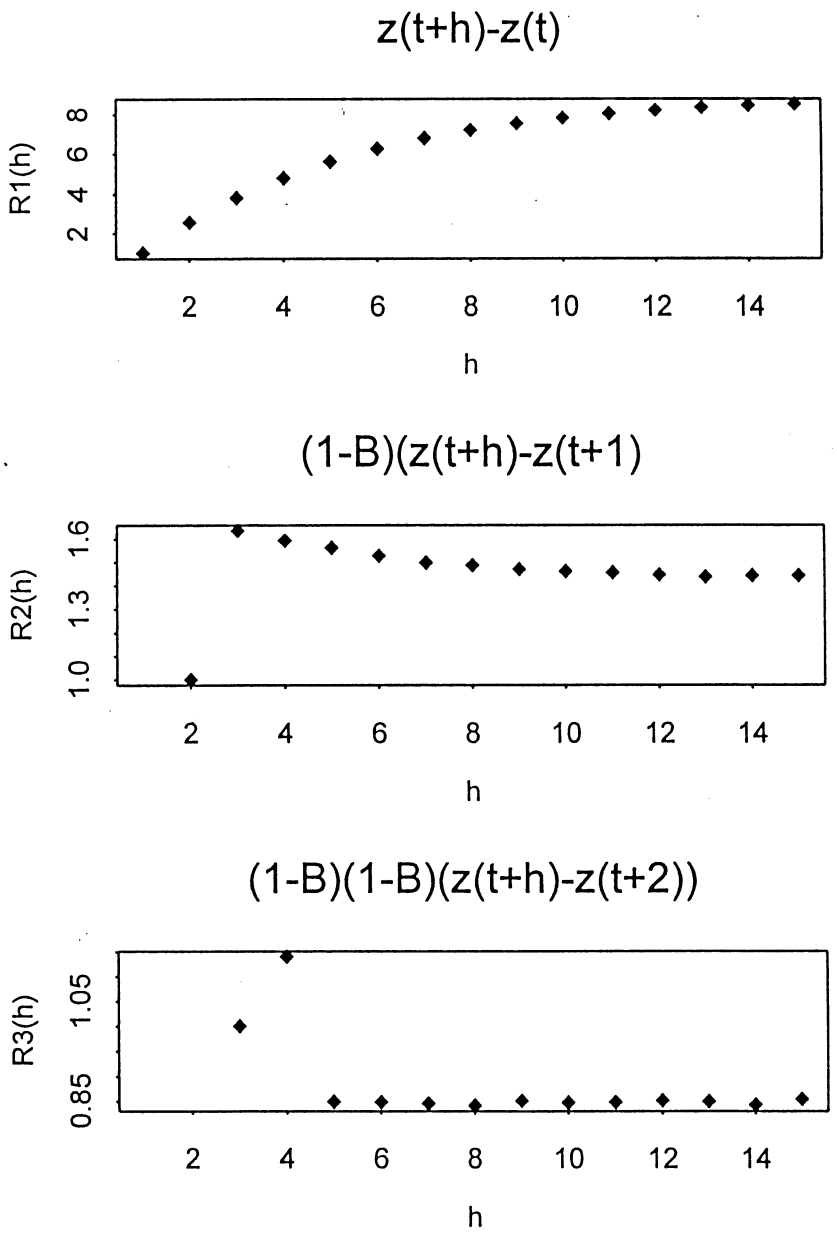


Figure 3.2: ARIMA(1,0,1), Phi=.8, Theta=-.5

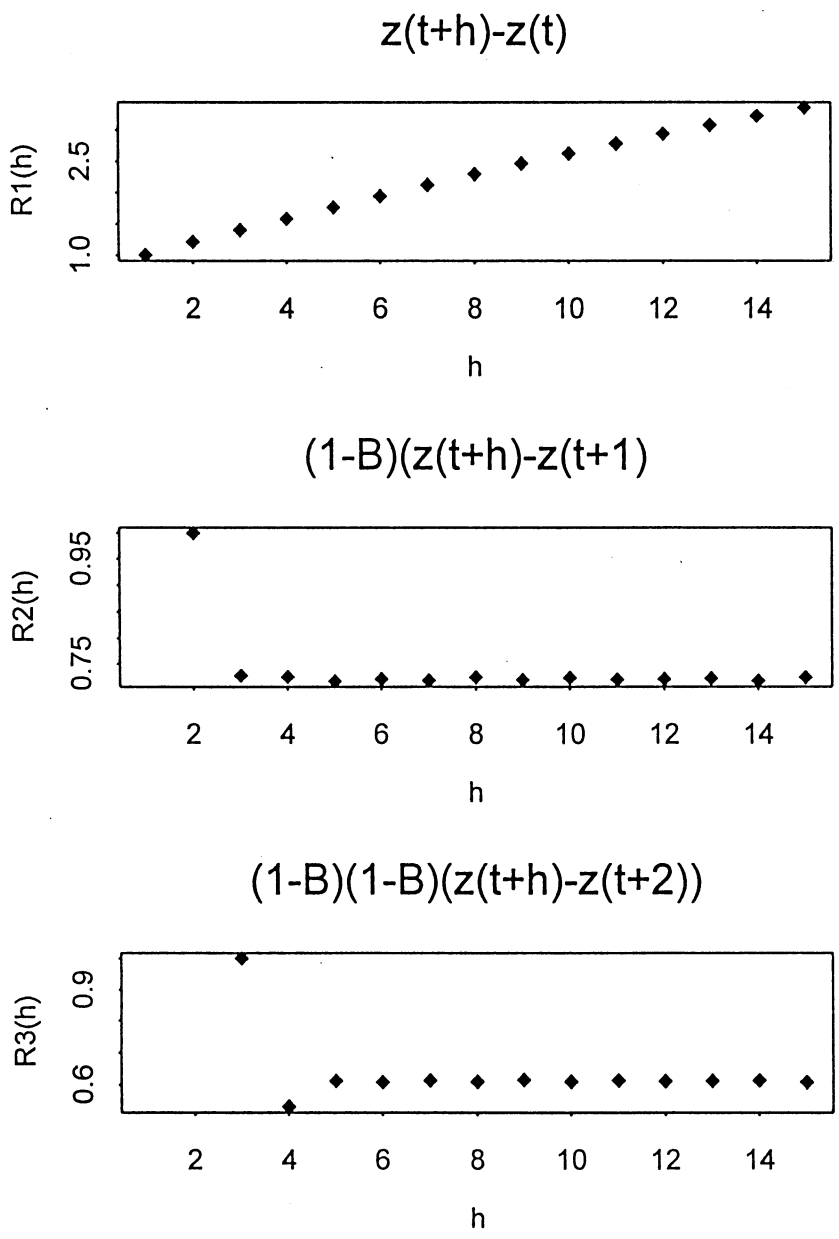


Figure 3.3: ARIMA(0,1,1), Theta=.5

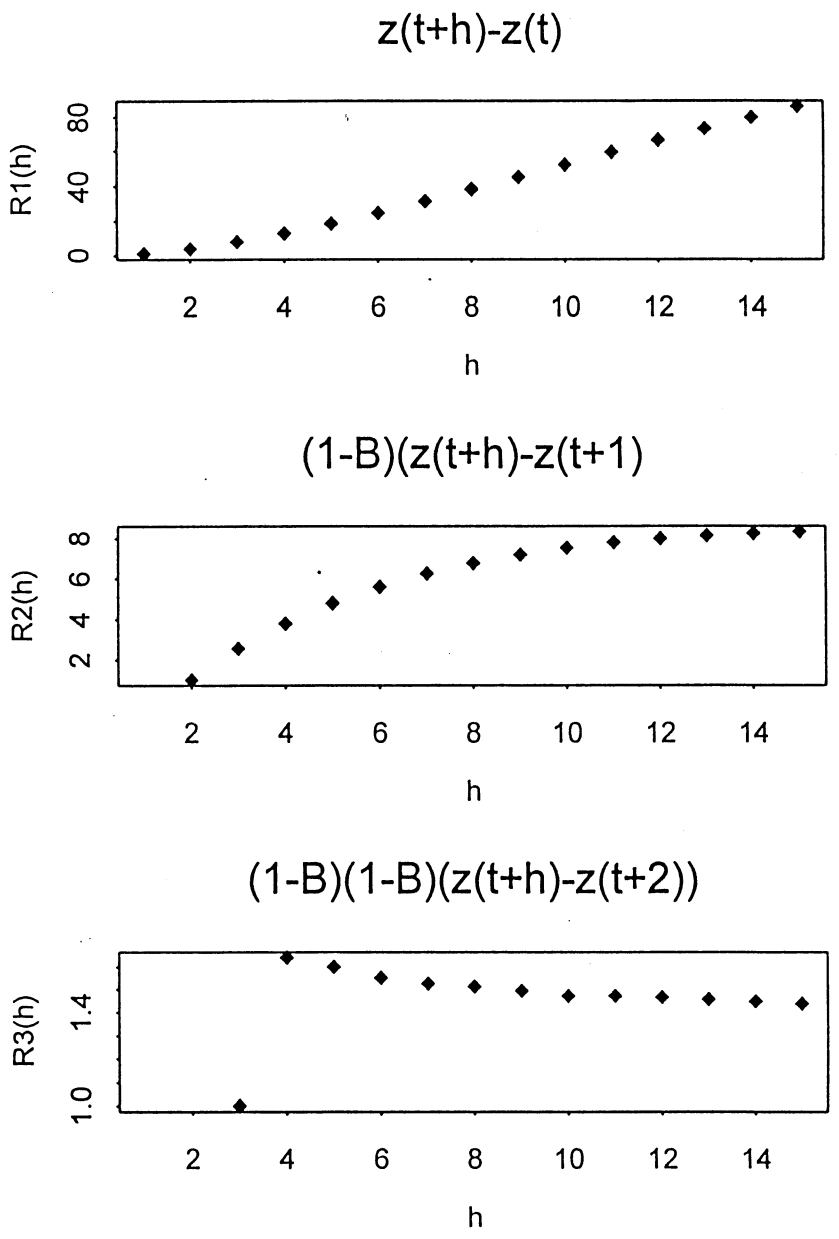


Figure 3.4: ARIMA(1,1,1), Phi=.8, Theta=-.5

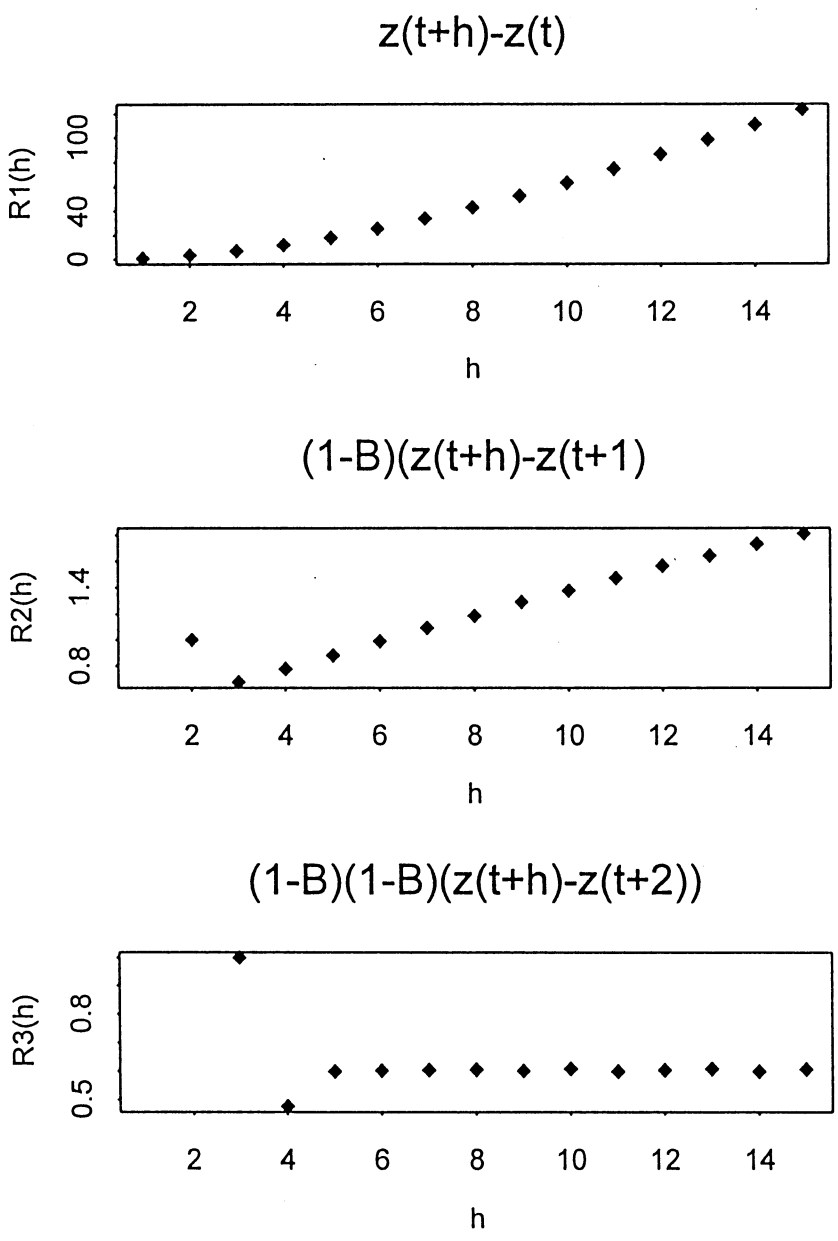


Figure 3.5: ARIMA(0,2,2), Theta=(1,-.5)

