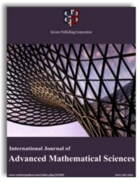
International Journal of Advanced Mathematical Sciences, 7 (1) (2019) 6-15

International Journal of Advanced Mathematical Sciences

Website: www.sciencepubco.com/index.php/IJAMS

Research paper



Penalty of model misspecification in time series dominated with trend

Ogbonna, Chukwudi Justin[1], Nweke, Chijioke Joel [2]*, Ojide, Kelechi Charity[2]

[1]*Department of Statistics, Federal University of Technology, Owerri, Nigeria*
[2]*Department of Mathematics and Statistics, Alex Ekwueme Federal University, Ndufu-Alike, Ikwo, Nigeria*
**Corresponding author E-mail:cj_nweke@yahoo.com*

Abstract

Model specification is consequential in mathematical science and statistics in particular. This work seeks to ascertain the consequences of model mis-specification in the analysis of a time series dominated by trend. It further discusses the statistical properties of various types of trend as well as when they are combine with AR (1) and MA (1) process. It recommends the use of spectrum analysis in detection of trend type in a given series.Illustrations were carried out using simulated series. The results from the simulated series was in harmony with the theoretical results.

Keywords: *Deterministic; Mis-Specification; Spectrum; Stochastic; Trend.*

1. Introduction

History, when ignored could be devastating because of its possibility of repeat in the future. Therefore, many resourceful policies and decisions are largely based on the available information of the past observations and possibly the process that generate such observations. To develop an understanding of such phenomena, there are two possible approaches. The first is to consider the fundamental processes that are believed to be operating and to build a more or less detailed model of these processes that can be used to make predictions and explore alternative scenarios .The second approach is to analyse the available data, either to look for relationships that could explain how the system works or to test hypotheses suggested by the process based considerations (Saunders, 1999; Solomon et al., 2007; Chandler and Scott, 2011). Such information could be accessed through observation. When such observations are made sequentially at regular or approximately regular interval of time, it is called time series data (Box et al, 1994; Chatfield, 2004; Wei, 1990). Time series is used to represent the characterized time course of behaviour of wide range of several systems which could be biological, physical or economical (Ademola, 2007). To utilize the aforementioned observations, they are subjected to analysis using time series techniques. This would help to achieve the aim of the observation which could be for description, explanation forecast and control (Chatfield, 2004; Ogbonna et al, 2016).

Often, most real life data are characterized with trend, hence such data requires a proper trend analysis for adequate modelling. Trend is a long-term temporal variation in the statistical properties of a process. It is a long-term change in the mean level (Chatfield, 2004; Kendall and Ord, 1990; Chandler and Scott, 2011). Robinson, (2003) has however tried to distinguish between trend and fluctuation. A series is said to show trend if on average, the series is progressively increasing or decreasing, but is said to show fluctuation if on average, the series changes noticeably through time but not in any consistent direction. In any case, Chardler and Scott (2011) have noted among others, the reasons trend analysis could be useful, namely:(a) to describe the past behaviour of a process. (b) to try and understand the mechanisms behind observed changes. (c) to make assessments of possible future scenarios by forecast. (d) to enable the analysis of systems where long-term changes serve to obscure the aspects of real interest. (e) to set up an effective control mechanism.

Different tests for assessing the presence of trend as well as nature of trend in the series have been advocated for in the literature. Such methods range from graphical to functional methods. However to ascertain the nature of the trend in some cases usually poses a serious challenge. Therefore, the main aim of this research is to ascertain the consequences of model mis-specification of time series dominated with trend and as well as recommending an appropriate way to detect various trend types. This would be achieved via studying the various common trend types that could characterize a series, their properties as well as the differences.

2. Literature review

So many time series data that are non-stationary are characterized by mean and variance changes, seasonality and other local behaviours like outliers and discontinuities (Ademola, 2007; Wei, 1990). Such series may not only be non-stationary in mean and variance but could incorporate distortion due to both known and unknown causes (Granger, 1994). The presence of these characteristics in time series has led to considerable research debate on the desirability of data pre-processing (removal of deterministic component, transformation and sometimes detection and removal of outliers) to model such non-stationary series (Nelson et al., 1999; Zhang et al., 2001; Zhang and Qi, 2005). Wei (1990) had noted three broad forms of non-stationarity, namely: (i) Non-stationarity in mean (ii) Non-stationarity in variance and (iii) Non-stationarity in mean and variance. Nelson and Plosser (1982) were among the first to point out non-stationarity and its economic implications, hence, they advocated for a unit root test. Dickey and Fuller (1979) developed a simplest and most common unit root test. Dickey and Fuller test developed in 1979 has some deficiencies which were remedied in 1984 by Said and Dickey (Said and Dickey, 1984) and the new test is called Augmented Dickey-Fuller test (ADF). Phillips and Perron (1988) developed a non-parametric statistical method to take care of series correlation in the error term without adding lagged difference terms. The test statistic follows exactly the same asymptotic distributions with Augmented Dickey-Fuller test statistic (ADF). Stefan et al. (2011) applied unit root and stationarity testing in the analysis of industrial production of Central and Eastern Europe countries. Kwiatkowski et al. (1992) proposed a test of null hypothesis that an observable series is stationary around a deterministic trend. The test used Lagrange-Multiplier (LM) statistic and was applied on Nelson-Plosser data and for many of these series the hypothesis of trend stationarity were not rejected. Ruey (1988) discussed analysis of time series with outliers, level shifts and variance changes. He used the least squares technique and residual variance ratio method. His method was found to be very effective in the modelling of a time series that is non-stationary. To specify correctly a trended series, the nature of the trend must first be noticed.

The two types of trend commonly reported in the literature are deterministic trend and stochastic trend. A process with deterministic trend has a shock with transitory effect while those with stochastic trend have a shock with permanent effect (Heino, 2005). A process with deterministic trend is sometimes referred to as trend stationary process while that of stochastic trend is called differenced trend or unit root process.

One of the assumptions in regression and time series analysis is correct specification of the model, i.e., there is no specification bias in the model. When such assumption is violated, grave consequences could be incurred such as multicollinearity, model under fitting and over fitting (Gujariti, 2004). Heino (2005) has noted the test for stationarity of a series using both Dickey-Fuller test and Kwaiatkowski-Phillips-Schmidt-Shin test (KPSS).

