Abstract

This paper compares a Least-Squared Random Coefficient Autoregressive (RCA) model with a Least-Squared RCA model based on Autocorrelated Errors (RCA-AR). We looked at only the first order models, denoted RCA(1) and RCA(1)-AR(1). The efficiency of the Least-Squared method was checked by applying the models to Brownian motion and Wiener process, and the efficiency followed closely the asymptotic properties of a normal distribution. In a simulation study, we compared the performance of RCA(1) and RCA(1)-AR(1) by using the Mean Square Errors (MSE) as a criterion. The RCA(1) exhibited good power estimation in both cases where the data is stationary and nonstationary. On the other hand, when data oscillates around its mean, RCA(1)-AR(1) performed better. For real world data, we applied the two models to the daily volume of the Thai gold price and found that RCA(1)-AR(1) performed better than RCA(1).

Keywords: Autocorrelated Errors, Brownian motion, Mean Square Errors, Random Coefficient Autoregressive, Wiener process;

Contents

List of Tables

List of Figures

Chapter 1

The Least Square method of RCA-AR Model

1.1 Introduction

In recent years, models of time series data have been applied to the fields of finance, business, and economy. Time series data in the field of economy can either show a stationary or a nonstationary time series data. There are several stationary models fitted the stationary data such as the Autoregrssive (AR) model, Moving Average (MA) model, and Autoregressive Moving Average (ARMA) model.

For nonstationary time series data, the Autoregressive Integrated Moving Average (ARIMA) model can be used. The RCA Random Coefficient Autoregressive (RCA) model, introduced by Nicholls and Quinn (1982), is another one. Presently, it is very popular because it uses past data to help estimate parameters.

Nicholls and Quinn(1982) employed the least square method and the maximum likelihood method to estimate parameters. Wang and Ghosh(2002) used the Bayesian approach to obtain the first order estimate of an RCA model. Thavaneswaran and Abraham (1988) applied an estimating function to nonlinear time series data. The estimating function technique is equivalent to a weighted least square estimator discussed by Hwang and Basawa (1998) and Chandra and Taniguchi (2001).

Prediction based on an estimate of a time series model is difficult because of the effect of autocorrelated error. To address this type of error, various nonlinear autoregressive models have shown up in the literature: Haggan and Ozaki(1981) modeled nonlinear vibration by using an amplitude-dependent autoregressive time series model entitled Exponential Autoregressive (EXPAR) model; Tong(1983) introduced the Threshold Autoregressive (TAR) model of nonlinear time series; and Chan and Tong(1986) developed a TAR model into the Smooth-Transition Autoregressive (STAR) model.

This study investigated an RCA model based on autocorrelated error and proposed applying the Least Square (LS) method to estimate its parameters as well as compared it to a model without autocorrelated error. The efficiency of an LS estimator can be validated by applying it to Brownian motion and Wiener process. Finally, we compared the performance of the LS method with that of the Mean Square Error (MSE) method using both simulated and real data.

The rest of the paper is organized as follows: Section 1.2 describes two RCA Models, with and without autocorrelated error; Section 1.3 describes the LS method for parameter estimation; Section 1.4 showed the efficiency of the LS estimator. We applied our proposed method to simulated data and real data (Thai daily gold price) in Section 3.5 and 1.6. Finally, in Section 3.6, we discussed the results and indicated some directions for further research.

1.2 RCA model with Autocorrelated Errors

The general class of Random Coefficient Autoregressive model of order p, Wang and Ghosh(2002) written RCA(p) as

$$\begin{aligned} x_t &= \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \varepsilon_t \qquad , t = 1, 2, \ldots, n \qquad (1.1) \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \end{aligned}$$

where α is the scalar of constant, $\underline{\beta}_t = (\beta_{t1}, \ldots, \beta_{tp})^\top$ are the sequence of iid (independent and identically distributed) random variables, $\underline{\mu}_\beta = (\mu_{\beta 1}, \ldots, \mu_{\beta p})^\top$ and ε_t 's are the iid from a distribution with mean zero and unit variance. Empirically, we can see that the time series data are appeared the autocorrelated in error terms which is applied by AutoRegressive (AR) process following

$$\varepsilon_t = \sum_{j=1}^{q} \rho_j \varepsilon_{t-j} + e_t \qquad ,t = 1, 2, \ldots, n \tag{1.2}$$

The RCA model with autocorrelated errors denoted RCA(p)-AR(q) can be rewritten by

$$\begin{aligned} x_t &= \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \sum_{j=1}^{q} \rho_j \varepsilon_{t-j} + e_t \qquad ,t = 1, 2, \ldots, n \qquad (1.3) \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \end{aligned}$$

where $\underline{\rho} = (\rho_1, \ldots, \rho_q)^\top$ is the $q \times 1$ vector of constant.

In this paper, we will study the first order of RCA model with autocorrelated errors denoted by RCA(1)-AR(1) that can be written as

$$\begin{aligned} x_t &= \alpha + \beta_t x_{t-1} + \rho \varepsilon_{t-1} + e_t \qquad ,t = 1, 2, \ldots, n \qquad (1.4) \\ \beta_t &= \mu_\beta + \sigma_\beta u_t \end{aligned}$$

where x_t's are iid random variables with mean μ_β, and variance σ_β^2, ε_t's are iid random variables with mean 0 and variance σ_ε^2, and β_t's and ε_t's are independent.

For the parameter estimation of RCA(1)-AR(1), it can be seen from (2.4) that consisted of the intercept term α, the mean μ_β, variance σ_β^2 of the coefficient β_t, the variance σ_ε^2 of the ε_t, the coefficient ρ of AR process, and the variance σ_e^2 of the e_t, or defined as $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2, \sigma_e^2, \rho)^\top$.

1.3 Parameter Estimation for RCA(1)-AR(1)

To estimate parameter of RCA(1)-AR(1), we propose the concept of LS method to estimate parameter $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2, \sigma_e^2, \rho)^\top$ by minimizing sum of residuals square in 3 steps.

The first step, we consider the RCA(1) model following

$$\begin{aligned} x_t &= \alpha + \beta_t x_{t-1} + \varepsilon_t \qquad ,t = 1, 2, \ldots, n \qquad (1.5) \\ \beta_t &= \mu_\beta + \sigma_\beta u_t \end{aligned}$$

The parameter of RCA(1) is $(\mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2)$ based on the LS estimation. Given the sample $x_1, x_2, \ldots, x_n$, let $\varepsilon_t = x_t - \alpha - \mu_\beta x_{t-1}$, then $\hat{\alpha}$ is to estimate by minimizing sum of square errors,

$$\begin{aligned}
\varepsilon_t &= x_t - \alpha - \mu_\beta x_{t-1} \\
\sum_{t=1}^{n}(\varepsilon_t)^2 &= \sum_{t=1}^{n}(x_t - \alpha - \mu_\beta x_{t-1})^2 \\
\frac{\partial}{\partial\alpha}\sum_{t=1}^{n}(\varepsilon_t)^2 &= -2\sum_{t=1}^{n}(x_t - \alpha - \mu_\beta x_{t-1}) = 0 \\
\hat{\alpha} &= \frac{\sum_{t=1}^{n} x_t}{n} - \mu_\beta \frac{\sum_{t=1}^{n} x_{t-1}}{n} \qquad (1.6)
\end{aligned}$$

The LS estimate of $\hat{\mu}_\beta$ is given by,

$$\begin{aligned}
\frac{\partial}{\partial\mu_\beta}\sum_{t=1}^{n}(\varepsilon_t)^2 &= -2\sum_{t=1}^{n}(x_t - \alpha - \mu_\beta x_{t-1})x_{t-1} = 0 \\
\hat{\mu}_\beta &= \frac{\sum_{t=1}^{n} x_t x_{t-1} - \hat{\alpha}\sum_{t=1}^{n} x_{t-1}}{\sum_{t=1}^{n} x_{t-1}^2} \qquad (1.7)
\end{aligned}$$

From (1.7), let us replace in (1.6) and the solution of $\hat{\alpha}$ is,

$$\hat{\alpha} = \frac{\sum_{t=1}^{n} x_{t-1}^2 \sum_{t=1}^{n} x_t - \sum_{t=1}^{n} x_t x_{t-1} \sum_{t=1}^{n} x_{t-1}}{n\sum_{t=1}^{n} x_{t-1}^2 - (\sum_{t=1}^{n} x_{t-1})^2} \qquad (1.8)$$

Thus $\hat{\mu}_\beta$ is computed by,

$$\hat{\mu}_\beta = \frac{n\sum_{t=1}^{n} x_t x_{t-1} - \sum_{t=1}^{n} x_t \sum_{t=1}^{n} x_{t-1}}{n\sum_{t=1}^{n} x_{t-1}^2 - (\sum_{t=1}^{n} x_{t-1})^2} \qquad (1.9)$$

For RCA(1) model, it can be written as

$$x_t = \hat{\alpha} + \hat{\mu}_\beta x_{t-1} \qquad , t = 1, 2, \ldots, n$$

Hence, the estimated errors can be denoted by

$$\hat{\varepsilon}_t = x_t - \hat{\alpha} + \hat{\mu}_\beta x_{t-1} \qquad , t = 1, 2, \ldots, n$$

The second step, we assume that the errors of time series data have an autocorrelation function. The RCA(1)-AR(1) model used the concept of LS method by minimizing sum of square of autocorrelated errors as;

$$\begin{aligned}
e_t &= x_t - \hat{\alpha} - \hat{\mu}_\beta x_{t-1} - \rho\hat{\varepsilon}_{t-1} \qquad , t = 1, 2, \ldots, n \\
\sum_{t=1}^{n} e_t &= \sum_{t=1}^{n}(x_t - \hat{\alpha} - \hat{\mu}_\beta x_{t-1} - \rho\hat{\varepsilon}_{t-1})^2 \\
\frac{\partial}{\partial\rho}\sum_{t=1}^{n}(e_t)^2 &= -2\sum_{t=1}^{n}(x_t - \hat{\alpha} - \hat{\mu}_\beta x_{t-1} + \rho\hat{\varepsilon}_{t-1})\hat{\varepsilon}_{t-1} = 0
\end{aligned}$$

We get $\hat{\rho}$ as,

$$\hat{\rho} = \frac{\sum_{t=1}^{n} x_t \hat{\varepsilon}_{t-1} - \hat{\alpha} \sum_{t=1}^{n} \hat{\varepsilon}_{t-1} - \hat{\mu}_\beta \sum_{t=1}^{n} x_{t-1} \hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\varepsilon}_{t-1}^2} \tag{1.10}$$

For RCA(1)-AR(1) model, we write

$$x_t = \hat{\alpha} + \hat{\mu}_\beta x_{t-1} + \hat{\rho} \hat{\varepsilon}_{t-1} \qquad , t = 1, 2, \ldots, n$$

The final step, we let F_t be the information set up to time t, and denote $\varepsilon_t = x_t - \alpha + \mu_\beta x_{t-1}$, then it is seen that

$$E(\varepsilon_t | F_t) = 0 \text{ and } E(\varepsilon_t^2 | F_t) = \sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2$$

The estimation process is to the form of residuals $\hat{\varepsilon}_t = x_t - \hat{\alpha} - \hat{\mu}_\beta x_{t-1}$, and let $r_t^2 = \hat{\varepsilon}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2$. Nicholls and Quinn (1981) showed the LS estimators of σ_ε^2 and σ_β^2 by regressing $\hat{\varepsilon}_t^2$ on x_{t-1}^2, which are equivalent to minimizing $\sum_{t=1}^{n} \left(\hat{\varepsilon}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 \right)^2$ given by

$$\sum_{t=1}^{n} r_t^2 = \sum_{t=1}^{n} \left(\hat{\varepsilon}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 \right)^2$$

Consequently, the LS estimates of σ_ε^2 and σ_β^2 are obtained by,

$$\begin{aligned} \frac{\partial}{\partial \sigma_\varepsilon^2} \sum_{t=1}^{n} r_t^2 &= -2 \sum_{t=1}^{n} \left(\hat{\varepsilon}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 \right) = 0 \\ \sum_{t=1}^{n} \hat{\varepsilon}_t^2 &= n\sigma_\varepsilon^2 + \sigma_\beta^2 \sum_{t=1}^{n} x_{t-1}^2 \\ \hat{\sigma}_\varepsilon^2 &= \frac{\sum_{t=1}^{n} \hat{\varepsilon}_t^2 - \sigma_\beta^2 \sum_{t=1}^{n} x_{t-1}^2}{n} \end{aligned} \tag{1.11}$$

and

$$\begin{aligned} \frac{\partial}{\partial \sigma_\beta^2} \sum_{t=1}^{n} r_t^2 &= -2 \sum_{t=1}^{n} \left(\hat{\varepsilon}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 \right) x_{t-1}^2 = 0 \\ \sum_{t=1}^{n} \hat{\varepsilon}_t^2 x_{t-1}^2 &= \sigma_\varepsilon^2 \sum_{t=1}^{n} x_{t-1}^2 + \sigma_\beta^2 \sum_{t=1}^{n} x_{t-1}^4 \\ \hat{\sigma}_\beta^2 &= \frac{\sum_{t=1}^{n} \hat{\varepsilon}_t^2 x_{t-1}^2 - \sigma_\varepsilon^2 \sum_{t=1}^{n} x_{t-1}^2}{\sum_{t=1}^{n} x_{t-1}^4} \end{aligned} \tag{1.12}$$