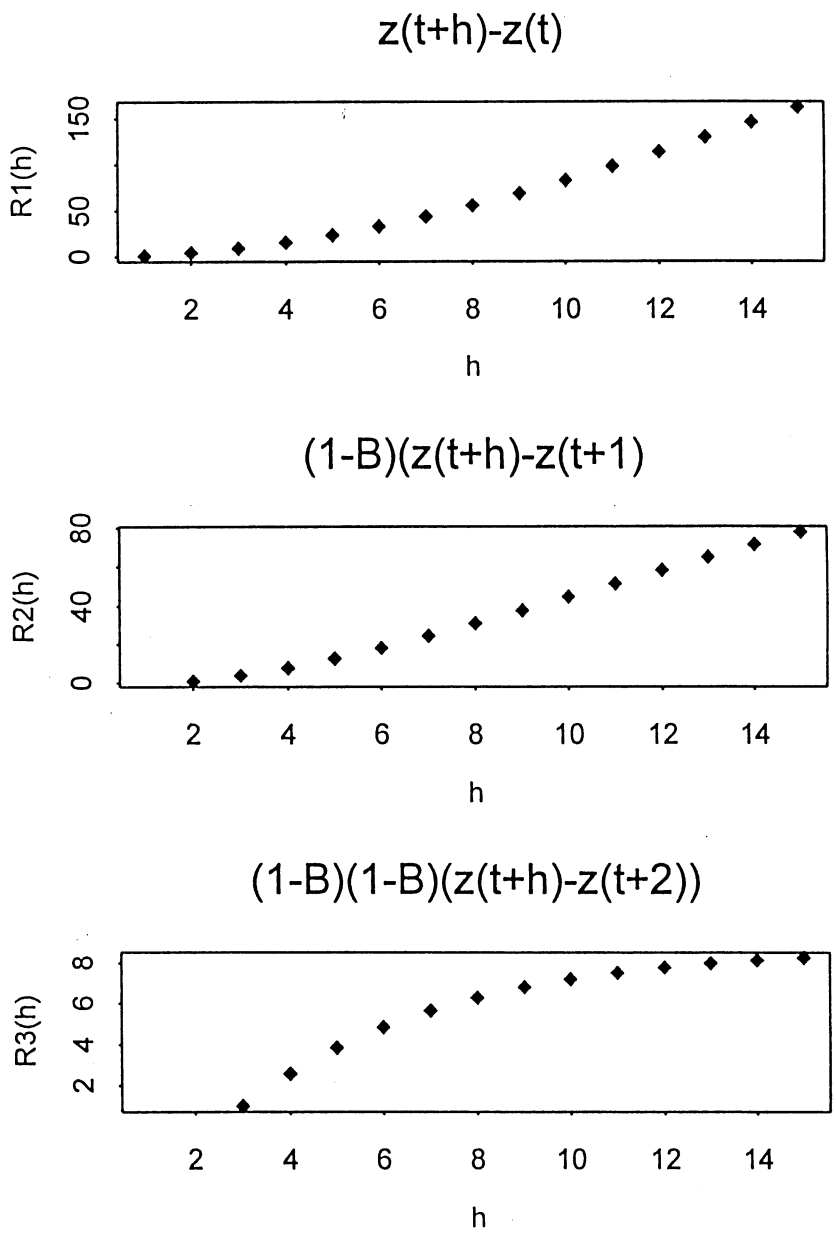


Figure 3.6: ARIMA(1,2,1), Phi=.8, Theta=-.5

4. Examples

In this section we consider the degree of differencing for some well known series. The following series are considered

(i) Resident U.S. Population (Cressie (1988))

(ii) Single-Family Housing Starts (Cressie (1988))

(iii) Series A (Box and Jenkins (1976))

(iv) Series B (Box and Jenkins (1976))

(v) Series C (Box and Jenkins (1976))

(i) Resident U.S. Population

The current plots are shown in Figure 4.1. The first plot shows a nonlinear increase while the second plot is more linear. The third plot indicates that $R_3(h)$ varies around a constant. Hence $d = 2$. Cressie's (1988) procedure leads to $d = 2$.

