รายงานการวิจัยฉบับสมบูรณ์

การประมาณค่าพารามิเตอร์โดยใช้เทคนิคการปรับให้เรียบด้วยการวิเคราะห์การถดถอยแบบไม่ใช้พารามิเตอร์

The Estimating Parameter by using Smoothing Techniques with Nonparametric Regression Analysis

นางสาวอัชฌา อระวีพร

ได้รับทุนสนับสนุนงานวิจัยจากเงินงบประมาณ ประจำปีงบประมาณ 2560

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Research Title : The Estimating Parameter by using Smoothing Techniques with Nonparametric Regression Analysis

Research : Autcha Araveeporn

Department : Statistics

Faculty : Science

ABSTRACT

This paper is to study the estimating parameter of nonparametric regression model that consisted of independent and dependent variables. The B-spline, smoothing spline, and penalized spline methods in a class of smoothing techniques are considered to estimate the unknown parameter. We also compare these methods by fitting nonparametric regression model on simulation data and real data. The nonlinear model is generated in different 2 models. According to the results, it can be found that the B-spline, smoothing spline, and penalized spline methods have a good performance to fit nonlinear data but the penalized spline method shows the minimum mean square errors. For a real data, we use the data from a light detection and ranging (LIDAR) experiment that contained the range distance travelled before the light as a independent variable and and logratio logarithm of the ratio of received light from two laser sources as a dependent variables. From the mean square errors of fitting data, the penalized spline shows the minimum values as the simulated data.

We proposed a Bayesian analysis for estimating an unknown parameter in a Random Coefficient Autoregressive (RCA) model and its AutoRegressive (AR) process errors. We called this model an RCA model with autocorrelated errors (RCA-AR). A Markov Chain Monte Carlo (MCMC) method was used to generate samples from a posterior distribution which, after having been averaged, gave the estimated value of the unknown parameter. We used a Gibbs sampling algorithm in our MCMC calculation. To compare the performances of the RCA and the RCA-AR models, a simulation was performed with a set of test data and then the mean square errors obtained were used to indicate their performance. The result was that the RCA-AR model worked better than the RCA model in every case. Lastly, we tried both models with real data. They were used to estimate a series of monthly averages of the Stock Exchange of Thailand (SET) index. The result was that the RCA-AR still worked better than the RCA model, similar to the simulation of test data.

Keywords: AutoRegressive process; Bayes Method; Smoothing Techniques;

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

b00268032
RE00010

CONTENTS

King Mongkut's Institute of Technology Ladkrabang

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 1

An Estimating Parameter of Nonparametric Regression Model

1.1 Introduction

In statistical modelling, regression analysis is a statistical process for estimating parameter of the relationships between dependent and independent variables in terms of regression function. However, regression analysis requires an assumption of the underlying regression function. If an inappropriate assumption is used, it is possible to produce misleading conclusion. To overcome this problems, the nonparametric regression is a choice to analyze data when the data are not approached to assumption of regression analysis.

The nonparametric regression is an alternative way for looking at scatter diagram smoothing to depict the relationship between dependent and independent variables. The single independent variable is called scatterplot smoothing that can be used to enhance the visual appearance to help our eyes pick out the trend in the plot.

Smoothing technique is a part of method to estimate unknown parameter (trend or smoothing estimators) of nonparametric regression model. There are many popular smoothing techniques such as B-spline (Eilers and Marx, 1996), smoothing spline (Wahba, 1990, Green and Silverman, 1994), and penalized spline (Ruppert, Wand and Carroll, 2003).

Smoothing Spline (SS) is a technique that estimated the natural polynomial spline by minimizing penalized sum of squares based on smoothing parameter. Penalized Spline (PS) smoother is approximated by minimizing truncated power function on a low rank thin-plate spline depended on smoothing parameter. The concept of B-spline is similar the smoothing spline and penalized spline. This is required the piecewise constant B-spline that could be obtained from truncated counterparts by differencing B-spline function.