The $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_\beta^2$ can be written in general form as

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^n \hat{\varepsilon}_t^2 \sum_{t=1}^n x_{t-1}^4 - \sum_{t=1}^n x_{t-1}^2 \sum_{t=1}^n \hat{\varepsilon}_t^2 x_{t-1}^2}{n\sum_{t=1}^n x_{t-1}^4 - \left(\sum_{t=1}^n x_{t-1}^2\right)^2} \quad (1.13)$$

$$\hat{\sigma}_\beta^2 = \frac{n\sum_{t=1}^n \hat{\varepsilon}_t^2 x_{t-1}^2 - \sum_{t=1}^n \hat{\varepsilon}_t^2 \sum_{t=1}^n x_{t-1}^2}{n\sum_{t=1}^n x_{t-1}^4 - \left(\sum_{t=1}^n x_{t-1}^2\right)^2} \quad (1.14)$$

Notice that for σ_e^2, we let F_t be the information set up to time t, and denote $e_t = x_t - \alpha + \mu_\beta x_{t-1} + \rho\varepsilon_{t-1}$, then it is seen that

$$E(e_t|F_t) = 0 \quad \text{and} \quad E(e_t^2|F_t) = \sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2 + \sigma_e^2\rho^2$$

The estimation of σ_e^2 is used by corresponding LS method as σ_ε^2, and σ_β^2. Let $\hat{e}_t = x_t - \hat{\alpha} - \hat{\mu}_\beta x_{t-1} - \hat{\rho}$, $s_t^2 = \hat{e}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 - \hat{\sigma}_e^2\hat{\rho}^2$, and regress $\hat{e}_t^2$ on ρ^2, which are equivalent to minimizing

$\sum_{t=1}^n(\hat{e}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 - \sigma_e^2\hat{\rho}^2)$, and get

$$\sum_{t=1}^n s_t^2 = \sum_{t=1}^n \left(\hat{e}_t^2 - \sigma_\varepsilon^2 - \sigma_\beta^2 x_{t-1}^2 - \hat{\sigma}_e^2\hat{\rho}^2\right)^2$$

Using the concept of LS method, σ_e^2 is computed by,

$$\frac{\partial}{\partial\sigma_e^2}\sum_{t=1}^n s_t^2 = -2\sum_{t=1}^n \left(\hat{e}_t^2 - \hat{\sigma}_\varepsilon^2 - \hat{\sigma}_\beta^2 x_{t-1}^2 - \hat{\sigma}_e^2\hat{\rho}^2\right)\hat{\rho}^2 = 0$$

$$\sum_{t=1}^n \hat{e}_t^2 = n\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\beta^2\sum_{t=1}^n x_{t-1}^2 - n\hat{\sigma}_e^2\hat{\rho}^2$$

$$\hat{\sigma}_e^2 = \frac{\sum_{t=1}^n \hat{e}_t^2 - n\hat{\sigma}_\varepsilon^2 - \hat{\sigma}_\beta^2\sum_{t=1}^n x_{t-1}^2}{n\hat{\rho}^2} \quad (1.15)$$

1.4 Efficiency of LS Estimator

1.4.1 RCA(1) model

Consider the RCA(1) model

$$x_t = \alpha + \mu_\beta x_{t-1} + \varepsilon_t \quad, t = 1, 2, \ldots, n \quad (1.16)$$

The RCA(1) at (3.1) can be written in terms of the matrix form as a regression model following

$$\mathbf{Y} = \mathbf{X}\underline{\beta} + \underline{\varepsilon}$$

where

$$\mathbf{Y} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_0 \\ \vdots & \vdots \\ 1 & x_{n-1} \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \alpha \\ \mu_\beta \end{bmatrix}, \text{and } \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Nicholls and Quinn (1982) showed that under the normality of $\underline{\beta}$ and $\underline{\varepsilon}$, if the process is second order stationary, then $\hat{\theta}$ is consistent for θ and moreover, if the forth moments of $\underline{\beta}$ and $\underline{\varepsilon}$, then $\sqrt{n}(\hat{\theta}-\theta)$ has a limiting normal distribution. In this case, it is shown that if the α and μ_β consistentcy estimates for $\hat{\alpha}$ and $\hat{\mu}_\beta$, then $\sqrt{n}(\hat{\alpha}-\alpha)$ and $\sqrt{n}(\hat{\mu}_\beta-\mu_\beta)$ has a limiting normal distribution.

Phillips(1987) showed that the sample moment of $\{x_t\}$ converges to random functions of Brownian motion or Wiener process :

$$\begin{aligned} n^{-3/2}\sum_{t=1}^{n} x_{t-1} &\xrightarrow{d} \sigma_\varepsilon \int_0^1 W(r)dr \\ n^{-2}\sum_{t=1}^{n} x_{t-1}^2 &\xrightarrow{d} \sigma_\varepsilon^2 \int_0^1 W(r)^2 dr \\ n^{-1}\sum_{t=1}^{n} x_{t-1}\,\varepsilon_t &\xrightarrow{d} \sigma_\varepsilon^2 \int_0^1 W(r)\, dW(r) \end{aligned}$$

where $W(r)$ denotes a standard Brownian motion. Using the above equation Phillips showed that

$$\begin{aligned} n(\hat{\alpha}-\alpha) &\xrightarrow{d} N(0,\sigma_\alpha) \\ n(\hat{\mu}_\beta-\mu_\beta) &\xrightarrow{d} N(0,\sigma_\beta) \end{aligned}$$

The LS estimator, the estimator $(\underline{\hat{\beta}})$ is given by $\underline{\hat{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Then

$$\underline{\hat{\beta}} = \begin{bmatrix} \hat{\alpha} \\ \hat{\mu}_\beta \end{bmatrix} = \begin{bmatrix} n & \sum_{t=1}^{n} x_{t-1} \\ \sum_{t=1}^{n} x_{t-1} & \sum_{t=1}^{n} x_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{n} x_t \\ \sum_{t=1}^{n} x_{t-1}x_t \end{bmatrix}$$

when following Dickey and Fuller(1979), we use $(\underline{\hat{\beta}}-\underline{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\underline{\varepsilon}$, so we have

$$\begin{bmatrix} \hat{\alpha}-\alpha \\ \hat{\mu}_\beta-\mu_\beta \end{bmatrix} = \begin{bmatrix} n & \sum_{t=1}^n x_{t-1} \\ \sum_{t=1}^n x_{t-1} & \sum_{t=1}^n x_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^n \varepsilon_t \\ \sum_{t=1}^n x_{t-1}\varepsilon_t \end{bmatrix}$$

$$= \frac{1}{\delta} \begin{bmatrix} \sum_{t=1}^n x_{t-1}^2 & -\sum_{t=1}^n x_{t-1} \\ -\sum_{t=1}^n x_{t-1} & n \end{bmatrix} \begin{bmatrix} \sum_{t=1}^n \varepsilon_t \\ \sum_{t=1}^n x_{t-1}\varepsilon_t \end{bmatrix}$$

$$= \frac{1}{\delta} \begin{bmatrix} \sum_{t=1}^n \varepsilon_t \sum_{t=1}^n x_{t-1}^2 - \sum_{t=1}^n x_{t-1} \sum_{t=1}^n x_{t-1}\varepsilon_t \\ n\sum_{t=1}^n x_{t-1}\varepsilon_t - \sum_{t=1}^n x_{t-1}\sum_{t=1}^n \varepsilon_t \end{bmatrix} \quad (1.17)$$

where $\delta = n\sum_{t=1}^n x_{t-1}^2 - (\sum_{t=1}^n x_{t-1})^2$. Araveeporn (2012) studied the limiting distribution using $\sqrt{n}(\hat{\theta}-\theta)$ as

$$n(\hat{\alpha}-\alpha) = \frac{n\sum_{t=1}^n x_{t-1}^2 \sum_{t=1}^n \varepsilon_t - n\sum_{t=1}^n x_{t-1}\sum_{t=1}^n x_{t-1}\varepsilon_t}{n\sum_{t=1}^n x_{t-1}^2 - (\sum_{t=1}^n x_{t-1})^2}$$

$$\xrightarrow{d} \frac{n^3\sigma_\varepsilon^2 \int_0^1 W(r)^2 dr W_t - n\, n^{3/2}\sigma_\varepsilon \int_0^1 W(r)dr n\sigma_\varepsilon^2 \int_0^1 W(r)dW(r)}{\delta}$$

where $W_t = \sum_{t=1}^n \varepsilon_t$. The t statistics of α can be written as

$$t_\alpha = \frac{\hat{\alpha}-\alpha}{S_{\hat{\alpha}}}$$

where

$$S_{\hat{\alpha}} = \left\{ \sigma_\varepsilon^2 \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} n & \sum_{t=1}^n x_{t-1} \\ \sum_{t=1}^n x_{t-1} & \sum_{t=1}^n x_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}^{1/2}$$

Therefore

$$S_{\hat{\alpha}}^2 = \sigma_\varepsilon^2 \left[\frac{\sum_{t=1}^n x_{t-1}^2}{n\sum_{t=1}^n x_{t-1}^2 - (\sum_{t=1}^n x_{t-1})^2} \right]$$

Hence,

$$n^2 S_{\hat{\alpha}}^2 \xrightarrow{d} n^2\sigma_\varepsilon^2 \left[\frac{n^2\sigma_\varepsilon^2 \int_0^1 W(r)^2 dr}{\delta} \right] \quad (1.18)$$

Rewriting t_α as

$$t_\alpha = \frac{n(\hat{\alpha}-\alpha)}{[n^2 S_{\hat{\alpha}}^2]^{1/2}} \quad (1.19)$$

and using (1.18) and (1.19), we get

$$t_\alpha = \frac{n(\hat{\alpha}-\alpha)}{[n^2 S_{\hat{\alpha}}^2]^{1/2}}$$

$$\xrightarrow{d} \frac{\int_0^1 W(r)^2 dr W_t - n^{1/2}\sigma_\varepsilon \int_0^1 W(r)dr \int_0^1 W(r)dW(r)}{\left[n\sigma_\varepsilon^2 \int_0^1 W(r)^2 dr \right]^{1/2}} \quad (1.20)$$

From (1.20), we show the limiting distribution of μ_β as

$$\begin{aligned} n(\hat{\mu}_\beta - \mu_\beta) &= \frac{n^2 \sum_{t=1}^n x_{t-1}\varepsilon_t - n\sum_{t=1}^n x_{t-1} \sum_{t=1}^n \varepsilon_t}{n\sum_{t=1}^n x_{t-1}^2 - (\sum_{t=1}^n x_{t-1})^2} \\ &\xrightarrow{d} \frac{n^2\ n\ \sigma_\varepsilon^2\ \int_0^1 W(r)dW(r) - n\ n^{3/2}\sigma_\varepsilon \int_0^1 W(r)drW_t}{\delta} \end{aligned}$$

The t statistics of μ_β can be written as

$$t_{\mu_\beta} = \frac{\hat{\mu}_\beta - \mu_\beta}{S_{\hat{\mu}_\beta}}$$

where

$$S_{\hat{\mu}_\beta} = \left\{ \sigma_\varepsilon^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} n & \sum_{t=1}^n x_{t-1} \\ \sum_{t=1}^n x_{t-1} & \sum_{t=1}^n x_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}^{1/2}$$

Therefore

$$S_{\hat{\mu}_\beta}^2 = \sigma_\varepsilon^2 \left[\frac{n}{n\sum_{t=1}^n x_{t-1}^2 - (\sum_{t=1}^n x_{t-1})^2} \right]$$

Hence,

$$n^2 S_{\hat{\mu}_\beta}^2 \xrightarrow{d} \left[\frac{n^3\sigma_\varepsilon^2}{\delta} \right] \tag{1.21}$$

Rewriting t_{μ_β} as

$$t_{\mu_\beta} = \frac{n(\hat{\mu_\beta} - \mu_\beta)}{[n^2 S_{\hat{\mu}_\beta}^2]^{1/2}} \tag{1.22}$$

and using (1.21) and (1.22), we have that

$$\begin{aligned} t_{\mu_\beta} &= \frac{n(\hat{\mu_\beta} - \mu_\beta)}{[n^2 S_{\hat{\mu_\beta}}^2]^{1/2}} \\ &\xrightarrow{d} \frac{n\ \sigma_\varepsilon^2\ \int_0^1 W(r)dW(r) - n^{1/2}\sigma_\varepsilon \int_0^1 W(r)drW_t}{[n\sigma_\varepsilon^2]^{1/2}} \end{aligned} \tag{1.23}$$

By Slutsky's Theorem, under $t \to z$ in distribution as $n \to \infty$, where z is a standard normal distribution. Therefore the $\hat{\alpha}$ and $\hat{\mu}_\beta$ are consistent estimator as $n \to \infty$.