(ii) Single Family Housing Starts (deseasonalized)

Figure 4.2 indicates that $R_1(h)$ is linear in h and $R_2(h)$ is constant approximately. Hence $d = 1$. This is the same conclusion in Dickey et al. (1986) using a unit root test. Cressie's (1988) procedure also leads to $d = 1$.

(iii) Series A (Box and Jenkins (1976))

This is a set of chemical process concentration readings and Box and Jenkins consider the degree of differencing to be either 0 or 1. Figure 4.3 gives the plots and this is consistent with the pattern of a process with an AR component and $d = 1$.

(iv) Series B

This is the IBM Common Stock closing prices (daily, May 17, 1961 - Nov. 2, 1962). The plots are given in Figure 4.4. We conclude that $d = 1$ since $R_1(h)$ is linear in h and $R_2(h)$ and $R_3(h)$ are varying around constants. Box and Jenkins (1976) concluded the same.

(v) Series C

This is a series of chemical process temperature readings every minute. The plots are given in Figure 4.5. This is consistent with the pattern of an ARIMA$(1,1,0)$ process and hence we take $d = 1$. This is the same conclusion in Box and Jenkins (1976) as well.

5. Concluding Remarks

In this paper we introduced some graphical tools based on modifications to the variogram to specify the degree of differencing,d, in a time series. The behavior of these tools were studied using a modest simulation study. We also used the tools to specify d in some well known series and the results agreed with those from other methods. We feel that these tools are very simple and can be adapted easily in any software dealing with time series.

Acknowledgements

B. Abraham was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

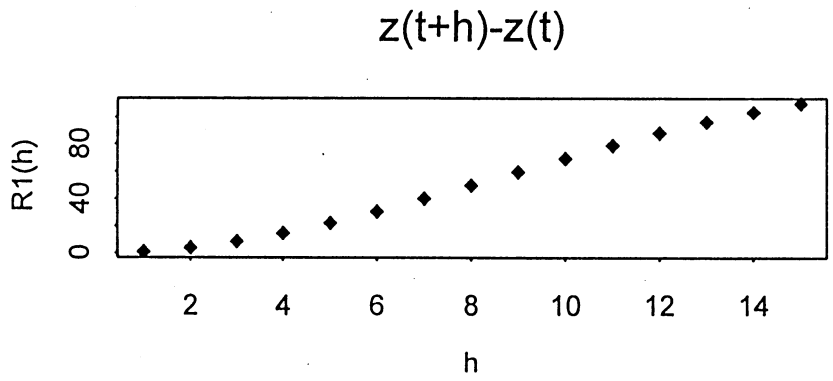
Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis, Forecasting and Control*, San Francisco, Holden-Day.

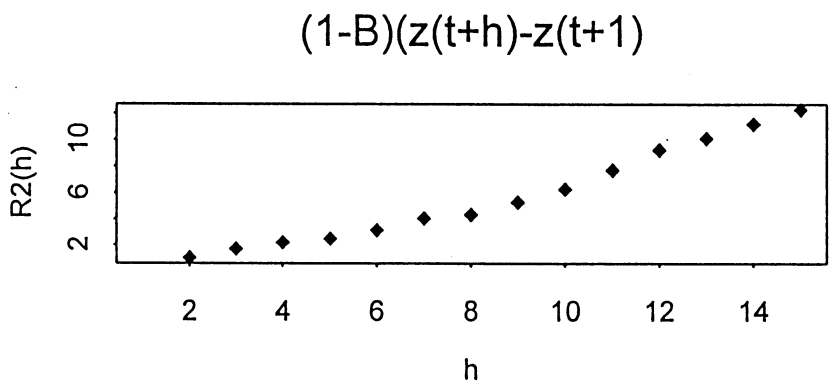
Box, G.E.P. and Lucerno, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*, New York, John Wiley and Sons Inc.

Box, G.E.P. and Kramer, T. (1992). Statistical Process Monitoring and Feedback Adjustment. A discussion. *Technometrics, 34*, 251-285.

Cressie, N. (1988). A Graphical Procedure for Determining Nonstationarity in Time Series, *Jour. of the American Statistical Association, 83(404)*, 1108-1116.

Dickey, D.A., Bell, W.R., and Miller, R.B. (1986). Unit Roots in Time Series Models: Tests and Implications, *The American Statistician, 40*, 12-26.





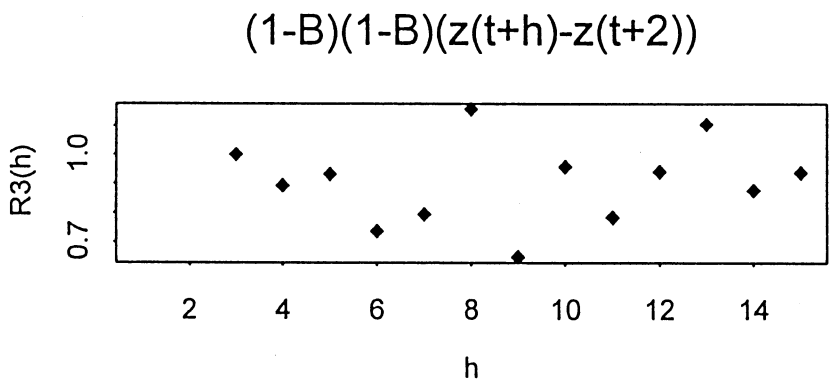
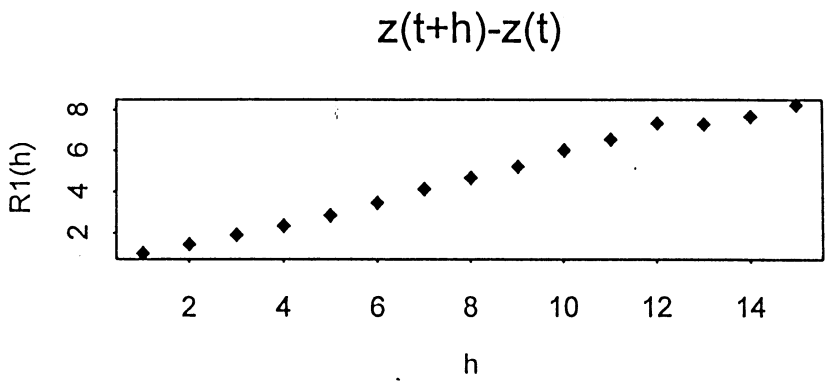
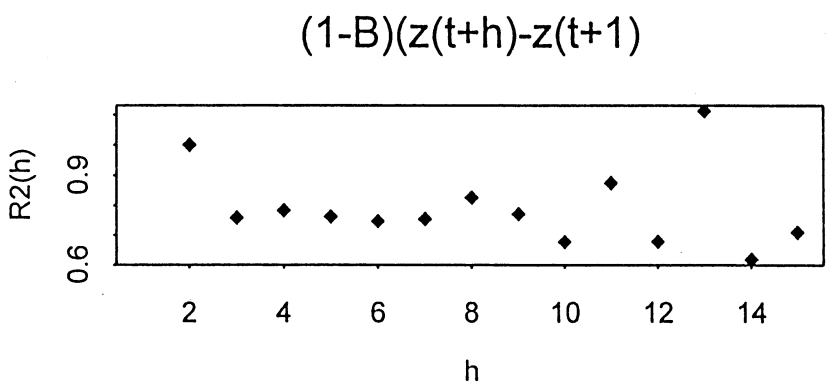