In this paper, we consider the nonparametric regression model in Section 2, and use the B-spline, smoothing spline, and penalized spline methods to estimate unknown parameter of nonparametric regression model in Section 3. In the section

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

4 and , we show the estimation of these methods for simulation data and real data. The conclusion are present on section 6.

1.2 The Nonparametric Regression Model

The nonparametric regression model consists of the cubic spline of piecewise polynomials function based on independent variables ($S(x_t)$), error process (ε_t), and dependent variables (y_t) following

$$y_t \quad = \quad S(x_t)+\varepsilon_t \quad , \ t=1,2,3,\ldots,n. \tag{1}$$

The error process is assumed to follow the normal distribution with mean zero and variance one.

1.3 Method of Smoothing Techniques

The following smoothing techniques are shown the process to estimate parameter based on nonparametric regression model.

1.3.1 Smoothing Spline Method

Wahba (1990) defined the natural polynomial spline $S(x_t)=s_K^m(x_t)$ is a real-valued function on $[a,b]$ with the aid of K so-called knots $-\infty \le a < x_1 < x_2 < \ldots < x_K < b \le \infty$. The class of m-order splines with domain $[a,b]$ will be denoted by $W^m[a,b]$.

The natural measure associated with the function $f \in W^m[a,b]$ that used to measure the roughness of curve which is called the quadratic penalty function given by

$$\int_b^a \left\{S^{(m)}(x_t)\right\}^2 dx \tag{2}$$

where $S^{(m)}(x_t)$ is the mth derivative of $S(x)$ with respect to x.

Consider the simple nonparametric regression model, to estimate $\hat{S}(\cdot)$ minimizes $s_K^{(m)}(f)$ over the class of function $S(\cdot)$ following

$$S_K^{(m)}(f)=\min\sum_{t=1}^{n}\left\{y_t-S(x_t)\right\}^2+\lambda\int_a^b\left\{S^{(\cdot)}(x_t)\right\}^2 dx \tag{3}$$

where $\lambda>0$ denotes a smoothing parameter. In this study, we emphasize $m=2$ so-called the natural cubic spline which is commonly considered in the statistical literature (see Green and Silverman, 1994).

The natural cubic spline is given the value and second derivatives at each knots y_t as

$$\underset{\sim}{S} \quad = \quad S(x_t),$$
$$\underset{\sim}{\gamma} \quad = \quad S''(x_t), \qquad t=1,2,3,\ldots,n$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Let $\underset{\sim}{S}$ be the vector $(S_1,\ldots,S_t)^{\mathrm{T}}$ and let $\underset{\sim}{\gamma}$ be the vector $(\gamma_1,\ldots,\gamma_n)^{\mathrm{T}}$.

The condition of natural cubic spline depends on two matrices Q and R below

$$Q=\begin{pmatrix} h_1^{-1} & 0 & \cdots & 0 \\ -h_1^{-1}-h_2^{-1} & h_2^{-1} & \cdots & 0 \\ h_2^{-1} & -h_2^{-1}-h_3^{-1} & \cdots & 0 \\ 0 & h_3^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{n-1}^{-1} \end{pmatrix}_{(n)\times(n-2)},$$

where $h_t = x_{t+1}-x_t$, for $t=1,2,\ldots,n$, then Q is an $n\times(n-2)$ matrix.

R is a symmetric $(n-2)\times(n-2)$ matrix with elements below

$$R=\begin{pmatrix} \frac{1}{3}(h_1+h_3) & \frac{1}{6}h_2 & \cdots & 0 \\ \frac{1}{6}h_2 & \frac{1}{3}(h_2+h_3) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{3}(h_{n-2}+h_{n-1}) \end{pmatrix}_{(n-2)\times(n-2)}.$$

The matrix K can be decomposed by

$$K = QR^{-1}Q^{\mathrm{T}}. \tag{4}$$

The roughness penalty will satisfy

$$\int_a^b \{S''(x_t)\}^2\,dx_t = \gamma^{\mathrm{T}} R\,\gamma = S^{\mathrm{T}} K\,S. \tag{5}$$

To illustrate, it can be written in matrix form introduced by Green and Silverman (1994) as

$$RSS=\sum_{t=2}^{n}\{y_t - S(x_t)\}^2 = (\underset{\sim}{y}-\underset{\sim}{S})^{\mathrm{T}}(\underset{\sim}{y}-\underset{\sim}{S}) \tag{6}$$

,where $\underset{\sim}{y}=(y_1,\ldots,y_n)^{\mathrm{T}}$ and $S=\left(S(x_1),\ldots,S(x_n)\right)^{\mathrm{T}}$.