3. Methodology

In time series, non-stationarity could either be in mean or variance or in both. A non-stationarity in mean could be as a result of trend. When such trend exists in a time series, two (possible) approaches to specify it are

$$X_t = T_t + \psi(B)e_t \quad (1)$$

$$(1-B)^d X_t = \alpha + \psi(B)e_t \quad (2)$$

Where, X_t

X_t = original value of the series at time t , B = backward shift operator such that $(1-B)^d = X_t - X_{t-d}$ $\forall\ d \geq 1$, d is any positive integer, T_t = The trend function, e_t = any stationary process not necessarily white noise, $\psi(B) \neq 0$

3.1. Case I

We start here by assuming that the trend function, T_t is linear, $\psi(B)=1$, and $d=1$ and the error term, $e_t \sim N(0,\sigma^2)$. Following these assumptions, Equations 1 and 2 are re-written as;

$$X_t = \beta_0 + \beta_2 t + e_t \quad (3)$$

$$(1-B)X_t = \alpha + e_t \quad (4)$$

The dependent variable X_t , in Equation 3, changes at a constant rate over time. Equation 3 has a deterministic trend while Equation 4 has a stochastic trend. The summary of the properties of Equations 3 and 4 are shown in Table 1.

3.1.1. Specification and transformation for stationarity

If the process in Equation 3 is generated with trend stationarity, then it could be expressed symbolically as:

$$X_t = \beta_0 + \beta_1 t + e_t \quad (6)$$

And

$$X_{t-1} = \beta_0 + \beta_1 (t-1) + e_{t-1} \quad (7)$$

Fitting a deterministic trend to Equation 3 by detrending yields Equation 8:

$$X_t - \hat{T}_t = e_t \quad (8)$$

Fitting a stochastic model to Equation 3 by differencing yields Equation 9:

$$\Delta X_t = -\beta_1 + e_t - e_{t-1} \quad (9)$$

If the process in Equation 4 is generated with difference stationary, then it could be expressed symbolically as:

$$\left(1-B\right)X_t=\alpha+e_t \Rightarrow X_t=\alpha+X_{t-1}+e_t \tag{10}$$

Fitting a deterministic trend to Equation 4 by detrending yields Equation 11:

$$X_t-T_t=-\beta_1 t+X_{t-1}+e_t \tag{11}$$

Fitting a stochastic model to Equation 4 by differencing yields Equation 12:

$$\Delta X_t=\alpha+e_t \tag{12}$$

3.2. Case II

Suppose that in Equations 1 and 2 the trend function, T_t is linear and the error term, e_t follows a zero mean ARMA process of order (p,q) ,then the equations could be written as:

$$\left(1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_p B^p\right)e_t=\left(1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q\right)u_t \tag{13}$$

Where, $u_t \sim N\left(0,\sigma^2\right)$, $1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_p B^p$ is the autoregressive part while $1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q$ is the moving average part. Supposing the moving average part is invertible, factorizing the autoregressive part gives

$$1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_p B^p=\left(1-\lambda_1 B\right)\left(1-\lambda_2 B\right)\ldots\left(1-\lambda_p B\right)$$

And solving for e_t gives

$$e_t=\frac{1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q}{\left(1-\lambda_1 B\right)\left(1-\lambda_2 B\right)\ldots\left(1-\lambda_p B\right)}u_t=\psi\left(B\right)u_t \tag{14}$$

With $\sum_{i=1}^{\infty}\left|\psi\right|<\infty$ and root of $\psi\left(z\right)=0$ lies outside the unit circle.
Suppose in Equation 13 we set $q=0, p=1$ so that it now becomes Equation 15:

$$\left(1-\varphi_1 B\right)e_t=u_t \Rightarrow e_t=\frac{u_t}{\left(1-\Phi B\right)}=\left(1-\Phi B\right)^{-1}u_t \tag{15}$$

$$=\left(1+\varphi_1 B+\varphi_2 B^2+\cdots\right)u_t=u_t+\varphi_1 u_{t-1}+\varphi_2 u_{t-2}+\cdots$$

Hence Equation 1, becomes

$$X_t=\beta_0+\beta_1 t+u_t+\varphi_1 u_{t-1}+\varphi_2 u_{t-2}+\cdots=\beta_0+\beta_1 t+u_t+\varphi_1\left(X_{t-1}-\beta_0-\beta_1\left(t-1\right)\right)$$

$$=\alpha^*+\beta^* t+\varphi_1 X_{t-1}+u_t \tag{16}$$

Where, $\alpha^*=\left(1-\varphi_1\right)\beta_0+\varphi_1\beta_1$, $\beta^*=\left(1-\varphi_1\right)\beta_1$
If $\left|\varphi_1\right|<1$, Equation 16 isAR(1) process around a deterministic trend.
Secondly, suppose in Equation 14, $\lambda_1=1, \lambda_i>1, i=2,3,\ldots,p$ and $q=0$, then it is rewritten as Equation 17:

$$e_t=\frac{u_t}{\left(1-B\right)\left(1-\lambda_2 B\right)\left(1-\lambda_3 B\right)\ldots\left(1-\lambda_p B\right)} \tag{17}$$

$$\left(1-B\right)e_t=\frac{u_t}{\left(1-\lambda_2 B\right)\left(1-\lambda_3 B\right)\ldots\left(1-\lambda_p B\right)}=\psi^*\left(B\right)u_t$$

Here again if in Equation 2, $d=1,$ and $p=2$ in Equation 13, then Equation 2 becomes

$$\left(1-B\right)X_t=\alpha+\varphi_1 X_{t-1}+u_t \tag{18}$$

Equation 18 is a differenced trend around AR(1), where α is the mean level and $u_t \sim N\left(0,\sigma^2\right)$ .Therefore,

$$X_t=\alpha+X_{t-1}+\varphi_1 X_{t-1}+u_t=\alpha+\left(1+\varphi_1\right)X_{t-1}+u_t \tag{19}$$

The summary of the properties of the AR(1) process around a deterministic trend in Equation 17 and a diferenced process in Equation 19 are shown in Table 2.