Figure 4.1: U.S. Population





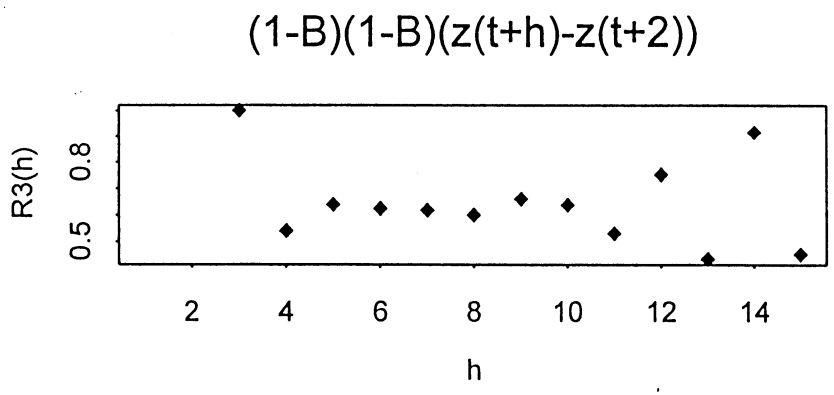


Figure 4.2: Housing Starts

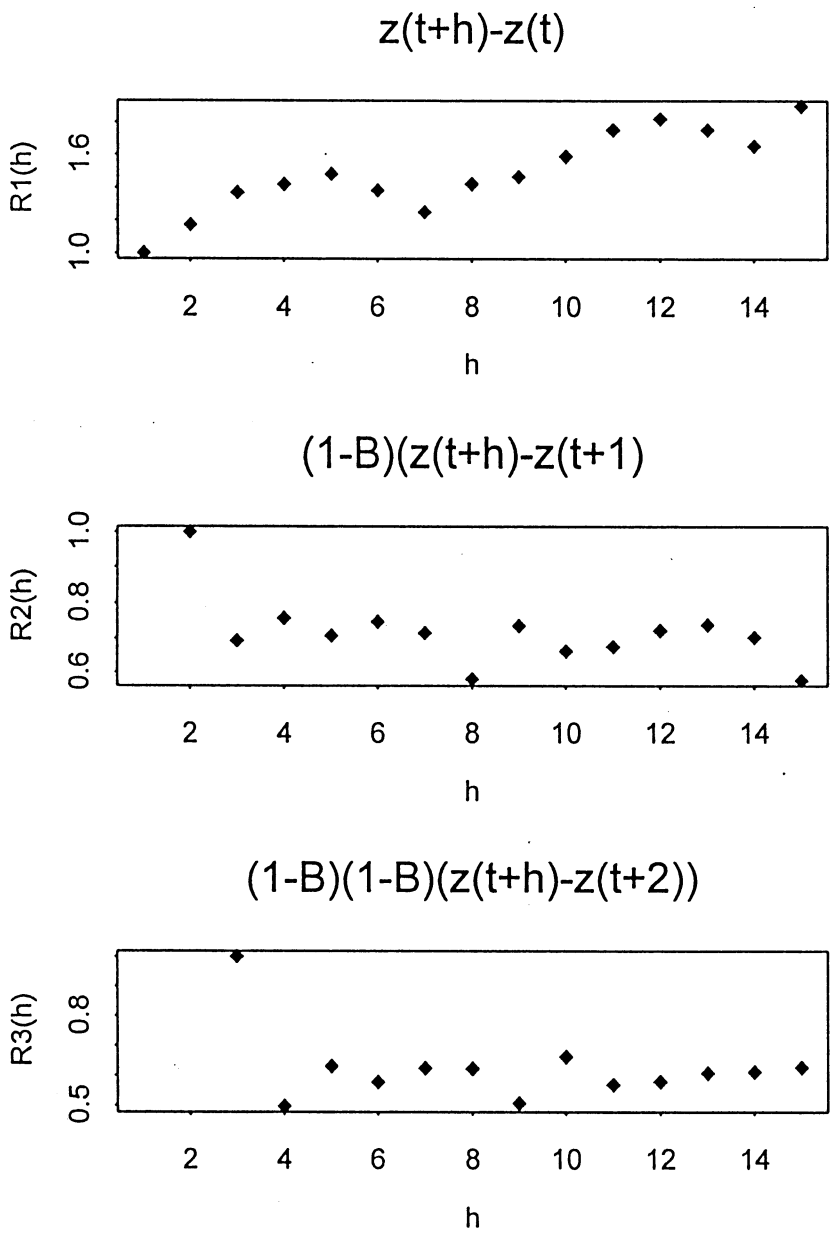


Figure 4.3: Series A

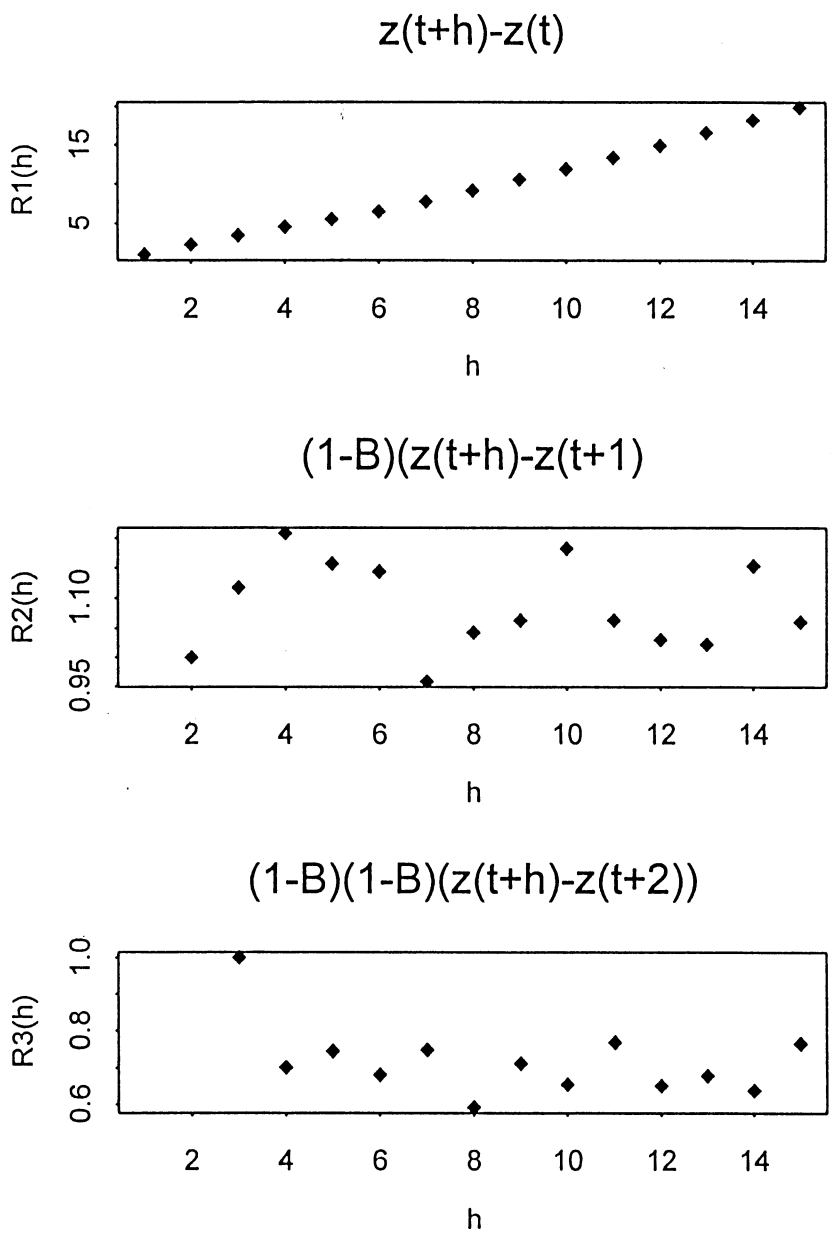


Figure 4.4: Series B

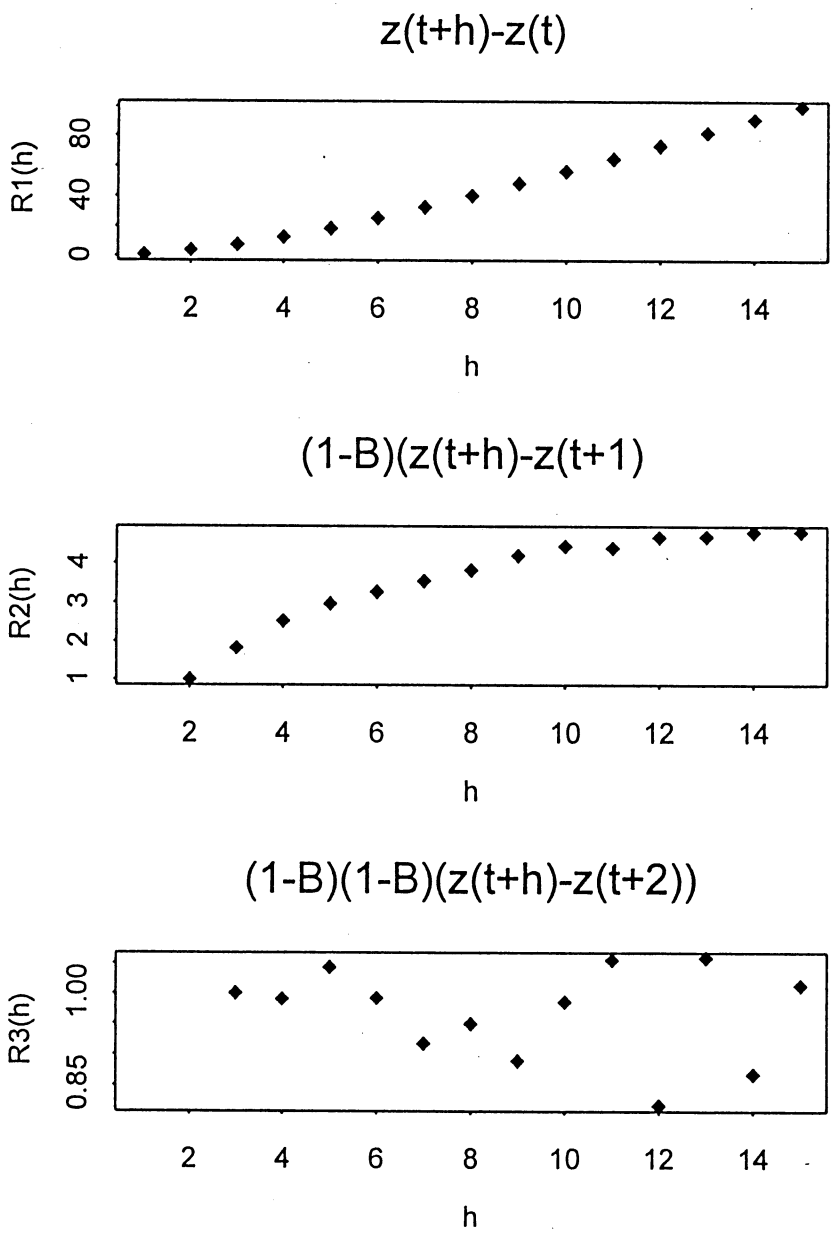


Figure 4.5: Series C

APPENDIX

Table A.1 $V_i(h)$ and $E(S_i^2(h))$

	ARIMA(1, 0, 1) $(1-\phi B)z_t = (1-\theta B)a_t$	ARIMA(1, 1, 1) $(1-\phi B)(1-B)z_t = (1-\theta B)a_t$	ARIMA(0, 2, 2) $(1-B)^2 = (1-\theta_1 B - \theta_2 B^2)a_t$
$V_1(h)$	$\gamma_{(1,0,1)}(u_{1t}(h))$ see equation (A.1)	$\gamma_{(1,1,1)}(u_{1t}(h))$ see equation (A.3)	Does not exist
$E(S_1^2(h))$	$\gamma_{(1,0,1)}(u_{1t}(h)) - \frac{\delta_{(1,0,1)}(u_{1t}(h)}{(n-h)^2}$ see equation (A.2)	$\gamma_{(1,1,1)}(u_{1t}(h)) - \frac{\delta_{(1,1,1)}(u_{1t}(h))}{(n-h)^2}$ see equation (A.5)	see equation (A.7)