1.4.2 RCA(1)-AR(1) model

The RCA(1)-AR(1) model can be described in its simplest from as follows,

$$x_t = \alpha + \mu_\beta x_{t-1} + \rho\varepsilon_{t-1} + e_t \quad, t = 1, 2, \ldots, n \tag{1.24}$$

The RCA(1)-AR(1) from (1.24) can be written in terms of the matrix form as a regression model following

$$\mathbf{Y} = \mathbf{X}\underline{\beta} + \underline{e}$$

where

$$\mathbf{Y} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} z_0 & \hat{\varepsilon}_0 \\ \vdots & \vdots \\ z_{n-1} & \hat{\varepsilon}_{n-1} \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \upsilon \\ \rho \end{bmatrix}, \text{and } \underline{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

where $\hat{\alpha} + \hat{\mu}_\beta x_{t-1} = z_{n-1}$. The limiting normal distribution is focused on the α and μ_β, hence the Brownian motion is showed that

$$n(\hat{\rho} - \rho) \xrightarrow{d} N(0, \sigma_\rho)$$

The LS estimator, the estimator ($\underline{\hat{\rho}}$) is given by $\underline{\hat{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Then

$$\begin{aligned} \underline{\hat{\beta}} &= \begin{bmatrix} \hat{\upsilon} \\ \hat{\rho} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{t=1}^n z_{t-1}^2 & \sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1} \\ \sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1} & \sum_{t=1}^n \hat{\varepsilon}_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^n x_t z_{t-1} \\ \sum_{t=1}^n x_t \hat{\varepsilon}_{t-1} \end{bmatrix} \end{aligned}$$

when following Dickey and Fuller(1979), we use $(\underline{\hat{\beta}} - \underline{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\underline{e}$, so we have

$$\begin{aligned} \begin{bmatrix} \hat{\upsilon} - \upsilon \\ \hat{\rho} - \rho \end{bmatrix} &= \begin{bmatrix} \sum_{t=1}^n z_{t-1}^2 & \sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1} \\ \sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1} & \sum_{t=1}^n \hat{\varepsilon}_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^n e_t z_{t-1} \\ \sum_{t=1}^n e_t \hat{\varepsilon}_{t-1} \end{bmatrix} \\ &= \frac{1}{\gamma} \begin{bmatrix} \sum_{t=1}^n \hat{\varepsilon}_{t-1}^2 & -\sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1} \\ -\sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1} & \sum_{t=1}^n z_{t-1}^2 \end{bmatrix} \begin{bmatrix} \sum_{t=1}^n e_t z_{t-1} \\ \sum_{t=1}^n e_t \hat{\varepsilon}_{t-1} \end{bmatrix} \end{aligned}$$

where $\gamma = \sum_{t=1}^n z_{t-1}^2 \sum_{t=1}^n \hat{\varepsilon}_{t-1}^2 - (\sum_{t=1}^n \hat{\varepsilon}_{t-1} z_{t-1})^2$, then

$$\begin{aligned} n(\hat{\rho} - \rho) &= \frac{\sum_{t=1}^n z_{t-1}^2 \sum_{t=1}^n e_t \hat{\varepsilon}_{t-1} - C \sum_{t=1}^n e_t z_{t-1}}{\gamma} \\ &\xrightarrow{d} \frac{n^3\ (\sigma_e^2)^2 \int_0^1 W_2(r)^2 d(r)\ \int_0^1 W_1(r) dW_1(r) - Cn\ \sigma_e^2\ \int_0^1 W_2(r) dW_2(r)}{\gamma} \end{aligned}$$

where C is constant. The t statistics of ρ can be written as

$$t_\rho = \frac{\hat{\rho} - \rho}{S_{\hat{\rho}}}$$

where

$$S_{\hat{\rho}} = \left\{ \sigma_e^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{n} z_{t-1}^2 & \sum_{t=1}^{n} \hat{\varepsilon}_{t-1} z_{t-1} \\ \sum_{t=1}^{n} \hat{\varepsilon}_{t-1} z_{t-1} & \sum_{t=1}^{n} \hat{\varepsilon}_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}^{1/2}$$

Therefore

$$S_{\hat{\rho}}^2 = \sigma_e^2 \left[\frac{\sum_{t=1}^{n} z_{t-1}^2}{\gamma} \right]$$

Hence,

$$n^2 S_{\hat{\rho}}^2 \xrightarrow{d} n^2 \sigma_e^2 \left[\frac{n^2 \ \sigma_e^2 \int_0^1 W_2(r)^2 d(r)}{\gamma} \right] \tag{1.25}$$

Rewriting t_ρ as

$$t_\rho = \frac{n(\hat{\rho} - \rho)}{[n^2 S_{\hat{\rho}}^2]^{1/2}} \tag{1.26}$$

and using (1.25) and (1.26), we have that

$$\begin{aligned} t_\rho &= \frac{n(\hat{\rho} - \rho)}{[n^2 S_{\hat{\rho}}^2]^{1/2}} \qquad (1.27) \\ &\xrightarrow{d} \frac{n^2 \ \sigma_e^2 \int_0^1 W_2(r)^2 d(r) \ \int_0^1 W_1(r) dW_1(r) - C \ \int_0^1 W_2(r) dW_2(r)}{[n^3 \ \sigma_e^2 \int_0^1 W_2(r)^2 d(r)]^{1/2}} \end{aligned}$$

The $\hat{\rho}$ is consistent estimator as α and μ_β.

1.5 Simulation Study

The simulation study is to compare the performance of RCA(1) model and RCA(1)-AR(1) model. We generated data from RCA(1) that fixed $\alpha = 1$ and $\sigma_\varepsilon^2 = 1$ and used different values from $(\mu_\beta, \sigma_\beta^2)$ with parameter values as the following set.

Case 1 : $\mu_\beta = 0.8$ and $\sigma_\beta^2 = 1$

Case 2 : $\mu_\beta = 1$ and $\sigma_\beta^2 = 0$

Case 3 : $\mu_\beta = -0.995$ and $\sigma_\beta^2 = 0.1$

Case 4 : $\mu_\beta = -0.1$ and $\sigma_\beta^2 = 0.995$

The Figures 2.1 and 2.2 show 100 and 500 sample sizes generated from each of the above cases. Notice Case 1 looks the stationary process, whereas in Case 2 presents nonstationary process, and Case 3 and 4 tends to oscillate around its mean.

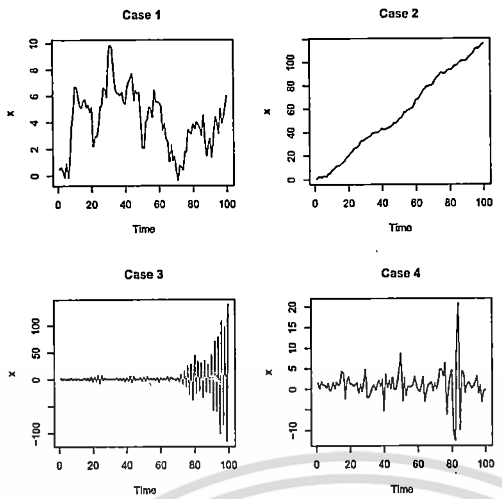

Figure 1.1: The time series plot for generated data (100 sample sizes)

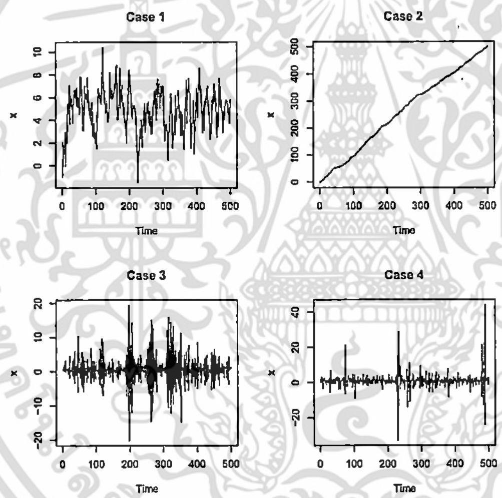

Figure 1.2: The time series plot for generated data (500 sample sizes)

Table 1 provides the average and standard deviation of Mean Square Error (MSE). We also computed the Mean Square Error (MSE) defined as following:

$$MSE = \frac{\sum_{t=1}^{n}(x_t - \hat{x}_t)^2}{n}$$

where x_t denotes the real values and $\hat{x}_t$ denotes the estimated values.

It appears from Table 2.1 that RCA(1) model is well estimated in Case 1 and 2 for all sample sizes. However the RCA(1)-AR(1) model performs

Table 1.1: Average and Standard Deviation of Mean Square Errors with RCA(1) and RCA(1)-AR(1) model (sample size $n = 100, 500$ and 500 replications)

	n=100		n=500	
Case	RCA(1)	RCA(1)-AR(1)	RCA(1)	RCA(1)-AR(1)
Case 1	**0.985**	3.185	**0.996**	3.230
	(0.139)	(0.489)	(0.060)	0.222)
Case 2	**0.970**	3.983	**0.995**	3.230
	(0.136)	(0.542)	(0.063)	(0.222)
Case 3	61.302	**0.258**	3091.059	**8.582**
	(714.202)	(3.053)	(62224.38)	170.104)
Case 4	21.720	**14.986**	21.275	**19.759**
	(91.771)	(42.217)	(42.202)	(57.966)

reasonably well in Case 3 and 4 for all sample sizes. For Case 3 and 4, we see that the volatility is high and cluster around μ_β.

1.6 Application in Real Data

In this section, we will applied the RCA(1), and RCA(1)-AR(1) model using the LS method that we developed in previous section. The data set, we use daily volume of the Thai gold price for selling per 1.5244 grams or called 1 Baht. This data are collected from March 1, 2012 to February 28, 2013 giving a total 288 observations which is collected from http://www.goldpricethai.com/ and shown in Figure 1.3.

For starters, the RCA(1) model following , following

$$x_t = \alpha + \beta_t x_{t-1} + \varepsilon_t$$
$$\beta_t = \mu_\beta + \sigma_\beta u_t$$

where ε_t's are independently and identically distributed with mean 0 and vari-

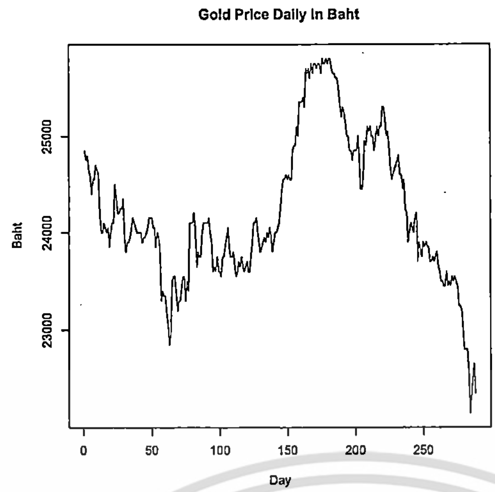

Figure 1.3: The time series plot of Thai gold price from March 1, 2012 to February 28, 2013

ance σ^2.

Next step for parameter estimation, we fitted the RCA(1) model to obtain LS estimator $\hat{\theta} = (\hat{\alpha}, \hat{\mu}_\beta, \hat{\sigma}^2_\beta, \hat{\sigma}^2_\varepsilon)^\top$. We get

$$\hat{\varepsilon}_{t-1} = x_t - \hat{\alpha} - \hat{\mu}_\beta x_{t-1}$$

Finally, the RCA(1)-AR(1) model is fitted by LS method and obtained $\hat{\theta} = (\hat{\alpha}, \hat{\mu}_\beta, \hat{\sigma}^2_\beta, \hat{\sigma}^2_\varepsilon, \hat{\sigma}^2_e, \hat{\rho})^\top$. We have

$$\hat{x}_t = \hat{\alpha} + \hat{\mu}_\beta x_{t-1} + \rho\hat{\varepsilon}_{t-1}$$

Let x_t denote the daily volume of the Thai gold price and $\hat{x}_t$ denote the the daily volume of the Thai gold price estimated from RCA(1) and RCA(1)-AR(1) model.

We use the MSE to compare the performance of RCA(1) model and RCA(1)-AR(1) model. The MSE of the RCA(1) model (MSE = 377,077) is larger than the RCA(1)-AR(1) model (MSE = 1093.138), so the RCA(1)-AR(1) model performs good estimates of Thai gold price.

In Figure 1.4 the bottom panel is the plot of Thai gold price ,the dashed line is RCA(1) model, and the solid line is RCA(1)-AR(1) model. It can be

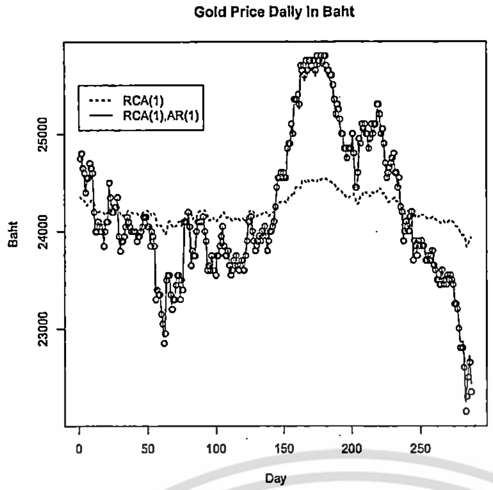

Figure 1.4: The scatter plot of Thai gold price and estimated parameters of RCA(1), and RCA(1)-AR(1) model

seen that the RCA(1)-AR(1) model are competitive model when the data has volatility.