3.2.1. Specification and transformation for stationarity

Given the process $X_t = \alpha^* + \beta^* t + \varphi_1 X_{t-1} + u_t$
Fitting a deterministic trend by detrending gives

$$X_t - \hat{T}_t = Y_t = \varphi_1 Y_{t-1} + u_t$$

Where, Y_t is the detrended series and u_t is a white noise.
Fitting a stochastic model by difference gives

$$\Delta X_t = -\beta_1 + \varphi_1 (X_{t-1} - X_{t-2}) + u_t - u_{t-1}$$

Table 1: Summary of the Process Generated Using Deterministic and Stochastic Trend for CASE 1

Function	Deterministic Trend	Stochastic Trend
Model	$X_t = \beta_0 + \beta_2 t + e_t$	$(1-B)X_t = \alpha + e_t$
Other nomenclature	Trend stationary process	Difference stationary or unit root process.
Mean	$E(X_t) = E(\beta_0 + \beta_1 t + e_t) = \beta_0 + \beta_1 t$	$E(X_t) = E\left(t\alpha + X_0 + \sum_{i=1}^{t} e_i\right) = t\alpha + X_0 + \sum_{i=1}^{t} E(e_i) = t\alpha + X_0$
Mean at lag k	$E(X_{t-k}) = E(\beta_0 + \beta_1 (t-k) + e_{t-k}) = \beta_0 + \beta_1 (t-k)$	$E(X_{t-k}) = E\left((t-k)\alpha + X_0 + \sum_{i=1}^{t} e_{i-k}\right)(t-k)\alpha + X_0$
Variance	$Var(X_t) = E(\beta_0 + \beta_1 t + e_t - \beta_0 - \beta_1 t)^2 = Var(e_t) = \sigma^2$	$Var(X_t) = E(X_t - E(X_t))^2$ $= \sum_{i=1}^{t} E(e_i)^2 + 2\sum_i^t \sum_j^t E(e_i e_j) = \sum_{i=1}^{t} E(e_i)^2 = t\sigma^2$
Variance at lag k	$Var(X_{t-k}) = E(\beta_0 + \beta_1 (t-k) + e_{t-k} - \beta_0 - \beta_1 (t-k))^2$ $= Var(e_{t-k}) = \sigma^2$	$Var(X_{t-k}) = (t-k)\sigma^2$
Auto-covariance	$r_k = Cov(X_t, X_{t-k})$ $= E(e_t e_{t-k}) = \begin{cases} \sigma^2, & k = 0 \\ 0, & k \neq 0 \end{cases}$	$\gamma_k = \text{cov}(X_t, X_{t-k}) = E(X_t - E(X_t))(X_{t-k} - E(X_{t-k}))$ $= E\left(\sum_{i=1}^{t} e_i\right)\left(\sum_{i=1}^{t-k} e_t\right) = (t-k)\sigma^2$
Autocorrelation	$\rho_k = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t-k})}} = E(e_t e_{t-k}) = \begin{cases} \sigma^2, & k = 0 \\ 0, & k \neq 0 \end{cases}$	$\rho_k = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)Var(X_{t-k})}} = \frac{(t-k)\sigma^2}{\sqrt{(t\sigma^2)(t-k)\sigma^2}} = \sqrt{\frac{t-k}{t}}$
Requirement to achieve Stationarity	Detrending	Differencing
Misspecification penalty	$-\beta_1 - e_{t-1}$	$-\beta_1 t + X_{t-1}$

Given the process $(1-B)X_t = \alpha + \varphi_1 X_{t-1} + u_t \Rightarrow X_t = \alpha + X_{t-1} + \varphi_1 X_{t-1} + u_t$

Fitting a deterministic trend by detrending gives $X_t - \hat{T}_t = Y_t = -\beta_1 t + Y_{t-1} + \varphi_1 Y_{t-1} + u_t$ where Y_t is the detrended series and u_t is a white noise.

Fitting a stochastic model by difference gives $\Delta X_t = \alpha + \varphi_1 X_{t-1} + u_t$

3.3. Case III

Suppose that in Equations 1 and 2 the trend function, T_t, is linear and the error term, e_t, follows a zero mean ARMA process of order (p,q), then

$$(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p) e_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) u_t$$

Where $u_t \sim N(0, \sigma^2)$, $1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$ is the autoregressive part while $1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ is the moving average part. Hence

$$e_t = \left(\frac{1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q}{1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p} \right) u_t = \psi^*(B) u_t$$

Suppose $p = 0$, $q = 1$, $\psi^*(B) = (1 + \theta_1 B)$, then $\psi^*(B) u_t = (1 + \theta_1 B) u_t = u_t + \theta_1 u_{t-1}$. Following this, Equation 1 becomes

$$X_t = \beta_0 + \beta_1 t + \theta_1 u_{t-1} + u_t \qquad (21)$$

Equation 21 is a deterministic trend around a moving average process of order 1.
Similarly, Equation 2 becomes

$(1-B)X_t = \alpha + u_t + \theta_1 u_{t-1}$ such that

$$X_t = \alpha + X_{t-1} + u_t + \theta_1 u_{t-1} \quad (22)$$

Equation 22 is a differenced trend around a moving average process of order 1 (or a unit root process in economic time series).The properties of Equations 21 and 22 are displayed in Table 3.

3.3.1. Specification and transformation for stationarity

Let $X_t = \beta_0 + \beta_1 t + \theta_1 u_{t-1} + u_t$

Then fitting a deterministic trend to it by detrending gives

$X_t - T_t = \theta_1 u_{t-1} + u_t$,

While fitting a stochastic trend to it by differencing yields

$\Delta X_t = -\beta_1 + \theta_1(u_{t-1} - u_{t-2}) + u_t - u_{t-1}$.

Similarly, let

$X_t = \alpha + X_{t-1} + u_t + \theta_1 u_{t-1}$

Then fitting a deterministic trend to it by detrending, the following is obtained

$X_t - T_t = -\beta t + X_{t-1} + u_t + \theta_1 u_{t-1}$.