$V_2(h)$		$\gamma_{(1,1,1)}(u_{2t}(h))$ see equation (A.4)	$\gamma_{(0,2,2)}(u_{2t}(h))$ see equation (A.8)
$E(S_2^2(h))$		$\gamma_{(1,1,1)}(u_{2t}(h)) - \frac{\delta_{(1,1,1)}(u_{2t}(h))}{(n-h)^2}$ see equation (A.6)	$\gamma_{(0,2,2)}(u_{2t}(h)) - \frac{\delta_{(0.2,2)}(u_{2t}(h))}{(n-h)^2}$ see equation (A.9)
$V_3(h)$			$\gamma_{(0,2,2)}(u_{3t}(h))$, see equation (A.10)
$E(S_3^2(h))$			$\gamma_{(0,2,2)}(u_{3t}(h)) - \frac{\delta_{(0,2,2)}(u_{3t}(h))}{(n-h)^2}$ see equation (A.11)

Variances and Expected Values of Sample Variances

ARIMA(0,01): $z_t = (1-\theta B)a_t$; $\gamma_{(0,0,1)}(u_{1t}(h)) = V_1(h) = 2\sigma^2(1+\theta^2)$

ARIMA(1,0,0): $(1-B)z_t = a_t$; $\gamma_{(1,0,0)}(u_{1t}(h)) = V_1(h) = 2\sigma^2(1-\phi^h)/(1-\phi^2)$

ARIMA(1,0,1): $(1-\phi B)z_t = (1-\theta B)a_t$

$$\gamma_{(1,0,1)}(u_{1t}(h)) = V_1(h) = \frac{2\sigma^2}{1-\phi^2}[1-\phi^2+(\phi-\theta)^2+(\phi-\theta)(\phi\theta-1)\phi^{h-1}] \qquad (A.1)$$