1.7 Conclusion

In this paper, we studied the LS method for RCA(1) model without correlated error and for RCA(1)-AR(1) model with autocorrelated error. We proposed to validate the efficiency of LS estimators by using Brownian motion or Wiener process that asymptotically approach the normal distribution. Through a Monte Carlo simulation study, we evaluated the performance of the LS method and showed its MSEs for different data in 4 cases at the sample size of 100 and 500. For a stationary (Case 1) and a nonstationary data (Case 2), RCA(1) model worked reasonably well for both the 100 and 500 sample size. When the data tends to oscillate around its mean (Case 3 and 4), the RCA(1)-AR(1) model with the LS method worked better.

For real data, we were interested in the power of estimation of each model. Using the Mean Square Error (MSE) as a criterion, we found that the RCA(1)-

AR(1) model performed significantly better than the RCA(1) model.

As part of further work, we are going to study the following aspects of these same 2 models:

- Simultaneous estimation of the RCA-AR model using the Maximum Likelihood (ML) method approach.
- a higher order RCA model and RCA-AR model.

Chapter 2

Bayesian Analysis of RCA-AR Model

2.1 Introduction

In finance it is very common to see that the data are collected in the forms of time series. Researchers have been interested the time series data in modeling the time dependent feature of condition variance or volatility. There are several the volatility models in financial time series, such as the AutoRegressive Conditional Heteroscedastic (ARCH) model introduced by Engle (1982), who is expressed the volatility process in this model. The ARCH model has been extended by Bollerslev (1986) who produced the Generalized ARCH (GARCH). The GARCH model allowed the incorporation of the past observations and past volatility in this model.

An alternative way to model of random coefficients to volatility model is the Conditional Heteroscedastic Autoregressive Moving Average (CHARMA) model by Tsay (1987). The CHARMA model was allowed the coefficient of Autoregressive Moving Average (ARMA) and employed the second-order properties similar to the ARCH model. The special case of CHARMA model was reduced the number of parameter to the Random Coefficient Autoregressive (RCA) model which was studied by Nicholls and Quinn (1982).

The parameter estimation in RCA model, Nicholls and Quinn (1982) employed the least square method and the maximum likelihood method. Wang and Ghosh (2002) used the Bayesian approach to estimate the first order of RCA model. For a RCA model, the estimation using the estimating function technique is equivalent to the weighted least squares estimator which has been discussed by Hwang and Basawa (1998) and Chandra and Taniguchi (2001).

The estimation of the time series model is difficult because of the volatility and the autocorrelated errors, so the autoregressive process has been applied on the error terms. Various autoregressive models have been appeared in literature; Haggan and Ozaki (1981) modelled nonlinear vibration by using and amplitude-dependent autoregressive time series model defined as Exponential Autoregressive (EXPER) model; Tong (1982) introduced the Threshold AutoRegressive (TAR) model in nonlinear time series.

The rest of the article is organized as follows : Section 2 describes the RCA model and the autoregressive process. Section 3 describes the parameter estimation procedure of Bayesian approach. We apply our proposed methods to simulation data and real data using the observation of the Stock Exchange of Thailand (SET) index in Section 4 and 5. Finally, in Section 6, we conclude the results and indicate some directions for further research.

2.2 The RCA model with Autocorrelated Errors

The general class of Random Coefficient Autoregressive model of order p, RCA(p) is given by

$$\begin{aligned} x_t &= \sum_{i=1}^{p} [\alpha_i + \beta_{ti}]\, x_{t-i} + \varepsilon_t \qquad , t = 2, 3, \ldots, n \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \end{aligned} \tag{2.1}$$

where $\underline{\alpha} = (\alpha_1, \ldots, \alpha_p)^\top$ is the $p \times 1$ vector of constant, $\underline{\beta}_t = (\beta_{t1}, \ldots, \beta_{tp})^\top$ are the sequence of iid (independent and identically distributed) random variables, $\underline{\mu}_\beta = (\mu_{\beta 1}, \ldots, \mu_{\beta p})^\top$ and ε_t 's are the iid from a distribution with mean zero and unit variance. Empirically, we can see that the time series data that appeared the autocorrelated in error terms which can be applied by AutoRegressive (AR) process following

$$\varepsilon_t = \sum_{j=1}^{q} \rho_j \varepsilon_{t-j} + e_t \qquad ,t = 2,3,\ldots,n \tag{2.2}$$

The RCA model with autocorrelated errors denoted RCA(p),AR(q) can be rewritten by

$$\begin{aligned} x_t &= \sum_{i=1}^{p} [\alpha_i + \beta_{ti}]\, x_{t-i} + \sum_{j=1}^{q} \rho_j \varepsilon_{t-j} + e_t \qquad ,t = 2,3,\ldots,n \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \end{aligned} \tag{2.3}$$

where $\underline{\rho} = (\rho_1, \ldots, \rho_q)^\top$ is the $q \times 1$ vector of constant.

In this article, we will consider the first order of RCA model with autocorrelated errors denoted by RCA(1),AR(1) that can be written as

$$\begin{aligned} x_t &= (\alpha + \beta_t)x_{t-1} + \rho\varepsilon_{t-1} + e_t \qquad ,t = 2,3,\ldots,n \\ \beta_t &= \mu_\beta + \sigma_\beta u_t \end{aligned} \tag{2.4}$$

where x_t's are iid random variables with mean μ_β, and variance σ_β^2, ε_t's are iid random variables with mean 0 and variance σ^2, and β_t's and ε_t's are independent.

For the parameter estimation of RCA(1),AR(1), it is seen from (2.4) that consisted of the intercept term α, the mean μ_β, variance σ_β^2 of the coefficient β_t, the variance σ^2 of the ε_t, and the coefficient of AR process, or defined as $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \rho)^\top$.

2.3 Parameter Estimation for RCA(1),AR(1)

To perform Bayesian data analysis for the RCA(1),AR(1) model, we propose the three-level hierarchical model. At the first level is the conditional distribution of the data x_t's given the unobserved random variables x_{t-1}, coefficient α, β_t, ρ, and σ^2 . The second level consists of the conditional distribution β_t given the parameter μ_β and σ_β^2 . Finally the last level shows the prior distribution of θ . Consequently, given the sample variables $x_1, x_2, \ldots, x_n$, we are able to express the RCA(1),AR(1) model in the following hierarchical structure,

$$\begin{aligned} x_t|x_{t-1},\alpha,\beta_t,\rho,\sigma^2 &\sim N((\alpha+\beta_t)x_{t-1}+\rho\varepsilon_{t-1},\sigma^2) \\ \beta_t|\mu_\beta,\sigma^2_\beta &\sim N(\mu_\beta,\sigma^2_\beta) \\ (\alpha,\mu_\beta,\sigma^2_\beta,\sigma^2,\rho) &\sim p(\alpha,\mu_\beta,\sigma^2_\beta,\sigma^2,\rho) \end{aligned} \tag{2.5}$$

where $p(\cdot)$ is the prior density of θ which unknown parameter. Following (2.5) , we can express the likelihood function of θ as,

$$\begin{aligned} L(\theta) &= L(\theta|x_1,x_2,\ldots,x_n) \\ &= \prod_{i=1}^{n}\phi\left(x_t;(\alpha+\beta_t)x_{t-1}+\rho\varepsilon_{t-1},\sqrt{(1+\alpha^2+\rho^2)\sigma^2+\sigma^2_\beta x^2_{t-1}}\right) \end{aligned} \tag{2.6}$$

where $\phi(x;\mu,\sigma)$ denotes the density function of a normal distribution with mean μ and standard deviation σ. Therefore, the joint posterior density of the parameters is given by,

$$f(\theta|x_1,x_2,\ldots,x_n) \propto L(\theta|x_1,x_2,\ldots,x_n)p(\theta)$$

where $p(\theta)$ is a prior density of θ . From the hierarchical structure in (2.6), the joint posterior density can be written as

$$f(\mu_\beta|x_1,x_2,\ldots,x_n) \propto \int f(\alpha,\mu_\beta,\sigma^2_\beta,\sigma^2,\rho|x_1,x_2,\ldots,x_n)\,d\alpha\,d\sigma^2_\beta\,d\sigma^2\,d\rho$$

To deal with the complicated likelihood function, we used the so-called Markov Chain Monte Carlo (MCMC) methods to generate samples from the posterior distribution of $\theta = (\alpha,\mu_\beta,\sigma^2_\beta,\sigma^2,\rho)^{\mathrm{T}}$. We will carry out the Gibbs sampler see Gelfand and Smith (1990), a widely used MCMC method, to obtain the parameter from the posterior distribution using the software Win-BUGS.

2.3.1 Prior Distribution

Bayesian analysis combines prior information about model parameters with information from observed data, thereby generating a posterior distribution. Such an analysis requires prior distribution but it's difficult to specify prior about parameters.

For the parameter estimation of in RCA(1),AR(1) model, the prior distribution of $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \rho)^\top$ is denoted that ρ is the uniform distribution, α and μ_β are the normal distribution, and σ_β^2, σ^2 are the inverse gamma distribution.

2.3.2 Posterior Distribution via Gibbs Sampling

To manage Bayesian analysis for RCA(1),AR(1) model, we are interested in the properties of the density of $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \rho)^\top$. Deriving the joint posterior density for θ amounts to integrating out the unobserved coefficients $\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2$, and ρ, which are difficult to perform both analytically and numerically when as the number of observed is equal to the sample sizes.

At this points, MCMC methods consist of algorithms to construct a Markov Chain of the parameters such that its stationary distribution is our distribution of interest, i.e. the posterior distribution of the parameter of interest. That means, under some regularity conditions the realization of this Marcov Chain can be thought of as points sample from the posterior distribution.

We use the Gibbs sampler (Geman, 1984) which is the most popular MCMC method, to obtain dependent samples from the posterior distribution. Specifically, we derive the condition densities, $f(\alpha|\mu_\beta, \sigma_\beta^2, \sigma^2, \rho, \underline{x})$, $f(\mu_\beta|\alpha, \sigma_\beta^2, \sigma^2, \rho, \underline{x})$, $f(\sigma_\beta^2|\alpha, \mu_\beta, \sigma^2, \rho, \underline{x})$, $f(\sigma^2|\alpha, \mu_\beta, \sigma_\beta^2, \rho, \underline{x})$, and $f(\rho|\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \underline{x})$ as the full conditional densities of $\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2$, and ρ, respectively, based on Model 2.5.

The Gibbs sampling algorithm is :

1. Initialize $\alpha^{(0)}, \mu_\beta^{(0)}, \sigma_\beta^{2(0)}, \sigma^{2(0)}$, and $\rho^{(0)}$, for $k = 1, 2, \ldots, m + M$

2. Draw $\alpha^{(k)}$ from $f(\alpha|\mu_\beta^{(k-1)}, \sigma_\beta^{2(k-1)}, \sigma^{2(k-1)}, \rho^{(k-1)}, \underline{x})$
   Draw $\mu_\beta^{(k)}$ from $f(\mu_\beta|\alpha^{(k)}, \sigma_\beta^{2(k-1)}, \sigma^{2(k-1)}, \rho^{(k-1)}, \underline{x})$
   Draw $\sigma_\beta^{2(k)}$ from $f(\sigma_\beta^2|\alpha^{(k)}, \mu_\beta^{(k)}, \sigma^{2(k-1)}, \rho^{(k-1)}, \underline{x})$
   Draw $\sigma^{2(k)}$ from $f(\sigma^2|\alpha^{(k)}, \mu_\beta^{(k)}, \sigma_\beta^{2(k)}, \rho^{(k-1)}, \underline{x})$
   Draw $\rho^{(k)}$ from $f(\rho|\alpha^{(k)}, \mu_\beta^{(k)}, \sigma_\beta^{2(k)}, \sigma^{2(k)}, \underline{x})$

where m is burn-in and M is the number of samples generated after burn-in. Repeating the above sampling steps, we obtain a discrete-time Markov chain $\left\{\left(\alpha^{(k)}, \mu_\beta^{(k)}, \sigma_\beta^{2(k)}, \sigma^{2(k)}, \rho^{(k)}\right); k = 1, 2, \ldots\right\}$ whose stationary distribution is the joint posterior density of the parameters.

We carry out the Gibbs sampling algorithm as proposed above by means of a software package known as WinBUGS. In WinBUGS, a Markov chain of parameters of $\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2$, and ρ are constructed by direct sampling from the corresponding full conditional densities as they are standard distribution. We can perform this procedure using a freeware WinBUGS (Spiegelhalter, 2003).

The MCMC samples of $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \rho)^\top$ obtain via WinBUGS which is computed to approximate posterior summaries of the parameters as the posterior estimation of θ. In particular, we use the posterior mean as point estimates of θ.

2.4 A Simulation Study

The objective of this study is to estimate parameter $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \rho)^\top$ for the Bayesian analysis based on the RCA(1),AR(1) model. The results have showed to compare the average estimators in the sample sizes 100 and 500. In this case, simulation study is divided in three parts.