The roughness penalty term $\int S''^2$ as $S^{\mathrm{T}}K\,S$ in (5) to obtain

$$\begin{aligned} \hat{S}_\lambda(x_t) &= (\underset{\sim}{y}-\underset{\sim}{S})^{\mathrm{T}}(\underset{\sim}{y}-\underset{\sim}{S})+\lambda \underset{\sim}{S}^{\mathrm{T}}K\underset{\sim}{S} \\ &= \underset{\sim}{S}^{\mathrm{T}}(I+\lambda K)\underset{\sim}{S}-2\underset{\sim}{y}^T\underset{\sim}{S}+\underset{\sim}{y}^T\underset{\sim}{y} \end{aligned} \tag{7}$$

since λK is non-negative definite, the matrix $I+\lambda K$ is strictly positive definite. It therefore follows that (7) has a unique minimum, other smoothing spline estimator is obtained by

$$\hat{S}_\lambda(x_t) = (I+\lambda K)^{-1}\underset{\sim}{y}, \tag{8}$$

where I denote the n-dimensional identity matrix.

In this paper, we also select the smoothing parameter using the method of generalized cross-validation (GCV) suggested by Wahba (1976) and Craven and Wahba (1979). In practice, this step can be implemented by using the function of smooth.spline in the software R.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

1.3.2 Penalized Spline Method

Eubank (1988,1999) introduced the regression spline that the local neighborhoods are specified by a group of locations:

$$\tau_0, \tau_1, \tau_2, \ldots, \tau_K, \tau_{K+1} \tag{9}$$

in the range of interval $[a,b]$, where $a = \tau_0 < \tau_1 < \ldots < \tau_K < \tau_{K+1} < b$. These locations are known as knots, and $\tau_r, r = 1,2,\ldots,K$ are called interior knots.

A regression spline can be constructed using the k-th degree truncated power basis or called the B-spline basis with K knots $\tau_1, \tau_2, \ldots, \tau_K$:

$$1, x_t, \ldots, x_t^k, (x_t - \tau_1)_+^k, \ldots, (x_t - \tau_K)_+^k, \tag{10}$$

where w_+^k denotes k-th power of the positive part of w where $w_+ = \max(0,w)$. The first $(k+1)$ basis functions of the truncated power basis (12) are polynomials of degree up to k, and the others are all the truncated power functions of degree k. A regression spline can be expressed as

$$S(x_t) = \sum_{s=0}^{k} \beta_s x_t^s + \sum_{r=1}^{K} \beta_{k+r} (x_t - \tau_r)_+^k \tag{11}$$

where $\beta_0, \beta_1, \ldots, \beta_{k+K}$ are the unknown coefficients to be estimated by a suitable loss minization.

The penalized spline is a method to estimate a unknown smooth function using the truncated power function (Ruppert and Carroll, 2000), and the penalized spline can be expressed as

$$S(x_t) = \sum_{j=0}^{m-1} \alpha_j x_s^j + \sum_{k=1}^{K} \beta_k (x_t - \tau_k)^{2m-1} \tag{12}$$

where $\beta = (\beta_1, \ldots, \beta_K)^T \sim N\left(0, \sigma_\beta^2 \Omega^{-1/2} \left(\Omega^{1/2}\right)^T\right)$, and the (l,k) th entry of Ω is $|\tau_l - \tau_k|^{2m-1}$ and only the coefficient of $|x_t - \tau_k|^{2m-1}$ are penalized so that a reasonably large order K can be used.