$$\begin{aligned}\delta_{(1,0,1)}(u_{1t}(h)) &= \sigma^2[(n-h)(\frac{1-\theta}{1-\phi})^2 - \frac{2\phi(\phi-\theta)(1-\theta)(1-\phi^{n-h})}{(1-\phi)^2} \\ &\quad + \frac{(\phi-\theta)^2}{(1-\phi)^2(1-\phi^2)}\{\phi^2(1-\phi^{2(n-h)}) + (1-\phi^{n-h})^2\}]. \qquad (A.2)\end{aligned}$$

By setting $\phi = 0$, and $\theta = 0$ we can obtain the results for MA(1) and AR(1) processes respectively.

ARIMA$(0,1,1)$: $(1-B)z_t = (1-\theta B)a_t$

$\gamma_{(0,1,1)}(u_{1t}(h)) = V_1(h) = [(h-1)(1-\theta)^2 + (1+\theta^2)]\sigma^2$, $\gamma_{(0,1,1)}(u_{2t}(h)) = V_2(h) = 2\sigma^2(1+\theta^2)$

ARIMA$(1,1,0)$: $(1-\phi B)(1-B)z_t = a_t$

$\gamma_{(1,1,0)}(u_{1t}(h)) = V_1(h) = \sigma^2[\frac{h}{(1-\phi)^2} - \frac{2\phi(1-\phi^h)}{(1-\phi)^2(1-\phi^2)}]$, $\gamma_{(1,1,0)}(u_{2t}(h)) = V_2(h) = 2\sigma^2(1-\phi^{h-1})/(1-\phi^2)$