On the other hand, a fit of a stochastic trend to it by differencing gives

$\Delta X_t = \alpha + u_t + \theta_1 u_{t-1}$.

3.4. Remedy for misspecification

We will adopt spectral analysis approach to detect an appropriate trend in a series. The method will be to evaluate the estimate of the population spectrum at frequency zero of the first differenced series.
Let $X_t, t \in T$ be an observed time series at time t, let $\Delta X_t = X_t - X_{t-1}$ be the first difference of the series. If $\Delta X_t = X_t - X_{t-1}$ is covariance stationary with constant mean and summable auto-covariance given by r_k , then the auto-covariance generating function is given by:

$$g_{\Delta X_t}(z) = \sum_{k=-\infty}^{\infty} r_k z^k = \psi(z)\sigma^2\psi(z^{-1}) \quad (26)$$

Where z denotes a complex scalar which could be represented as $e^{-i\omega}$ for $i = \sqrt{-1}$. Therefore, the population spectral is given by

$$S_{\Delta X_t}(\omega) = \frac{1}{2\pi} g_{\Delta X_t}(z) = \frac{1}{2\pi} g_{\Delta X_t}(e^{-i\omega}) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r_k e^{-\omega k} \quad (27)$$

Where ω is a real scalar.
Putting z =1 in Equation 26, we have

$$g_{\Delta X_t}(1) = \sum_{k=-\infty}^{\infty} r_k = [\psi(1)]^2 \sigma^2 \quad (28)$$

Therefore, the population spectrum in Equation 27 evaluated at z =1 and frequency zero becomes:

$$S_{\Delta X_t}(0) = \frac{1}{2\pi} g_{\Delta X_t}(1) = \frac{1}{2\pi}[\psi(1)]^2 \sigma^2 \quad (29)$$

Hence when a series is difference stationary, the population spectrum of the first difference at frequency zero is positive.
On the contrary, if the process is trend stationary, Hamiton (1994) has shown the autocovariance-generating function is given by

$$g_{\Delta X_t}(z) = (1-z)\psi(z)\sigma^2\psi(z^{-1})(1-z^{-1}) \quad (30)$$

Therefore, at z = 1 Equation 30 becomes 0, hence the population spectrum of the first difference of the trend stationary series at zero frequency is zero.

4. Illustration using simulated data

4.1. Simulated series with deterministic trend (X1t) and stochastic trend (X2t)

Fitting a deterministic trend to X_{1t} produced a significant trend line and appropriate model with uncorrelated residuals (i.e white noise) while fitting a stochastic trend on X_{1t} introduces amoving average process to the residual. This is evident in the ACF and PACF of the residual as shown in table 4 and is in harmony with the theoretical result in table 1. To avoid such specification error, the nature of the trend must first be detected.

Fitting a deterministic trend to X_{2t} resulted to a significant trend line but introduces AR(1) around a deterministic trend to the residual. On the contrary, fitting a stochastic trend to X_{2t} by difference produced an uncorrected residual (white noise).This again is in agreement with the theoretical result at Table 1. The ACF and PACF for the residual when deterministic and stochastic trend are fitted to X_{2t} are displayed in Table 5.

Table 4: The ACF and PACF of the Residual when Deterministic and Stochastic Trend Is Fitted to X_{1t}

lag	When deterministic trend is fitted				When stochastic trend is fitted			
	ACF	T	PACF	T	ACF	T	PACF	T
1	-0.0580	-1.2979	-0.0580	-1.2979	-0.5297	-11.8336	-0.5297	-11.8336
2	0.0066	0.1466	0.0032	0.0720	0.0056	0.1009	-0.3823	-8.5391
3	0.0601	1.3404	0.0609	1.3622	0.0688	1.2296	-0.2110	-4.7129
4	-0.0302	-0.6703	-0.0234	-0.5226	-0.0961	-1.7135	-0.2538	-5.6697
5	0.0744	1.6509	0.0711	1.5894	0.0897	1.5888	-0.1447	-3.2319

Table 5: The ACF and PACF of the Residual when Deterministic and Stochastic Trend Is Fitted to X_{1t}

Lag	When deterministic trend is fitted				When stochastic trend is fitted			
	ACF	T	PACF	T	ACF	T	PACF	T
1	0.9775	21.8571	0.9775	21.8571	-0.0583	-1.3014	-0.0583	-1.3014
2	0.9580	12.5560	0.0577	1.2910	0.0073	0.1635	0.0040	0.0886
3	0.9378	9.6252	-0.0229	-0.5124	0.0623	1.3858	0.0631	1.4101
4	0.9145	8.0175	-0.0815	-1.8227	-0.0310	-0.6863	-0.0239	-0.5333
5	0.8922	6.9760	-0.0003	-0.0058	0.0746	1.6526	0.0711	1.5876

Table 2: Summary of the AR (1) Process Generated Using Deterministic and Stochastic Trend for Case II