At the first, we generate data $x_t, t = 1, 2, \ldots, n$ from RCA(1) by taking $\alpha = 0, \sigma^2 = 1$, and $\rho = 0.5$ following ;

$$
\begin{aligned}
x_t &= (\alpha + \beta_t)x_{t-1} + \rho\varepsilon_{t-1} + e_t \\
\beta_t &= \mu_\beta + \sigma_\beta u_t
\end{aligned}
$$

To illustrate the implication of RCA(1) model, Figure 2.1 and 2.2 show generating data of the 100 and 500 sample sizes for each 3 cases as;

1. $\mu_\beta = 0.5, \sigma_\beta^2 = 0.25$
2. $\mu_\beta = 0, \sigma_\beta^2 = 1$
3. $\mu_\beta = 1, \sigma_\beta^2 = 0$

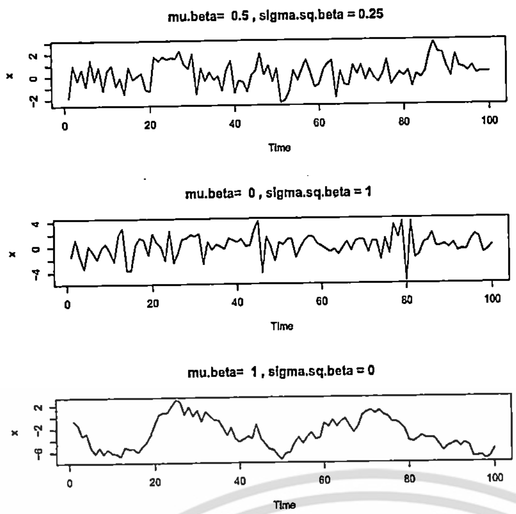

Figure 2.1: The time series plot for generated data of RCA(1) (100 sample sizes)

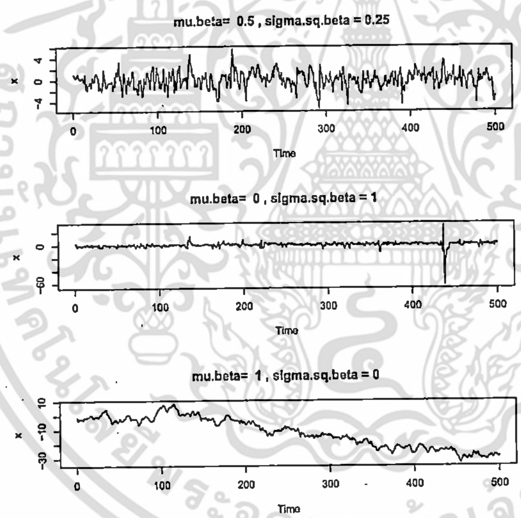

Figure 2.2: The time series plot for generated data of RCA(1) (500 sample sizes)

It should be noted that case 1 is the stationary process, case 2 also tends to oscillate around its mean zero, and case 3 is displayed the nonstationarity or random walk case.

In the second, we obtain the estimator $(\hat{\alpha}, \hat{\mu}_{\beta}, \hat{\sigma}^2_{\beta})$ from MCMC method and we get x_t from

$$\begin{aligned} x_t &= (\hat{\alpha} + \hat{\beta}_t)x_{t-1} + \hat{\varepsilon}_t \quad ,t = 2, 3, \ldots, n \\ \hat{\beta}_t &= \hat{\mu}_\beta + \hat{\sigma}_\beta u_t \end{aligned}$$

then

$$\hat{\varepsilon}_t = x_t - (\hat{\alpha} + \hat{\mu}_\beta)x_{t-1} \quad ,t = 2, 3, \ldots, n$$

Thus the MCMC method is used to estimate parameter $\theta = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma^2, \rho)^\top$ again and we get

$$\begin{aligned} x_t &= (\hat{\alpha} + \hat{\beta}_t)x_{t-1} + \hat{\rho}\hat{\varepsilon}_{t-1} \qquad ,t = 2, 3, \ldots, n \\ \hat{\beta}_t &= \hat{\mu}_\beta + \hat{\sigma}_\beta u_t \end{aligned} \tag{2.7}$$

Finally, in the third part, we simulate data at 500 replications and compute the average of posterior mean. For the Table 1-4, the first column shows the true parameter of simulation, and the second and the third columns represent average and standard deviation of posterior mean. The posterior means for lower and upper bounds of 95% confidence interval are given in next two columns, and the last two columns list the T-statistics and p-value for hypothesis testing of the average of parameters.

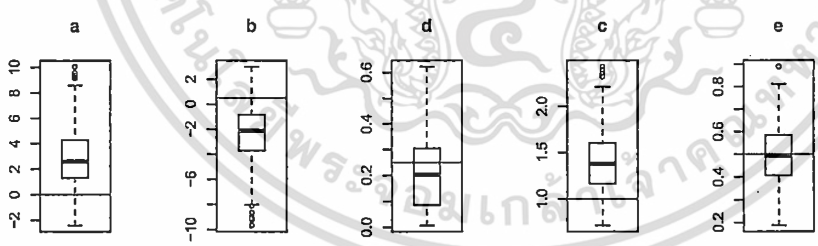

Figure 2.3: The box plot of RCA(1),AR(1) model estimated parameters in case $\alpha = 0$ (a), $\mu_\beta = 0.5$ (b), $\sigma_\beta^2 = 0.25$ (c), $\sigma^2 = 1$ (d), $\rho = 0.5$ (e)

By observing the p-value, the results appear following : from Table1-2, the mean of ρ provide unbiased estimator for all generating models, and the

Table 2.1: Monte Carlo average of Bayesian analysis for different parameters of RCA(1),AR(1) model (sample size $n = 100$ and 500 replications)

true	mean	s.d.	95% Posterior Interval	T-stat	p-value
$\alpha = 0$	2.937	2.174	(2.745,3.128)	30.199	0.000*
$\mu_\beta = 0.5$	-2.449	2.165	(-2.639,-2.259)	-30.451	0.000*
$\sigma_\beta^2 = 0.25$	0.211	0.138	(0.199,0.223)	-6.252	0.000*
$\sigma^2 = 1$	1.390	0.308	(1.363,1.417)	28.349	0.000*
$\rho = 0.5$	0.497	0.126	(0.486,0.508)	-0.438	0.661
$\alpha = 0$	2.319	2.828	(2.070,2.567)	18.335	0.000*
$\mu_\beta = 0$	-2.317	2.827	(-2.566,-2.069)	-18.326	0.000*
$\sigma_\beta^2 = 1$	1.000	0.292	(0.974,1.026)	0.045	0.963
$\sigma^2 = 1$	1.355	0.416	(1.318,1.391)	19.096	0.000*
$\rho = 0.5$	0.502	0.161	(0.487,0.561)	0.299	0.764
$\alpha = 0$	0.536	0.374	(0.503,0.569)	32.080	0.000*
$\mu_\beta = 1$	0.450	0.384	(0.416,0.483)	-31.963	0.000*
$\sigma_\beta^2 = 0$	0.004	0.006	(0.003,0.004)	14.752	0.000*
$\sigma^2 = 1$	1.150	0.193	(1.133,1.167)	17.411	0.000*
$\rho = 0.5$	0.506	0.495	(0.495,0.517)	1.139	0.255

*indicates significance at 5% level

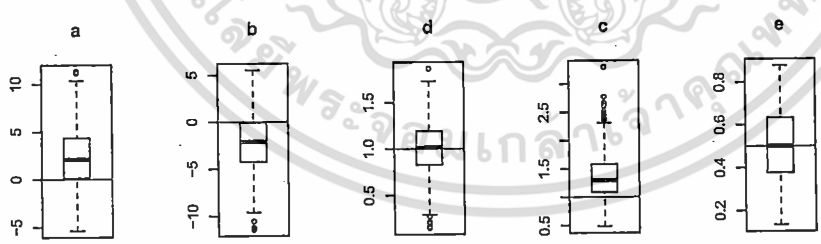

Figure 2.4: The box plot of RCA(1),AR(1) model estimated parameters in case $\alpha = 0$ (a), $\mu_\beta = 0$ (b), $\sigma_\beta^2 = 1$ (c), $\sigma^2 = 1$ (d), $\rho = 0.5$ (e)

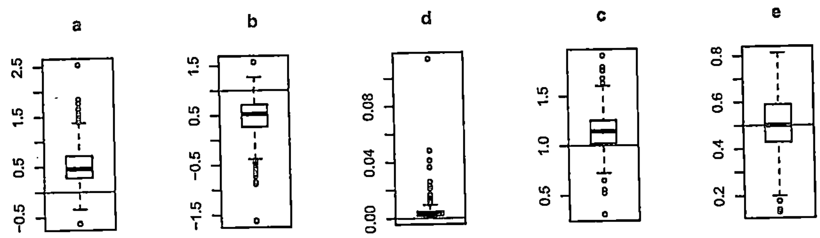

Figure 2.5: The box plot of RCA(1),AR(1) model estimated parameters in case $\alpha = 0$ (a), $\mu_\beta = 1$ (b), $\sigma^2_\beta = 0$ (c), $\sigma^2 = 1$ (d), $\rho = 0.5$ (e)

Table 2.2: Monte Carlo average of Bayesian analysis for different parameters of RCA(1),AR(1) model (sample size $n = 500$ and 500 replications)

true	mean	s.d.	95% Posterior Interval	T-stat	p-value
$\alpha = 0$	1.286	0.925	(1.186,1.349)	30.640	0.000*
$\mu_\beta = 0.5$	-0.771	0.926	(-0.852,-0.689)	-30.673	0.000*
$\sigma^2_\beta = 0.25$	0.245	0.056	(0.240,0.250)	-1.963	0.050
$\sigma^2 = 1$	1.263	0.119	(1.252,1.273)	49.047	0.000*
$\rho = 0.5$	0.497	0.064	(0.492,0.503)	-0.733	0.463
$\alpha = 0$	0.310	1.113	(0.213,0.408)	6.242	0.000*
$\mu_\beta = 0$	-0.315	1.112	(-0.413,-0.218)	-6.347	0.000*
$\sigma^2_\beta = 1$	0.999	0.115	(0.989,1.009)	-0.132	0.894
$\sigma^2 = 1$	1.254	0.146	(1.242,1.267)	38.851	0.000*
$\rho = 0.5$	0.506	0.102	(0.497,0.515)	1.372	0.170
$\alpha = 0$	0.131	0.081	(0.124,0.138)	36.115	0.000*
$\mu_\beta = 1$	0.865	0.084	(0.858,0.872)	-35.469	0.000*
$\sigma^2_\beta = 0$	0.0007	0.0005	(0.0006,0.0008)	28.016	0.000*
$\sigma^2 = 1$	1.135	0.093	(1.127,1.143)	32.510	0.000*
$\rho = 0.5$	0.500	0.055	(0.495,0.505)	0.207	0.835

*indicates significance at 5% level

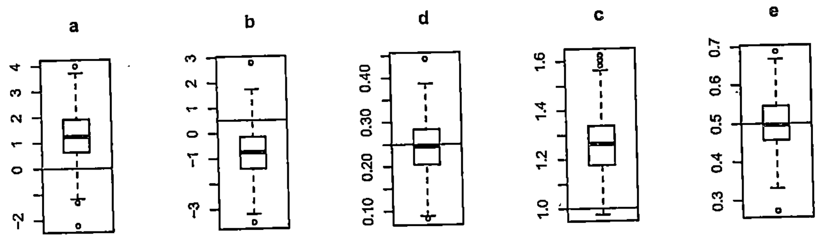

Figure 2.6: The box plot of RCA(1),AR(1) model estimated parameters $\alpha = 0$ (a), $\mu_\beta = 0.5$ (b), $\sigma_\beta^2 = 0.25$ (c), $\sigma^2 = 1$ (d), $\rho = 0.5$ (e)

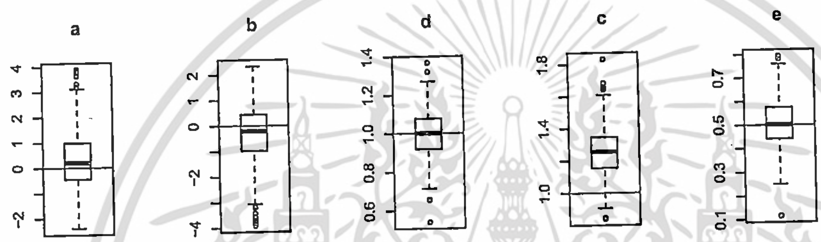

Figure 2.7: The box plot of RCA(1),AR(1) model estimated parameters in case $\alpha = 0$ (a), $\mu_\beta = 0$ (b), $\sigma_\beta^2 = 1$ (c), $\sigma^2 = 1$ (d), $\rho = 0.5$ (e)

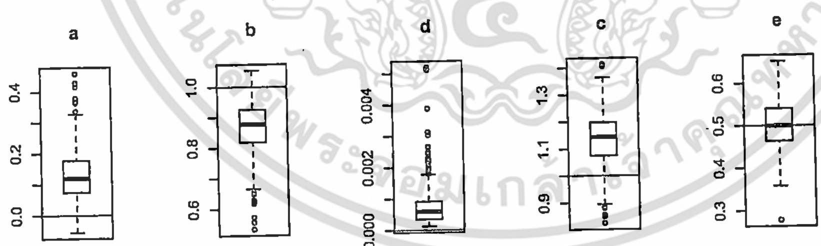

Figure 2.8: The box plot of RCA(1),AR(1) model estimated parameters in case $\alpha = 0$ (a), $\mu_\beta = 1$ (b), $\sigma_\beta^2 = 0$ (c), $\sigma^2 = 1$ (d), $\rho = 0.5$ (e)

mean estimators seem to decrease with increasing sample sizes. However, the estimates of α, μ_β, and σ^2 are not good performance in RCA(1),(AR1) model, but the estimates of σ^2_β appear to be unbiased estimator when the data were generated in case 2.