ARIMA$(1,1,1)$: $(1-\phi B)(1-B)z_t = (1-\theta B)a_t$

$$\gamma_{(1,1,1)}(u_{1t}(h)) = V_1(h) = \sigma^2[(\frac{1-\theta}{1-\phi})^2 h - \frac{2(\phi-\theta)(1-\phi\theta)}{(1-\phi)^2(1-\phi^2)}(1-\phi^h)], \quad h = 1,2,... \quad \text{(A.3)}$$

$$\gamma_{(1,1,1)}(u_{2t}(h)) = V_2(h) = 2\sigma^2[1-(\phi-\theta)\phi^{h-2} + \frac{(\phi-\theta)^2}{1-\phi^2}(1-\phi^{h-1})], \quad h = 2,3,... \quad \text{(A.4)}$$

$$\begin{aligned}\delta_{(1,1,1)}(u_{1t}(h)) &= \sigma^2[\sum_{j=1}^{h}\{1+(j-1)(\frac{1-\theta}{1-\phi}) - \frac{\phi(\phi-\theta)}{(1-\phi)}(\frac{1-\phi^{j-1}}{1-\phi})\}^2 + \{1+(h-1)\frac{1-\theta}{1-\phi} \\ &\quad -\frac{\phi(\phi-\theta)}{1-\phi}\frac{(1-\phi^{h-1})}{1-\phi} + \frac{(\phi-\theta)(1-\phi^h)}{(1-\phi)^2}(n-2h-\frac{\phi(1-\phi^{n-2h})}{1-\phi}\}^2 \\ &\quad +\sum_{j=1}^{h-1}\{(h-j)(\frac{1-\theta}{1-\phi}) - \phi^j\frac{(\phi-\theta)}{1-\phi}(\frac{1-\phi^{h-j}}{1-\phi}) + \frac{(\phi-\theta)(1-\phi^h)}{1-\phi}(\frac{1-\phi^{n-2h+j}}{1-\phi})\}^2 \\ &\quad +\frac{(\phi-\theta)^2(1-\phi^h)^2}{(1-\phi)^2(1-\phi)^2}(1-\phi^{n-h})^2\frac{1}{1-\phi^2}] \end{aligned} \quad \text{(A.5)}$$

$$\begin{aligned}\delta_{(1,1,1)}(u_{2t}(h)) &= \sigma^2[1+\sum_{j=1}^{h-2}\{1+(\phi-\theta)(\frac{1-\phi^j}{1-\phi})\}^2 + \sum_{j=1}^{h-2}\{(\frac{\phi-\theta}{1-\phi})(1-\phi^j+\phi^{n-h+j}-\phi^{n-2h+j+1})+1\}^2 \\ &\quad +\frac{(\phi-\theta)^2}{1-\phi^2}(\frac{1-\phi^{h-1}}{1-\phi})^2\{2(1-\phi^{n-h})+\phi^{2n-2h}(1-\phi^{4-2h})\}] \text{ if } h>2. \end{aligned} \quad \text{(A.6)}$$

By setting $\phi = 0$, and $\theta = 0$ we can obtain the expressions for the processes $(0,1,1)$ and $(1,1,0)$ respectively.

ARIMA(0,2,2): $(1-B)^2 z_t = (1-\theta_1 B - \theta_2 B^2)a_t$

$$\begin{aligned}E(S_1^2(h)) &= \frac{\sigma^2}{(n-h)^3}[\{\theta_2^2+(\theta_1+3\theta_2)^2\}(n-h) + (\frac{n-h}{4})\sum_{i=3}^{h}\{(1-\theta_1-\theta_2)i^2+(\theta_1-\theta_2-3)i+2\}^2 \\ &\quad +(\frac{n-h}{4})\{h(h-1)-\theta_1 h(h+1)-\theta_2 h(h+3)\}^2 \\ &\quad +\sum_{t=4}^{n-h}\{h^2\sum_{i=h+1}^{t-2}[i-\frac{h+1}{2}-\theta_1(i+1-\frac{h+1}{2}) \\ &\quad -\theta_2(i+2-\frac{h+1}{2})]^2 + (n-h)[h(t-1-\frac{h+1}{2}) - \theta_1 h(t-\frac{h+1}{2}) - \theta_2(h-1)(t-\frac{h}{2})]^2 \\ &\quad +(n-h)[h(t-\frac{h+1}{2}) - \theta_1(h-1)(t-\frac{h}{2}) - \theta_2(h-2)(t-\frac{h-1}{2})]^2\} \end{aligned}$$

$$+\sum_{t=2}^{n-h}\{\sum_{i=1}^{h-3}[(h-i)(t-\frac{h-i+1}{2})-\theta_1(h-i-1)(t-\frac{h-i}{2}) \tag{A.7}$$
$$-\theta_2(h-i-2)(t-\frac{h-i-1}{2})]^2\}$$
$$+\frac{(5-8\theta_1+\theta_1^2)}{6}(n-h-1)(n-h)(2n-2h-1)]$$

$$\begin{aligned}\gamma_{(0,2,2)}(u_{2t}(h)) &= \mathrm{Var}(u_{2t}(h)) = V_2(h)\\ &= \{1+(1-\theta_1)^2+(\theta_1+\theta_2)^2+\theta_2^2+(1-\theta_1-\theta_2)^2(h-3)\}\sigma^2,\quad h\geq 3. \qquad \text{(A.8)}\end{aligned}$$

$$\begin{aligned}\delta_{(0,2,2)}(u_{2t}(h)) &= \frac{\sigma^2}{(n-h)^2}[1+\theta_2^2+\sum_{j=1}^{h-2}\{j\theta_1+(j+1)\theta_2-(j-1)\}^2+(h-1)(\theta_1+\theta_2)-(h-2)\}^2\\ &\quad +(n-2h)(h-1)^2(\theta_1+\theta_2-1)+\{(h-1)\theta_1+(h-2)\theta_2-(h-1)\}^2\\ &\quad +\sum_{j=2}^{h-1}\{(j-1)\theta_1+(j-2)\theta_2-j\}^2] \quad h>2 \qquad \text{(A.9)}\end{aligned}$$

$$\gamma_{(0,2,2)}(u_{3t}(h)) = \mathrm{Var}(u_{3t}(h)) = V_3(h) = 2\sigma^2(1+\theta_1^2+\theta_2^2) \tag{A.10}$$

$$\delta_{(0,2,2)}(u_{3t}(h)) = \frac{2\sigma^2}{(n-h)^2}\{1+(1-\theta_1)^2+\theta_2^2+(\theta_1+\theta_2)^2+(h-4)(1-\theta_1-\theta_2)^2\} \quad h>3 \tag{A.11}$$