Function	Deterministic Trend	Stochastic Trend
Model	$X_t = \alpha^* + \beta^* t + \varphi_1 X_{t-1} + u_t$	$X_t = \alpha + (1+\varphi_1) X_{t-1} + u_t$
Mean	$E(X_t) = (1+\varphi_1+\varphi_1^2+\ldots+\varphi_1^{t-1})\alpha^* + \begin{pmatrix} t+(t-1)\varphi_1+ \\ (t-2)\varphi_1^2+\ldots+\varphi_1^{t-1} \end{pmatrix}\beta^* + \varphi_1^t X_0$	$E(X_t) = E\begin{pmatrix} \alpha^{**}+(1+\varphi_1)^t X_0 + u_1(1+\varphi_1)^{t-1} \\ +u_2(1+\varphi_1)^{t-2}+\ldots+u_{t-1}(1+\varphi_1)+u_t \end{pmatrix}$ $= \alpha^{**}+(1+\varphi_1)^t X_0$
Mean at lag k	$E(X_{t-k}) = (1+\varphi_1+\varphi_1^2+\ldots+\varphi_1^{t-k-1})\alpha^*$ $+((t-k)+(t-k-1)\varphi_1+(t-k-2)\varphi_1^2+\ldots+\varphi_1^{t-k-1})\beta^* + \varphi_1^{t-k} X_0$	$E(X_{t-k}) = E\begin{pmatrix} \alpha^{**}+(1+\varphi_1)^{t-k} X_0 + u_1(1+\varphi_1)^{t-k-1} \\ +u_2(1+\varphi_1)^{t-k-2}+\ldots+u_{t-k-1}(1+\varphi_1)+u_{t-k} \end{pmatrix}$ $= \alpha^{**}+(1+\varphi_1)^{t-k} X_0$
Variance	$Var(X_t) = (\varphi_1^{2(t-1)}+\varphi_1^{2(t-2)}+\ldots+\varphi_1^2+1)\sigma^2$	$Var(X_t) = \left[(1+\varphi_1)^{2(t-1)}+(1+\varphi_1)^{2(t-2)}+\ldots+(1+\varphi_1)^2+1\right]\sigma^2$
Variance at lag k	$Var(X_{t-k}) = (\phi_1^{2(t-k-1)}+\phi_1^{2(t-k-2)}+\ldots+\phi_1^2+1)\sigma^2$	$Var(X_{t-k}) = \left[(1+\phi_1)^{2(t-k-1)}+(1+\phi_1)^{2(t-k-2)}+\ldots+(1+\phi_1)^2+1\right]\sigma^2$
Auto-covariance	$\gamma_k = (\varphi_1^{t-1}\varphi_1^{t-k-1}+\varphi_1^{t-2}\varphi_1^{t-k-2}+\ldots+\varphi_1^k)\sigma^2$	$\gamma_k = ((1+\varphi_1)^{2t-k-2}+(1+\varphi_1)^{2t-k-4}+\ldots+(1+\varphi_1)^k)\sigma^2$
Autocorrelation	$\rho_k = \frac{\varphi_1^{t-1}\varphi_1^{t-k-1}+\varphi_1^{t-2}\varphi_1^{t-k-2}+\ldots+\varphi_1^k}{\sqrt{(\varphi_1^{2(t-1)}+\ldots+\varphi_1^2+1)(\varphi_1^{2(t-k-1)}+\ldots+\varphi_1^2+1)}}$	$\rho_k = \frac{(1+\varphi_1)^{2t-k-2}+(1+\varphi_1)^{2t-k-4}+\ldots+(1+\varphi_1)^k}{\sqrt{((1+\varphi_1)^{2(t-1)}+\ldots+(1+\varphi_1)^2+1)((1+\varphi_1)^{2(t-k-1)}+\ldots+(1+\varphi_1)^2+1)}}$
Requirement to achieve Station-arity	Detrending	Differencing
Misspecification Penalty	$-\beta_1 + \varphi_1 X_{t-2} - u_{t-1}$	$-\beta_1 t + \varphi_1 Y_{t-1}$

$\alpha^{**} = \alpha\left(\frac{(1+\varphi_1)^{t-1}-1}{\varphi_1}\right)$,

Table 3: Summary of the MA (1) Process Generated Using Deterministic and Stochastic Trend for Case III

Function	MA(1)+Deterministic Trend	MA(1)+Stochastic Trend
Model	$X_t = \beta_0 + \beta_1 t + u_t - \theta_1 u_{t-1}$	$X_t = \alpha + X_{t-1} + u_t - \theta_1 u_{t-1}$
Mean	$E(X_t) = \beta_0 + \beta_1 t$	$E(X_t) = t\alpha + X_0 + \theta_1 \sum_{i=0}^{t-1} E(u_i) + \sum_{i=1}^{t} E(u_i) = t\alpha + X_0$
Mean at lag k	$E(X_{t-k}) = \beta_0 + \beta_1(t-k)$	$E(X_{t-k}) = (t-k)\alpha + X_0 + \theta_1 \sum_{i=0}^{t-k-1} E(u_i) + \sum_{i=1}^{t-k} E(u_i) = (t-k)\alpha + X_0$
Variance	$Var(X_t) = \sigma^2(1+\theta_1^2)$.	$Var(X_t) = \theta_1^2\left[(t-1)\sigma^2\right] + t\sigma^2 = \sigma^2\left[t+\theta_1^2(t-1)\right]$
Variance at lag k	$Var(X_{t-k}) = \sigma^2(1+\theta_1^2)$.	$Var(X_{t-k}) = \sigma^2\left[t-k+\theta_1^2(t-k-1)\right]$
Auto-covariance	$\gamma_k = \begin{cases} \sigma^2(1+\theta_1^2), & k=0 \\ \sigma^2\theta_1^2, & k=1 \\ 0, & k>1 \end{cases}$	$r_k = \sigma^2(2\theta_1(t-1)+t)$

Autocorrelation	$\rho_k = \begin{cases} 1, & k=0 \\ \theta_1^2/(1+\theta_1^2), & k=1 \\ 0, & k>1 \end{cases}$	$\rho_k = \dfrac{2\theta_1(t-1)+t}{\sqrt{\{[t+\theta_1^2(t-1)]*[t-k+\theta_1^2(t-k-1)]\}}}$
Requirement to achieve Stationarity	Detrending	Differencing
Misspecification Penalty	$-\beta_1+\theta_1(u_{t-1}-u_{t-2})-u_{t-1}$	$-\beta t+X_{t-1}+\theta u_{t-1}$

4.2. Ascertaining the nature of Trend in a given series using Spectrum

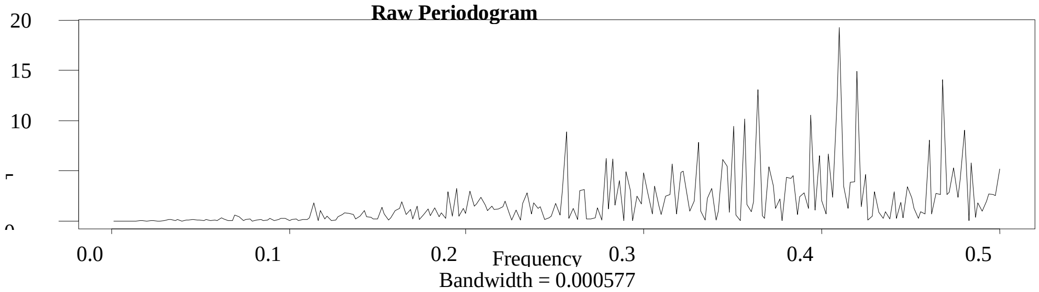


Fig. 1: The Plot of Spectrum (Periodiogram) of X_{1t}.