The boxplot of the coefficient estimates under RCA(1),(AR1) model are presented in Figure 2.3-2.8. From the boxplot is apparent that the variation of σ^2_β is minimum for case 3, but the hypothesis testing is showed the the statistical significance. However, it can be seen that the posterior mean of ρ is closed the true parameter, so ρ is non-significant results.

2.5 Applications in Real Data

In this section, we will consider the application of RCA(1),AR(1) model using the Bayesian approach that we developed in the previous section. The real data, we use the monthly volume of the Stock Exchange of Thailand (SET) index that is started for trading on April 30, 1975. This data are collected from 1976 to 2011 giving a total of 430 observations which is collected from `http://www.set.or.th/th/market/market_statistics.html`.

We estimated an RCA(1),AR(1) model using the prior distribution from Gibbs sampling algorithm. We constructed the Markov chains for initial values and executed the Gibbs sampling iteration 5000 times. For each chains, the first 2000 iterations were discarded and the last 3000 iterations were used to obtain the posterior distributions of the parameter. Table 2.3 shows the posterior mean, standard mean, and 95% posterior interval of the parameter. We also computed the Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) defined as following:

$$MSE = \frac{\sum_{t=1}^{n}(x_t - \hat{x}_t)^2}{n}$$

$$MAPE = \frac{\sum_{t=1}^{n}\left|\frac{x_t - \hat{x}_t}{x_t}\right|}{n} \times 100$$

Table 2.3: Posterior summaries of a RCA(1),AR(1) estimate of SET index

Parameters	Posterior Mean	Standard Deviation	95% Posterior Interval
α	0.05	0.001	(0.049,0.050)
μ_β	0.96	0.004	(0.024,0.964)
σ^2_β	0.007	0.0004	(0.006,0.008)
σ^2	0.168	0.274	(0.002,1.078)
ρ	0.5	0.284	(0.024,0.964)

where x_t denotes the real values and $\hat{x}_t$ denotes the estimated values and showed at Table 2.4.

Table 2.4: Model selection based on MSE, and MAPE of SET index.

Model	MSE	MAPE
RCA(1)	0.0001	0.003
RCA(1),AR(1)	0.00005	0.002

From Table 2.4, the MSE and MAPE of the RCA(1),AR(1) model are smaller than the RCA(1) model, so the RCA(1),AR(1) model performs to estimate SET index better than the RCA(1) model.

In Figure 3.1, the bottom panel is the plot of SET index and the dashed line is the RCA(1),AR(1) estimated model. It is seen that the RCA(1),AR(1) model are competitive model for fitting SET index.

2.6 conclusion

In this article, we have developed the Bayesian estimation for RCA(1),AR(1) model based on stationary and nonstationarity data. Through a Monte Carlo simulation study, we evaluated the performance of Bayesian analysis and

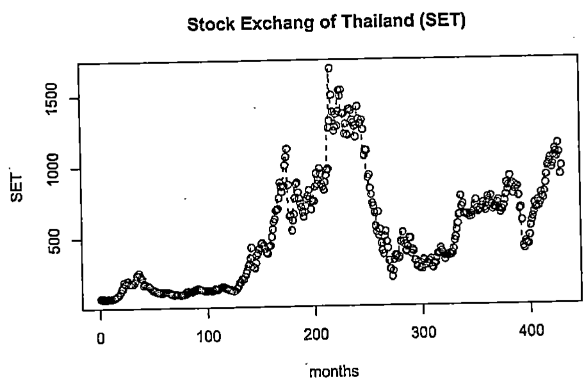

Figure 2.9: The time series plot for SET index and estimated of RCA(1),AR(1) model

showed the coefficient estimator. The coefficient estimator of AR process works reasonably well for all cases and sample sizes. The point estimates for parameters approach their corresponding true values as the sample sizes increase. For hypothesis testing based on the p-value, the estimator of α, μ_β, and σ^2 never support the null hypothesis, but σ_β^2 estimator can apply to estimate the null hypothesis is some cases.

For application in real data, we are also interested the power of estimating by considering the Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE). We can see that the RCA(1),AR(1) model performs better than the RCA(1).

As a part of future research, we suggested to estimated parameter of RCA(1),AR(1) by using the Least-square method and Maximum Likelihood method for the first and higher order.

Chapter 3

RCDR model

3.1 Introduction

Most data are collected in the form of time series that often exhibits nonstationarity and stationary models. The nonstationarity models might be caused by several aspects including changes in trend volatility and random walk. The heteroscedasticity or volatility has been modelled in the literature by various authors, for instance, Anderson and Lund (1977) evaluated risk in finance, Yao and Thong (1994) monitored the reliability of nonlinear prediction. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

There are several volatility models in time series, starting by Engle (1982) who introduced AutoRegressive Conditional Heteroscedastic model (ARCH) which was obtained the predictive variance for U.K. inflation rate. To obtain more flexibility, the ARCH model has been extended by Bollerslev (1986) who produced the Generalized ARCH (GARCH) model. The GARCH model is allowed the past data time series and the past volatility in this model. To overcome some weakness of the GARCH model, Nelson (1991) proposed the Exponential GARCH (EGARCH) model that is used the log condition variance to relax the positiveness constraint of coefficient model. Tsay (1987) proposed the Conditional Heteroscadastic AutoRegressive Moving Average

(CHARMA), in which is not similar to the GARCH model, but these two models possess similar second-order condition properties. The special case of CHARMA model that reduced to the Random Coefficient Autoregressive (RCA) model which was studied by Nicholls and Quinn(1982).

One of the common method is the estimation to practice, through Maximum Likelihood (ML) method that can be developed flexible statistics to a point estimation. The estimation of volatility model, Engle (1982) introduced a class of stochastic process called ARCH model. He derived the likelihood function of these processes and described the maximum likelihood estimators. Silverman (1985) reviewed the penalized maximum likelihood estimation in nonparametric regression and density estimation. Bollerslev (1986) extended the Engle's ARCH processes allowing for a much more flexible lag structure. Bollerslev (1986) derived the conditionals for stationarity of this class of processes and also discussed maximum likelihood estimation of the linear regression model with GARCH errors. Dahlhous and Polonik (2006) presented a methodology for nonparametric Maximum Likelihood estimation of time-varying model.

Our research goal is to present the ML method to model nonstationarity and stationary data. We introduce the new class of Random Coefficient Dynamic Regression (RCDR) model that is extended from the RCA model.

In this paper, the RCDR model is developed from RCA model in Section 3.2. Section 3.3 describes the parameter estimation procedure from the Maximum Likelihood (ML) method of RCDR(1,1) model and shows the properties of ML estimators. Section 3.5 illustrates the results of simulation study and we discuss the results based on AIC and BIC criterion. A conclusion of the results is presented in Section 3.6.

3.2 RCDR Model

In the case of univariate time series data, the RCA model is used the conditional variance to evolve with previous observations denoted RCA(p). The

RCA model is written as

$$\begin{aligned} x_t &= \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \sigma\varepsilon_t \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \end{aligned} \quad (3.1)$$

where

$$\begin{aligned} \underline{\beta}_t &= (\beta_{t1}, \ldots, \beta_{tp})' \\ \underline{\mu}_\beta &= (\mu_{\beta 1}, \ldots, \mu_{\beta p})' \end{aligned}$$

The ε_t and $\underline{\mu}_\beta$ are the sequences of independent of random vectors with mean zero and unit variance.

The RCA model consists of one variable but sometime it is not enough to estimate the coefficient of the time series model so the disadvantage can be improved the model by adding exogenous variables. The time series model can produce the time series dynamic modeling when the observations of time series data have correlated with exogenous variables, the dynamic modeling will help accurate the coefficient model.

Essentially, we will extend the RCA model by adding the exogenous variables of y_t denoted as

$$\begin{aligned} x_t &= \alpha_t + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \sigma\varepsilon_t \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \\ \alpha_t &= \sum_{j=1}^{q} \eta_j y_{t-j} + \varepsilon_t \end{aligned} \quad (3.2)$$

The ε_t and $\underline{\mu}_\beta$ are the sequences of independent of random vectors with mean zero and unit variance, so the model (3.2) is called Random Coefficient Dynamic Regression (RCDR) Model or RCDR(p,q) model.

We will consider the simplified case of RCDR model with $p = q = 1$ and $\sigma = 1$; denoted by the RCDR$(1, 1)$, and we can rewrite as

$$\begin{aligned} x_t &= \alpha_t + \beta_t x_{t-1} + \varepsilon_t \\ \beta_t &= \mu_\beta + \sigma_\beta u_t \\ \alpha_t &= \eta y_{t-1} + \varepsilon_t \end{aligned} \quad (3.3)$$

where β_t's are iid random variables with mean μ_β and variance σ^2_β, ε_t's are iid random variables with mean 0 and variance σ^2, and β_t's and ε_t's are independent.

The parameters of RCDR(1, 1) consist of the intercept term η, the mean μ_β and variance σ^2_β of the coefficient β_t and the variance σ^2 of the ε_t, or defined as $\theta = (\eta, \mu_\beta, \sigma^2_\beta, \sigma^2)'$. In the literature, there is the RCA(1) with the slight modifications to model setup that used the nature of problem at hand might motivate to assume the two random variables β_t and ϵ_t to be correlated, see Hwang and Carlin (1998).

3.3 Parameter Estimation for RCDR(1, 1)

The method of maximum likelihood has been widely used in estimation. For any set of observations, $x_1, \ldots, x_n$, time series or not, the likelihood function $L(\theta)$ is define to be the joint probability density of obtaining the data actually observed. However, it is considered as a function of the unknown parameters in the model with the observed data held fixed.

To estimate parameter of RCDR(1,1) model, we propose the maximum likelihood method to estimate parameter $\theta = (\eta, \mu_\beta, \sigma^2_\beta, \sigma^2)'$. The time series data $\{x_t\}$ and $\{y_t\}$from (3.3) obtain following:

$$E(x_t|\ x_{t-1}) = \alpha_t + \mu_\beta\ x_{t-1}$$

$$Var(x_t|\ x_{t-1}) = (1+\eta^2)\sigma^2 + \sigma^2_\beta x^2_{t-1}$$

The maximum likelihood method consider the likelihood function from (3.3) to estimate μ_β as

$$\begin{aligned} \mathrm{L}(\theta) &= \mathrm{L}(\theta|x_t\ x_{t-1}) = \prod_{t=2}^{n} f(x_t|\ x_{t-1}) \\ &= \left(\frac{1}{2\pi}\right)^{n/2} \prod_{t=2}^{n} \left[(1+\eta^2)\sigma^2 + \sigma^2_\beta x^2_{t-1}\right]^{-1/2} \\ &\quad \exp\left\{-\frac{1}{2}\sum_{t=2}^{n} \frac{(x_t - \alpha_t - \mu_\beta x_{t-1})^2}{(1+\eta^2)\sigma^2 + \sigma^2_\beta x^2_{t-1}}\right\} \end{aligned} \tag{3.4}$$

Constructing the new likelihood function by setting parameter, let $\eta^2 = \omega, \tau = \frac{\sigma_\beta^2}{\sigma^2}$ and substitute $\alpha_t = \eta\ y_{t-1}$, so it show that

$$\begin{aligned} L(\theta) &= \left(\frac{1}{2\pi}\right)^{n/2} \prod_{t=2}^{n} [\sigma^2]^{-1/2} \left[(1+\omega) + \tau x_{t-1}^2\right]^{-1/2} \\ & \exp\left\{-\frac{1}{2}\sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})^2}{\sigma^2[(1+\omega)+\tau x_{t-1}^2]}\right\} \end{aligned} \tag{3.5}$$

The ln likelihood function following;

$$\begin{aligned} \ln L(\theta) &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 \\ & -\frac{1}{2}\sum_{t=2}^{n} \ln\ \left[(1+\omega)+\tau x_{t-1}^2\right] \\ & -\left\{\frac{1}{2}\sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})^2}{\sigma^2[(1+\omega)+\tau x_{t-1}^2]}\right\} \end{aligned} \tag{3.6}$$

The next step is differentiable from (3.6) with respect to μ_β,η, and σ^2

$$\frac{\partial \ln L(\theta)}{\partial \mu_\beta} = \sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})x_{t-1}}{\sigma^2[(1+\omega)+\tau x_{t-1}^2]} \tag{3.7}$$

$$\frac{\partial \ln L(\theta)}{\partial \eta} = \sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})y_{t-1}}{\sigma^2[(1+\omega)+\tau x_{t-1}^2]} \tag{3.8}$$

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} \\ & +\frac{1}{2\sigma^4}\sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})^2}{(1+\omega)+\tau x_{t-1}^2} \end{aligned} \tag{3.9}$$

Now we get

$$\frac{\partial \ln L(\theta)}{\partial(\mu_\beta, \eta, \sigma^2)} = 0$$

We obtain the estimators;

$$\hat{\mu}_\beta = \frac{a_1 - \hat{\eta}\ a_2}{a_3} \tag{3.10}$$

$$\hat{\eta} = \frac{a_4 - \hat{\mu}_\beta\ a_2}{a_5} \tag{3.11}$$

$$\hat{\sigma}^2 = (n)^{-1}\sum_{t=2}^{n} \frac{(x_t - \hat{\eta}\ y_{t-1} - \hat{\mu}_\beta\ x_{t-1})^2}{(1+\omega)+\tau x_{t-1}^2} \tag{3.12}$$

where

$$\begin{aligned}
a_1 &= \sum_{t=2}^{n} \frac{x_t x_{t-1}}{(1+\omega)+\tau x_{t-1}^2}, a_2 = \sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{(1+\omega)+\tau x_{t-1}^2} \\
a_3 &= \sum_{t=2}^{n} \frac{x_{t-1}^2}{(1+\omega)+\tau x_{t-1}^2}, a_4 = \sum_{t=2}^{n} \frac{x_t y_{t-1}}{(1+\omega)+\tau x_{t-1}^2} \\
a_5 &= \sum_{t=2}^{n} \frac{y_{t-1}^2}{(1+\omega)+\tau x_{t-1}^2}
\end{aligned}$$

The ML estimates $\hat{\eta}, \hat{\mu}_\beta, \hat{\sigma}^2$, and $\hat{\sigma}_\beta^2$ can be obtained by calculating $\hat{\tau}$, where $\hat{\tau}$ is the minimizer of the following function of τ,

$$g(\tau) = \ln(\sigma^2) + \sum_{t=2}^{n} \ln((1+\omega) + \tau x_{t-1}^2)$$

That is, we have profiled the log-likelihood as a function of τ only.

The ML estimates $\hat{\eta}, \hat{\mu}_\beta, \hat{\sigma}^2$, and $\hat{\sigma}_\beta^2$ are obtained by

$$\begin{aligned}
\ln L(\hat{\theta}) &= -\inf_{(\theta)} \ln L(\theta) \\
\ell(\hat{\eta}, \hat{\mu}_\beta, \hat{\sigma}^2, \hat{\sigma}_\beta^2) &= -\inf_{(\eta,\mu_\beta,\sigma^2,\sigma_\beta^2)} \ln L(\eta, \mu_\beta, \sigma^2, \sigma_\beta^2)
\end{aligned}$$

$$\hat{\sigma}^2 = (n)^{-1} \sum_{t=2}^{n} \frac{(x_t - \hat{\eta}\ y_{t-1} - \hat{\mu}_\beta x_{t-1})^2}{(1+\hat{\omega}) + \hat{\tau} x_{t-1}^2}$$

$$\hat{\sigma}_\beta^2 = \hat{\sigma}^2 \hat{\tau}$$

and

$$\hat{\omega} = \hat{\eta}^2$$

3.3.1 The Properties of ML Estimators

For the point estimation, we might consider properties as if the sample sizes becomes infinite. In this section, we will look at the properties of ML estimators : consistency and asymptotic efficiency.

1. Consistency

   The consistency of

$$\theta = (\eta_1, \ldots, \eta_q, \mu_{\beta 1}, \ldots, \mu_{\beta p}, \sigma_{\beta 1}^2, \ldots, \sigma_{\beta p}^2, \sigma^2)'$$

will be shown by examining

$$\lim_{n\to\infty} P_\theta(|W_n - \theta| \geq \epsilon) = 0$$

where $W_n = W_n(x_1, \ldots, x_n)$ is a consistent sequence of estimators of parameter θ. Recall that, for an estimator W_n, Chebychev's Inequality states

$$P_\theta(|W_n - \theta| \geq \epsilon) \leq \frac{E_\theta[(W_n - \theta)^2]}{\epsilon^2}$$

For the second term, we have

$$E\left(\frac{(x_t - \sum_{j=1}^{q}\eta_j\, y_{t-j} - \sum_{i=1}^{p}\mu_{\beta i}x_{t-i})^2}{(1+\sum_{j=1}^{q}\omega_j) + \sum_{i=1}^{p}\tau_i x_{t-i}^2}\right)$$

$$\begin{aligned} &= \left|E\frac{(x_t - \sum_{j=1}^{q}\eta_j\, y_{t-j} - \sum_{i=1}^{p}\mu_{\beta i}x_{t-i})^2}{(1+\sum_{j=1}^{q}\omega_j) + \sum_{i=1}^{p}\tau_i x_{t-i}^2}\right| \\ &\leq \left|E\frac{(x_t - \sum_{j=1}^{q}\eta_j\, y_{t-j} - \sum_{i=1}^{p}\mu_{\beta i}x_{t-i})^2}{(1+\sum_{j=1}^{q}\omega_j) + \sum_{i=1}^{p}\tau_i x_{t-i}^2}\right| \\ &< \infty \end{aligned}$$

Therefore, we can conclude that

$$\begin{aligned} \left(E\frac{(x_t - \sum_{j=1}^{q}\eta_j\, y_{t-j} - \sum_{i=1}^{p}\mu_{\beta i}x_{t-i})^2}{(1+\sum_{j=1}^{q}\omega_j) + \sum_{i=1}^{p}\tau_i x_{t-i}^2}\right) &= \frac{\sigma^2}{n} \\ &< \infty \end{aligned}$$

Moreover, (3.13) says that the sample size becomes infinite, the estimators will be arbitrarily close to the parameter with zero in probability.

2. Asymptotic Efficiency Casella and Berger (2002)

Let $x_1, \ldots, x_n$ be iid $f(x|\theta)$, let $\hat{\theta}$ denote the ML estimator of θ, and let W_n be a continuous function of θ.

$$\sqrt{n}\ [W_n - \theta] \to n[0, \upsilon(\theta)],$$

where $\upsilon(\theta)$ is the Cramér-Rao Lower Bound. That is, W_n is a consistent and asymptotically efficient estimator of θ.

Under the property of consistency, the variance of estimator is

$$\begin{aligned} Var(\hat{\theta}) &= E_\theta(W_n - \theta)^2 = \frac{\sigma^2}{n} \\ &\approx \frac{1}{E_\theta\left(\frac{\partial \ln L(\theta)}{\partial \theta}\right)^2} \end{aligned} \quad (3.13)$$

Suppose that

$$\sqrt{n}\left(\frac{W_n - \theta}{\sigma}\right) \to \text{Z in distribution}$$

where $Z \sim \text{Normal}(0,1)$. By applying Slutsky's Theorem we conclude

$$W_n - \theta = \left(\frac{\sigma}{\sqrt{n}}\right)\left(\sqrt{n}\frac{W_n - \theta}{\sigma}\right) \to \lim_{n\to\infty}\left(\frac{\sigma}{\sqrt{n}}\right) Z = 0$$

so $W_n - \theta \to 0$ in distribution. We know that convergence in distribution to a point is equivalent to convergence in probability, so W_n is consistent estimator of θ.

3.4 The Condition of ML Estimators

To use the RCDR(1,1) model to verify that a function of parameters has a local maximum at estimators, it must be shown that the following three conditions hold Casella and Berger (2002).

1. The first-order partial derivatives are 0.

$$\frac{\partial \ln L(\theta)}{\partial(\mu_\beta, \eta, \sigma^2)} = 0$$

2. At least one second-order partial is negative.

   This condition can see after the first-order partial derivatives from (3.7) and (3.8).

$$\begin{aligned} \frac{\partial^2 \ln L(\theta)}{\partial \eta^2} &= -\sum_{t=2}^{n} \frac{y_{t-1}^2}{\sigma^2[(1+\omega) + \tau x_{t-1}^2]} \\ \frac{\partial^2 \ln L(\theta)}{\partial \mu_\beta^2} &= -\sum_{t=2}^{n} \frac{x_{t-1}^2}{\sigma^2[(1+\omega) + \tau x_{t-1}^2]} \end{aligned}$$

3. The Jacobian of the second-order partial derivatives is positive.

For the *ln* likelihood function, the second-order partial derivatives can be written in the symmetric matrix and denoted

$$V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix}$$

where each element V_{ij} of V given by

$$\begin{aligned}
V_{11} &= \frac{\partial^2 \ln L(\theta)}{\partial \eta^2} = -\sum_{t=2}^{n} \frac{y_{t-1}^2}{\lambda_t} \\
V_{12} = V_{21} &= \frac{\partial^2 \ln L(\theta)}{\partial \eta \partial \mu_\beta} = -\sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{\lambda_t} \\
V_{13} = V_{31} &= \frac{\partial^2 \ln L(\theta)}{\partial \eta \partial \sigma^2} = -\sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2} \\
V_{22} &= \frac{\partial^2 \ln L(\theta)}{\partial \mu_\beta^2} = -\sum_{t=2}^{n} \frac{x_{t-1}^2}{\lambda_t} \\
V_{23} = V_{32} &= \frac{\partial^2 \ln L(\theta)}{\partial \mu_\beta \partial \sigma^2} = -\sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2} \\
V_{33} &= \frac{\partial^2 \ln L(\theta)}{\partial (\sigma^2)^2} = \sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n} \sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^3}
\end{aligned}$$

where $\lambda_t = \sigma^2[(1+\omega) + \tau x_{t-1}^2]$ and $u_t = x_t - \eta\, y_{t-1} - \mu_\beta x_{t-1}$.

$$\begin{aligned}
V_{11}V_{22}V_{33} &= \sum_{t=2}^{n} \frac{y_{t-1}^2}{\lambda_t} \sum_{t=2}^{n} \frac{x_{t-1}^2}{\lambda_t} \left[\sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n} \sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^3} \right] \\
V_{12}V_{23}V_{31} &= -\sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{\lambda_t} \sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2} \sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2} \\
V_{13}V_{21}V_{32} &= -\sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{\lambda_t} \sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2} \sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2} \\
V_{31}V_{22}V_{13} &= -\left(\sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2} \right)^2 \sum_{t=2}^{n} \frac{x_{t-1}^2}{\lambda_t} \\
V_{32}V_{23}V_{11} &= -\left(\sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2} \right)^2 \sum_{t=2}^{n} \frac{y_{t-1}^2}{\lambda_t} \\
V_{12}V_{21}V_{33} &= \left(\sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{\lambda_t} \right)^2 \left[\sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n} \sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^3} \right]
\end{aligned}$$

Hence, we can compute the Jacobian by

$$V_{11}V_{22}V_{33} + V_{12}V_{23}V_{31} + V_{13}V_{21}V_{32}$$

$$-V_{31}V_{22}V_{13} - V_{32}V_{23}V_{11} - V_{12}V_{21}V_{33} > 0$$

3.5 A Simulation Study

The simulation study to estimate parameter $\theta = (\eta, \mu_\beta, \sigma_\beta^2, \sigma^2)'$ for the performance of ML method. At the beginning, we generate data $y_t, t = 1, 2, \ldots, n$ from the AR(1) model by taking $\eta = 0.1,\ 0.5$ and 0.9 following;

$$\alpha_t = \eta y_{t-1} + \varepsilon_t : AR(1) \tag{3.14}$$

To illustrate the implication of AR model, Figure 3.1 shows the 100 sample sizes for each 3 coefficients ($\eta = 0.1,\ 0.5$ and 0.9). It should be noted that $\eta = 0.1$ is stationary, $\eta = 0.5$ is weakly stationary, and $\eta = 0.9$ is the nonstationarity case.