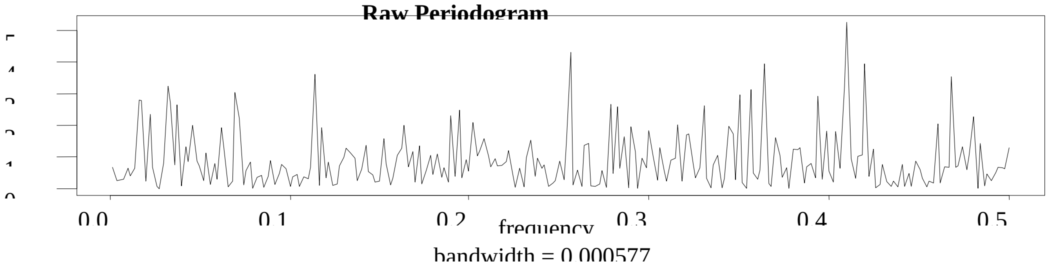


Fig. 2: The Plot of Spectrum (Periodiogram) Of X_{2t}.

Figures 4 and 5 are the periodogram of the simulated series (X_{1t}) and (X_{2t}) which are respectively generated using deterministic and stochastic trend. The spectrum at frequency zero for X1t is zero while for X2t is about 0.65 (positive). This result is in harmony with the theoretical result in section 3

4.3. Simulated series of autoregressive process of order 1 around a deterministic trend (Y1t) and stochastic trend (Y2t)

Table 6 contains the result for fitting a deterministic trend and stochastic trend plus AR(1) to Y_{1t}. The result review that both the detrended and differenced series have a significant AR(1) coefficient. However, the ACF and PACF of the residual is a white noise for the detrended but not for the differenced, rather a moving average component was introduced to the residual. This indicates that misspecifying the deterministic trend in Y_{1t} to be stochastic trend will produce an inadequate model. This result was also in harmony with the theoretical result in Table 2.

When a deterministic trend is fitted to Y_{2t}, a significant trend line was realized. But fitting AR(1) to the detrended series gave an estimate without convergence after 25 iterations. Conversely, fitting a stochastic trend to Y_{2t} by differencing and fitting AR (1) to the differenced series, gave a significant coefficients with uncorrelated residuals. The result of the estimate as well as the ACF and PACF of the residual of the AR (1) fitted to the differenced series is as shown in table 7

Table 6: Estimate of the Coefficient for Simulated Series (Y_{1t}) when Deterministic Trend and Stochastic Trend Is Fitted to It as Well as the ACF and PACF of the Residual when AR(1) Is Fitted to the Transformed Series

Model Estimate	Detrended series				Differenced series			
Constant	-0.012(0.947)				1.148(0.000)			
$\hat{\mu}$	-0.0021				0.897			
φ_1	0.4503(0.000)				-0.280(0.000)			
	When a deterministic trend and AR(1) is fitted				When a stochastic trend and AR(1) is fitted			
Lag	ACF	T	PACF	T	ACF	T	PACF	T
1	-0.0112	-0.2504	-0.0112	-0.2504	-0.0040	-0.0841	-0.0926	-2.0683
2	-0.0085	-0.1911	-0.0087	-0.1939	-0.0723	-1.5097	-0.1446	-3.2312
3	0.0127	0.2847	0.0125	0.2804	-0.0160	-0.3336	-0.0851	-1.9005
4	0.0354	0.7918	0.0356	0.7970	-0.0683	-1.4194	-0.1676	-3.7441
5	-0.0442	-0.9865	-0.0432	-0.9667	0.0563	1.1667	-0.0336	-0.7499

( ) P –Value

Table 7: Estimate of the Coefficient for Simulated Series (Y_{2t}) when Deterministic Trend and Stochastic Trend Is Fitted to It as Well as the ACF and PACF of the Residual when AR (1) Is Fitted to the Transformed Series

Model Estimate	Detrended series	differenced series
Constant	Convergence criteria not met after 25 iterations though estimation was made.	1.024(0.000)
$\hat{\mu}$		0.032
ϕ_1		-0.149(0.001)

ACF and PACF of the residuals of the model								
	When a deterministic trend and AR(1) is fitted				When a stochastic trend and AR(1) is fitted			
Lag	ACF	T	PACF	T	ACF	T	PACF	T
1	0.1545	0.9272	0.1545	0.9272	-0.0047	-0.1058	-0.0047	-0.1058
2	-0.2158	-1.2652	-0.2456	-1.4735	-0.0388	-0.8671	-0.0388	-0.8676
3	-0.0214	-0.1201	0.0634	0.3801	0.0664	1.4818	0.0662	1.4780
4	0.1281	0.7190	0.0734	0.4405	0.0671	1.4905	0.0665	1.4852
5	0.0280	0.1550	-0.0045	-0.0272	-0.0232	-0.5120	-0.0176	-0.3933

4.4. Ascertaining the nature of trend in Y1t and Y2t using spectrum

To ascertain the nature of trend and avoid the stress of trial and error as well as avoid specification error, the population spectrum of the first difference of Y_{1t} and Y_{2t} was estimated. The result is as displayed in figures 3 and 4. In agreement with the theoretical result in section 3, the spectrum of the first difference of Y_{1t} at frequency zero is zero indicating that Y1t is generated from a deterministic trend while the spectrum of the first difference of Y_{2t} at frequency zero is positive indicating that Y_{2t} is of stochastic trend.

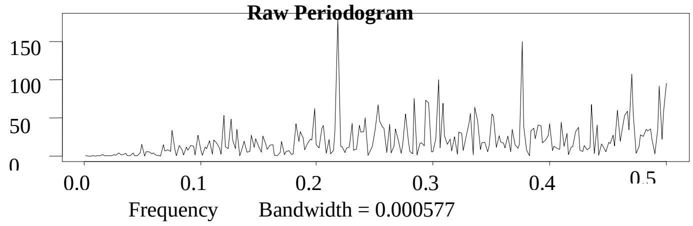


Fig. 3: The Spectrum (Periodiogram) of the First Differenced ofY_{1t}Series.