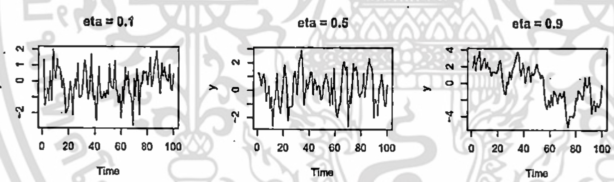

Figure 3.1: The time series plot for data generated from AR(1) of $y_t(\eta = 0.1, 0.5$ and $0.9)$

Next, we consider the RCDR(1, 1) model where y_t are generated from the AR(1). Therefore we obtain the x_t in term of RCDR(1, 1) written as

$$x_t = \eta y_{t-1} + \beta_t x_{t-1} + \varepsilon_t \qquad 3.15)$$

In Figures 3.3-3.4, we present the data generating in 6 cases :

1. $\sigma^2 = 1$, $\mu_\beta = 0.5$ and $\sigma_\beta^2 = 0.25$

2. $\sigma^2 = 1$, $\mu_\beta = 0.995$ and $\sigma_\beta^2 = 0.01$

3. $\sigma^2 = 1$, $\mu_\beta = 0.1$ and $\sigma_\beta^2 = 0.99$

4. $\sigma^2 = 1$, $\mu_\beta = -0.995$ and $\sigma_\beta^2 = 0.01$

5. $\sigma^2 = 1$, $\mu_\beta = -0.1$ and $\sigma_\beta^2 = 0.99$

6. $\sigma^2 = 1$, $\mu_\beta = 0$ and $\sigma_\beta^2 = 1$

It should be noted that Case 2 is the nonstationarity case and the Case 4 tends to be around its mean value of 0 as the stationary process. For Case 1, 3, 5, and 6, we can't define the character of the time series plot when there are slightly different figures that depended on the η under y_t. In order to

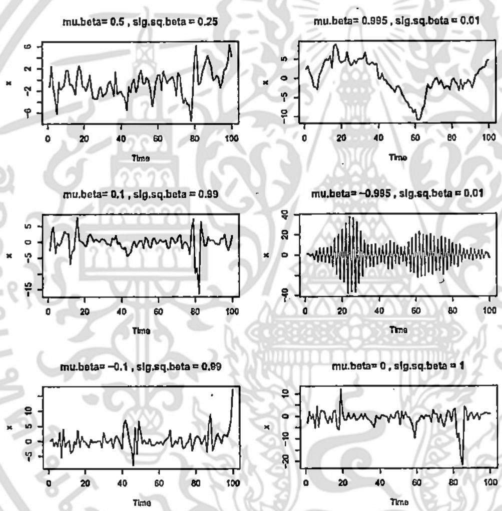

Figure 3.2: The time series plot for data generated from RCDR(1,1) of $x_t(\eta = 0.1)$

assess the model performance, it is customary to use some type of model selection criteria such as Akaike Information Criterion (AIC) introduced by Akaike (1984) and Bayesian Information Criterion (BIC) studied by Schwarz(1978). In our simulation studies we explore the performance of the RCDR(1,1) model

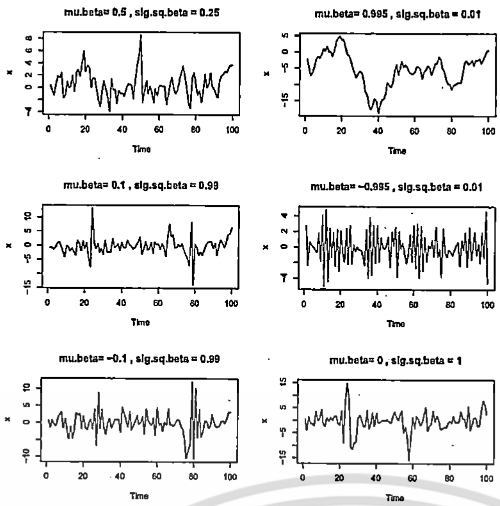

Figure 3.3: The time series plot for data generated from RCDR(1,1) of $x_t(\eta = 0.5)$

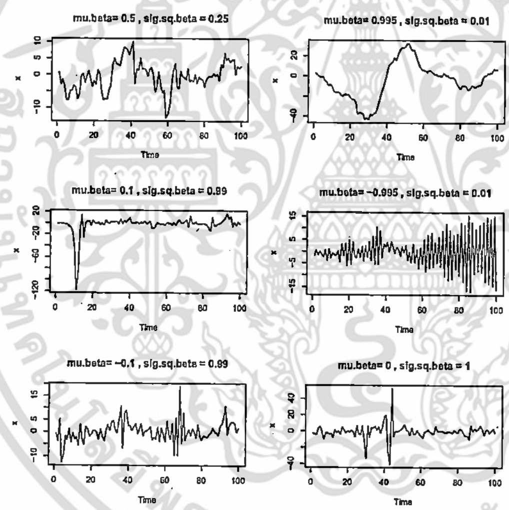

Figure 3.4: The time series plot for data generated from RCDR(1,1) of $x_t(\eta = 0.9)$

in picking up the true models. AIC and BIC are defined as,

$$AIC(\theta) = -2\ln L(\theta) + 2m$$
$$BIC(\theta) = -2\ln L(\theta) + m\ln(n)$$

where $\ln L(\cdot)$ is the log-likelihood function, m is the number of parameters in the model, and n is the number of the sample sizes.

The selection of the chosen model is then made by considering the smallest AIC and BIC in each case.

The results of the simulations were carried out using R program that was used to generate data and performed the parameter values from ML method. We simulated data with the sample sizes $n = 100$ and 500, and repeated the data generation for model fitting 500 times. Tables 3.1- 3.3 show various Monte Carlo (MC) of the estimates obtained by taking $\eta = 0.1, 0.5$ and 0.9 in 6 cases.

The third and the fourth columns of these tables represent the AIC and BIC criterion to perform in picking up the chosen model when the estimators are fitted. From Tables 3.1, it appears that both AIC and BIC are performing reasonably well when the data are generated from true parameter $\sigma^2 = 1, \mu_\beta = 0, \sigma_\beta^2 = 1$ at sample sizes $n = 100$, and $\sigma^2 = 1, \mu_\beta = 0.1, \sigma_\beta^2 = 0.99$ at sample sizes $n = 500$ when $\eta = 0.1$. However, the $\eta = 0.5$, model is good fit at parameter $\sigma^2 = 1, \mu_\beta = 0.1, \sigma_\beta^2 = 0.99$. On the other hand, the $\eta = 0.9$, parameters $\sigma^2 = 1, \mu_\beta = -0.1, \sigma_\beta^2 = 0.99$ prefer RCDR model at sample sizes $n = 100$, but the sample sizes $n = 500$ is fitted in parameter $\sigma^2 = 1, \mu_\beta = -0.995, \sigma_\beta^2 = 0.01$. For each η, the small sample sizes are fitted better than the large sample sizes.

Table 3.1: The average of AIC and BIC for different RCDR models($\eta = 0.1$, sample sizes n = 100, 500 at 500 replications)

Case	Parameters	n	AIC	BIC
Case 1	$\eta = 0.1, \sigma^2 = 1$	n=100	157.963	168.383
	$\mu_\beta = 0.5, \sigma_\beta^2 = 0.25$	n=500	893.249	910.107
Case 2	$\eta = 0.1, \sigma^2 = 1$	n=100	159.326	169.745
	$\mu_\beta = 0.995, \sigma_\beta^2 = 0.01$	n=500	894.414	911.273
Case 3	$\eta = 0.1, \sigma^2 = 1$	n=100	158.258	168.679
	$\mu_\beta = 0.1, \sigma_\beta^2 = 0.99$	n=500	**892.983**	**909.841**
Case 4	$\eta = 0.1, \sigma^2 = 1$	n=100	158.266	168.687
	$\mu_\beta = -0.995, \sigma_\beta^2 = 0.01$	n=500	893.110	909.968
Case 5	$\eta = 0.1, \sigma^2 = 1$	n=100	158.181	168.602
	$\mu_\beta = -0.1, \sigma_\beta^2 = 0.99$	n=500	893.029	909.888
Case 6	$\eta = 0.1, \sigma^2 = 1$	n=100	**157.920**	**168.341**
	$\mu_\beta = 0, \sigma_\beta^2 = 1$	n=500	893.483	990.341

Table 3.2: The average of AIC and BIC for different RCDR models($\eta = 0.5$, sample sizes n = 100, 500 at 500 replications)

Case	Parameters	n	AIC	BIC
Case 1	$\eta = 0.1, \sigma^2 = 1$	n=100	158.078	168.498
	$\mu_\beta = 0.5, \sigma_\beta^2 = 0.25$	n=500	893.443	910.302
Case 2	$\eta = 0.1, \sigma^2 = 1$	n=100	159.341	169.762
	$\mu_\beta = 0.995, \sigma_\beta^2 = 0.01$	n=500	893.975	910.833
Case 3	$\eta = 0.1, \sigma^2 = 1$	n=100	**157.879**	**168.300**
	$\mu_\beta = 0.1, \sigma_\beta^2 = 0.99$	n=500	**893.264**	**910.123**
Case 4	$\eta = 0.1, \sigma^2 = 1$	n=100	158.088	168.509
	$\mu_\beta = -0.995, \sigma_\beta^2 = 0.01$	n=500	893.275	910.133
Case 5	$\eta = 0.1, \sigma^2 = 1$	n=100	157.940	168.361
	$\mu_\beta = -0.1, \sigma_\beta^2 = 0.99$	n=500	893.302	910.161
Case 6	$\eta = 0.1, \sigma^2 = 1$	n=100	158.399	168.820
	$\mu_\beta = 0, \sigma_\beta^2 = 1$	n=500	893.353	910.211

Table 3.3: The average of AIC and BIC for different RCDR models($\eta = 0.9$, sample sizes n = 100, 500 at 500 replications)

Case	Parameters	n	AIC	BIC
Case 1	$\eta = 0.1, \sigma^2 = 1$	n=100	158.674	169.095
	$\mu_\beta = 0.5, \sigma_\beta^2 = 0.25$	n=500	893.654	910.512
Case 2	$\eta = 0.1, \sigma^2 = 1$	n=100	159.376	169.796
	$\mu_\beta = 0.995, \sigma_\beta^2 = 0.01$	n=500	894.817	911.676
Case 3	$\eta = 0.1, \sigma^2 = 1$	n=100	158.571	168.991
	$\mu_\beta = 0.1, \sigma_\beta^2 = 0.99$	n=500	893.640	910.499
Case 4	$\eta = 0.1, \sigma^2 = 1$	n=100	158.513	168.934
	$\mu_\beta = -0.995, \sigma_\beta^2 = 0.01$	n=500	**893.496**	**910.355**
Case 5	$\eta = 0.1, \sigma^2 = 1$	n=100	**158.387**	**168.808**
	$\mu_\beta = -0.1, \sigma_\beta^2 = 0.99$	n=500	893.777	910.635
Case 6	$\eta = 0.1, \sigma^2 = 1$	n=100	158.419	168.840
	$\mu_\beta = 0, \sigma_\beta^2 = 1$	n=500	893.676	910.535

3.6 Conclusion

We have proposed a new estimators for RCDR(1,1) model by using ML method. Through a Monte Carlo simulation study, we evaluated the performance of estimator to fit RCDR(1,1) model. The proposed estimator of stationary and weakly stationary data is a good performance when the exogenous variables are weakly stationary data. However, the estimator of the nonstationarity data is good fit when the exogenous variables is stationary process. Therefore, the proposed estimator $\eta, \sigma^2, \mu_\beta$, and σ_β^2 are based on the correlation between 2 variables, so the correlation is changed while the model fitting is changed too.

References

Akaike, H., 1984. A new look at the statistical model transactions. *IEEE Transactions on Automatic Control*, 19, 716-723.

Anderson, T.G., and Lund, J., 1997. Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics*, 77, 343-377.

Araveeporn, A., 2012. The Least-Squares Criteria of the Random Coefficient Dynamic Regression Model. *Journal of Statistical Theory and Practice*, 6(2), 315-333.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31, 307-327.

Casella, G., and Berger, R. L., 2002. *Statistical Inference*, CA : Duxbury.

Chan, K. S., Tong, H., 1986. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7, 179-190.

Chandra, S. A., Taniguchi M., 2001. Estimating Functions for Nonlinear Time Series Data. *Annals of the Institute of Statistical Mathematics*,53(1), 125-136.

Dahlhous, R., and Polonik, P., 2006. Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary process. *The Annals of Statistics*, 34, 2790-2824.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimates for autoregressive time series with a unit root.

Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007. *Journal of the American Statistical Association* 74, 427-431.

Gallant, A. R., and Tauchen, G., 1997. Estimation of continuous time models for stock returns and interests rates. *Macroeconomic Dynamics*, 1, 135-168.

Gelfand, A., Smith, A.F.M., 1990. Sampling-Based Approaches to Calculating Marginal Densities. *Journal of American Statistical Association*, 85, 398-409.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

Haggan, V., Ozaki, T., 1981. Modeling nonlinear vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, 68, 189-196.

Hwang, S. Y., Basawa, I. V., 1998. Parameter estimation for generalized random coefficient autoregressive process. *Journal of Statistical Planning and Inference*, 68, 323-337.

Hwang, C., and Carlin, B. P., 1998. Parameter estimation for generalized random coefficient autoregressive process. *Journal of Statistical Planning and Inference*, 68, 323-337.

Nicholls, D. F., Quinn, B. G., 1981. The Estimation of Random Coefficient Autoregressive Models, II. *Journal of Time Series Analysis*, 2, 185-203.

Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns : a new approach. *Econometrica*, 59, 347-370.

Nicholls, D. F., Quinn, B. G., 1982. *Random Coefficient Autoregressive Models: An introduction, Lecture Notes in Statistics, 11*, Springer-Verlag, New York.

Phillips, P . C. B., 1987. Time series regression with a unit roots. *Econometrica*, 55, 227-301.

Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.

Silverman, B. W., 1985. Penalized maximum likelihood estimation. *Encyclopaedia of Statistical Sciences*, 6, 664-667.

Spiegelhalter, D. J., Thomas, A., Best, N., Lunn, D., 2001. WinBUGS Version 1.4 User Manual.

Thavaneswaran, A., Abraham, B., 1988. Estimation for Nonlinear Time Series Model using Estimating Function. *Journal of Time Series Analysis*, 9, 99-108.

Tong, H., 1983. *Threshold models in nonlinear time series analysis.* Lecture Notes in Statistics 21, Heidelberg: Springer.

Tsay, R., 1987. Conditional heteroscedastic time series models. *Journal of the American Statistical Association*, 82, 590-604.

Wang, D., Ghosh, S.K., 2002. Bayesian analysis of random coefficient autoregressive models. *Model Assisted Statistics and Applications*, 3(2), 281-295.

Yao, Q., and Tong, H., 1994. Quantifying the influence of initial values on nonlinear prediction. *Journal of the Royal Statistical Society, Series B*, 56, 701-725.