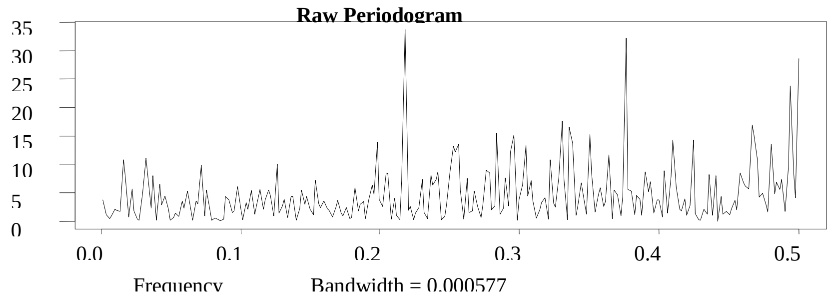


Fig. 4: The Spectrum (Periodiogram) of the First Differenced of Y_{2t}Series.

4.5. Simulated series with moving average process of order 1 around with deterministic trend (Z1t) and stochastic trend (Z2t)

Fitting MA(1) to the detrended (Z_{t1}-T) produced a model estimate displayed in table 7 with the coefficient of MA(1) being significant. But when MA(1) was fitted to the difference, convergence was unable to be attended even after 25 iteration, suggesting that the series when mis-specified could cause the MA process not to be invertible. Though a model estimate was realized, but this could not be used when convergence was not achieved. Another deceiving issue is that though convergence was not achieved, the residual from the wrongly specified model was a white noise just like the one correctly specified. Hence, caution should be taken to detect specification error when modelling a series characterized with trend. The result is in harmony with the theoretical result in Table 7 of section 3.
The result of the analysis of simulated series with MA(1) process around a stochastic trend is as displayed in table 9. The result showed that when deterministic trend is imposed on the series and the MA(1) fitted on the detrended series, a significant quadratic trend line is observed as well as a significant MA(1) coefficient, however, the residual was not a white noise making the model inadequate. When AR(2) is fitted to the detrended series, a significant coefficient of 0.6859 and 0.2638 for ϕ_1 *and* ϕ_2 respectively with a uncorrected (white noise) residual was observed. Confirming the theoretical result that mis-specifying stochastic trend to deterministic trend in a model introduces AR process to residual of such model. On the converse, when a stochastic trend was fitted to Z_{2t} and MA(1) fitted to the differenced series, a significant coefficient as well as a white noise residual was observed. The detail is as shown in table 9.

Table 8: Estimate of the Coefficient for Simulated Series (Z_{1t}) when Deterministic Trend and Stochastic Trend Is Fitted to It as Well as the ACF and PACF of the Residual when MA (1) Is Fitted to the Detrended/Differenced Series

Model Estimate	Detrended series				Differenced series			
Constant	-0.0016(0.996)				Convergence criteria not met after 25 iterations though estimation was made.			
$\hat{\mu}$	-0.0016							
θ_1	-0.7369(0.000)							
	When a deterministic trend and MA(1) is fitted				When a stochastic trend and MA1) is fitted			
Lag	ACF	T	PACF	T	ACF	T	PACF	T
1	0.0413	0.9235	0.0413	0.9235	0.0307	0.6850	0.0307	0.6850
2	0.0162	0.3617	0.0145	0.3247	0.0227	0.5077	0.0218	0.4876
3	0.0411	0.9173	0.0399	0.8927	0.0341	0.7611	0.0328	0.7331
4	0.0551	1.2276	0.0518	1.1578	0.0578	1.2880	0.0555	1.2391

5	-0.0320	-0.7097	-0.0375	-0.8393	-0.0364	-0.8083	-0.0413	-0.9226

( ) P –Value.

Table 9: Estimate of the Coefficient for Simulated Series (Z_{2t}) when Deterministic Trend and Stochastic Trend Is Fitted to It as Well as the ACF and PACF of the Residual when MA(1) Is Fitted to the Detrended/Differenced Series

Model Estimate			Detrended series			Differenced series		
Constant			0.0038(0.979)			0.0500 (0.134)		
$\hat{\mu}$			0.0038			0.0500		
θ_1			-0.7232(0.000)			0.2754(0.000)		
	When a deterministic trend and MA(1) is fitted				When a stochastic trend and MA(1) is fitted			
Lag	ACF	T	PACF	T	ACF	T	PACF	T
1	0.5150	11.5160	0.5150	11.5160	-0.0088	-0.1976	-0.0088	-0.1976
2	0.8355	15.1026	0.7762	17.3564	0.0649	1.4498	0.0648	1.4483
3	0.5537	7.2374	0.1974	4.4146	-0.0687	-1.5286	-0.0679	-1.5168
4	0.7027	8.3518	-0.0265	-0.5931	-0.1116	-2.4712	-0.1177	-2.6293
5	0.5699	5.9891	0.0507	1.1330	0.0065	0.1415	0.0138	0.3081

( ) p –value.

4.6. Ascertaining the nature of trend in Z1t and Z2t using spectrum

To ascertain the nature of trend, the population spectrum of the first difference of Z_{1t} and Z_{2t} was estimated and its periodiagram displayed in figure 20 and 21 respectively. In agreement with the theoretical result in section 3, the spectrum of the first difference of Z_{1t} at frequency zero is zero indicating that Z_1t is generated from a deterministic trend while the spectrum of the first difference of Z_{2t} at frequency zero is positive indicating that Z_{2t} is of stochastic trend.

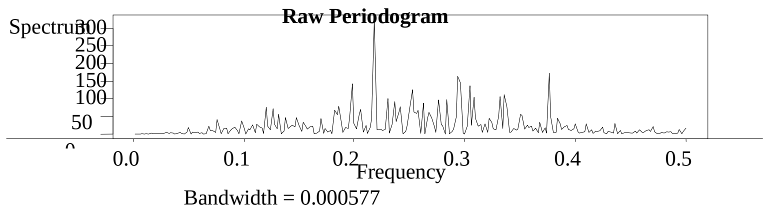


Fig. 5: The Spectrum (Periodiogram) of the First Differenced of Z_{1t}Series.

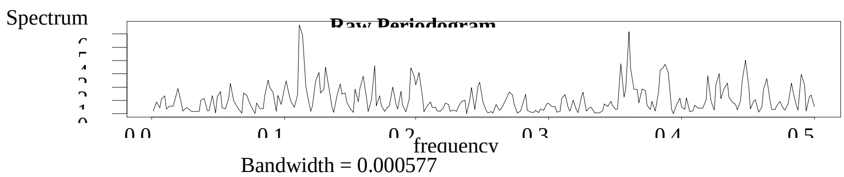


Fig. 6: The Spectrum (Periodiogram) of the First Differenced of Z_{2t}Series.

5. Conclusions

The penalty of misspecification in a time series dominated with trend was examined in this work. The two most common trends (deterministic and stochastic trend) were studied.Population spectrum approach for detection of nature of trend was recommended as this will save us the pitfall of consequences of model misspecification in time series analysis.

References

[1] Ademola, O. P. (2007). *Fuzzy-wavelet method for time series analysis*. University of Surrey Guildford, Surrey GU2 7XH, UK. (PhD thesis).
[2] Box, G.E.P. and Pierce, D.A. (1970). *Distributions of residual autocorrelations in autoregressive moving average time series models*. Journal of American Statistical Association, (65), 1509-1526.https://doi.org/10.1080/01621459.1970.10481180.
[3] Box, G.E.P., Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control (3rd Edition).* Prentice-Hall, Inc, Englewood Cliffs.
[4] Chandler, R. E. and Scott, E. M. (2011), Statistical Methods for Trend Detection and Analysis in the Environmental Sciences. John Wiley and Sons Ltd, USA.https://doi.org/10.1002/9781119991571.
[5] Chatfield, C. (2004). *The analysis of time series: An Introduction*, (6th Edition), Chapman & Hall/CRC Press Company Boca, New York Washington, D.C.
[6] Dickey, D.A. and Fuller, W.A. (1979). *Distribution of the estimates for autoregressive time series with a unit root*. Journal of the American Statistical Association,*74* (366), 427-431.https://doi.org/10.1080/01621459.1979.10482531.
[7] Granger, C. W. J. (1994). *Forecasting in Economics.Time Series Prediction: Forecasting the Future and Understanding the Past*. N. A. Gershenfeld and A. S. Weigend (eds.) Reading, MA: Addison-Wesley.
[8] Gujarati, D. N. (2004). *Basic Econometrics. (4th Edition).* The McGraw−Hill Companies, New York.
[9] Hamilton, J, D. (1994). *Time Series Analysis*. Princeton University Press Princeton, New Jersey.
[10] Heino, B. N. (2005), Non-Stationary Time Series and Unit Root Tests, Unpublished Note on Econometric Fall.
[11] Kwiatkowski, D., Philips, P.C.B., Schmitt, P. and Shin, Y. (1992). *Testing the null hypothesis of stationarity against the alternative of a unit root.* Journal of Econometrics, (54),159-178.https://doi.org/10.1016/0304-4076(92)90104-Y.
[12] Ljung, G. M. and Box, G. E. P. (1978). *On a measure of lack of fit in time series models*. Biometrika, 65 (2), 297-303.https://doi.org/10.1093/biomet/65.2.297.

[13] Nelson, C. R. and Plosser, C. I. (1982). *Trends and random walks in macroeconomic time series: Some evidence and implications*. Journal of Monetary Economics.139-162.https://doi.org/10.1016/0304-3932(82)90012-5.

[14] Nelson, M., Hill, T., Remus, T. and O'Connor, M. (1999). *Time series forecasting using neural networks: Should the data be deseasonalized first?* Journal of Forecasting,18, 359–367.https://doi.org/10.1002/(SICI)1099-131X(199909)18:5<359::AID-FOR746>3.0.CO;2-P.

[15] Ogbonna, C.J., Nweke, C.J., Nwogu E. C. and Iwueze, I.S. (2016). Wavelet Transform as an Alternative to Power Transformation in Time Series Analysis. *Bulletin of Mathematical Sciences and Applications Vol. 17, pp 57-74, SciPress Ltd., Switzerland.*https://doi.org/10.18052/www.scipress.com/BMSA.17.57.

[16] Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*.*75*(2), 335-346.https://doi.org/10.1093/biomet/75.2.335.

[17] Robinson, P. M. (2003), Time Series with Long Memory, Advanced Texts in Econometrics, Oxford Univ. Press, New York.

[18] Ruey, S.T. (1988). *Outliers, level shift and variance changes in time series*. Journal of Forecasting, *7*, 1-20.https://doi.org/10.1002/for.3980070102.

[19] Said, S. E. and Dickey, D. A. (1984). *Testing for unit roots in autoregressive-moving average models of unknown order.* Biometrika, *71*(3), 599-607.https://doi.org/10.1093/biomet/71.3.599.

[20] Stefan, L., Tomas, V. and Eduard, B. (2011). *Unit root and stationary testing with empiricalapplication on industrial production of Central Eastern Europe countries*. Munich personal repec Archive No. 29648.

[21] Tse, R. Y. C. (1997). *An application of the ARIMA model to real-estate prices in Hong Kong*. Journal of Property Finance, *8* (2) 152-163.https://doi.org/10.1108/09588689710167843.

[22] Wei, W. S. (1990). *Time Series Analysis: Univariate and Multivariate Methods (2nd Edition).* Pearson Addison-Wesley Publishing Company, Inc. USA.

[23] Zhang, B. L., Coggins, R., Jabri, M.A., Dersch, D. and Flower, B. (2001). *Multiresolution forecasting for future trading using wavelet decompositions.* Institute of electrical and Electronic Engineers Transactions on Neural Networks,*12*(4), 765-775.https://doi.org/10.1109/72.935090.

[24] Zhang, G. P. and Qi, M. (2005). *Neural network forecasting for seasonal and trend time series*. European Journal of Operational Research, *160*(2), 501-514.https://doi.org/10.1016/j.ejor.2003.08.037.