In this case, we focus $m = 2$, as the natural cubic spline, or called low-rank thin-plate spline which present of $S(\cdot)$ as

$$S(x_t) = \alpha_0 + \alpha_1 x_t + \sum_{k=1}^{K} \beta_k |x_t - \tau_k|^3 \tag{13}$$

where $\theta = \left(\alpha_0, \alpha_1, \beta_1, \ldots, \beta_K\right)^T$ is the vector of regression coefficients, and $\tau_1 < \tau_2 < \ldots < \tau_K$ are fixed knots. The number of knots, K can be selected using a cross validation method or information theoretic methods (e.g., BIC or AIC).

This class of penalized spline smoothers $\left(\hat{S}(.)\right)$ may also be expressed as

$$\hat{S}(x_t) = C\left(C^T C + \lambda^3 D\right)^{-1} C^T \underset{\sim}{y} \tag{14}$$

where

$$C = \begin{bmatrix} 1 & x_t & |x_t - \tau_k|^3_{1 \le k \le K} \end{bmatrix}_{1 \le t \le n},$$

$$D = \begin{bmatrix} 0_{2\times 2} & 0_{2\times K} \\ 0_{K\times 2} & \left(\Omega_K^{1/2}\right)^T \Omega_K^{1/2} \end{bmatrix},$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

and $\lambda = \sigma_\beta^2 / \sigma_\varepsilon^2$ is a smoothing parameter. The penalized spline smoothers is estimated by using the SemiPar package in the software R.

1.3.3 B-Splines Method

B-splines are very interested as a basic function for univariated independent variable of nonparametric regression function. De Boor (1978) gave an algorithm to compute B-spline of lower degree on piece wise polynomials function.

The m^{th} degree of B-spline function are evaluated from $(m-1)^{\text{th}}$ degree as

$$B_j^m(x) = \frac{x_t - \tau_j}{\tau_{j+m-1} - \tau_j} B_j^{m-1} + \frac{\tau_{j+m} - x_t}{\tau_{j+m} - \tau_{j+1}} B_{j+1}^{m-1}, \tag{15}$$

where basis of order m with knots $\{B_i^m \mid i = M - m + 1, \ldots, M + K\}$, and auxiliary knots τ_j . B-splines base on non-zero over domain spanned by at most $M+1$ knots. In this case, we focus the $m = 4$ or called the cubic B-spline with K knots has basis expansion as $S(x_t) = \sum_{j=1}^{K+4} B_j^4(x_t)\beta_j$.

The nonparametric regression model can written in form of B-splines as

$$y_t = \sum_{j=1}^{K+4} B_j^4(x_t)\beta_j + \varepsilon_t \ , \ t = 1,2,3,\ldots,n. \tag{16}$$

In matrix form, B-splines can be written in form a linear model

$$B = \begin{bmatrix} B_1^4(x_1) & \cdots & B_{K+4}^4(x_1) \\ \vdots & & \vdots \\ B_1^4(x_n) & \cdots & B_{K+4}^4(x_n) \end{bmatrix}, \ \underset{\sim}{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \text{ and } \underset{\sim}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

The B-splines estimators are approximated by least square problems as

$$\underset{\sim}{\hat{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_{K+4} \end{bmatrix} = \left(B^T B\right)^{-1} B^T \underset{\sim}{y} \ .$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

The B-spline and penalties are studied by Eilers and Marx (1996) that advocate the use of the equally spaced knots, instead of the order statistics of the independent variable. The B-spline coefficients can be estimated as

$$\hat{\underset{\sim}{\beta}} = \left(B^T B + \lambda D^T D\right)^{-1} B^T \underset{\sim}{y}, \tag{17}$$

where D is a banded matric which correspond to the difference penalty and denote by

$$D = \begin{bmatrix} 0 & 0 & 0 \dots 0 & 0 \\ -1 & 1 & 0 \dots 0 & 0 \\ 0 & -1 & 1 \dots 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots -1 & 1 \end{bmatrix}.$$

The fitting cubic B-splines are $\hat{S}(x_t) = \sum_{j=1}^{K+4} B_j^4(x_t)\hat{\beta}_j$. The smoothing parameter λ choosing by minimizing the ordinary function or the generalized cross-validation function.

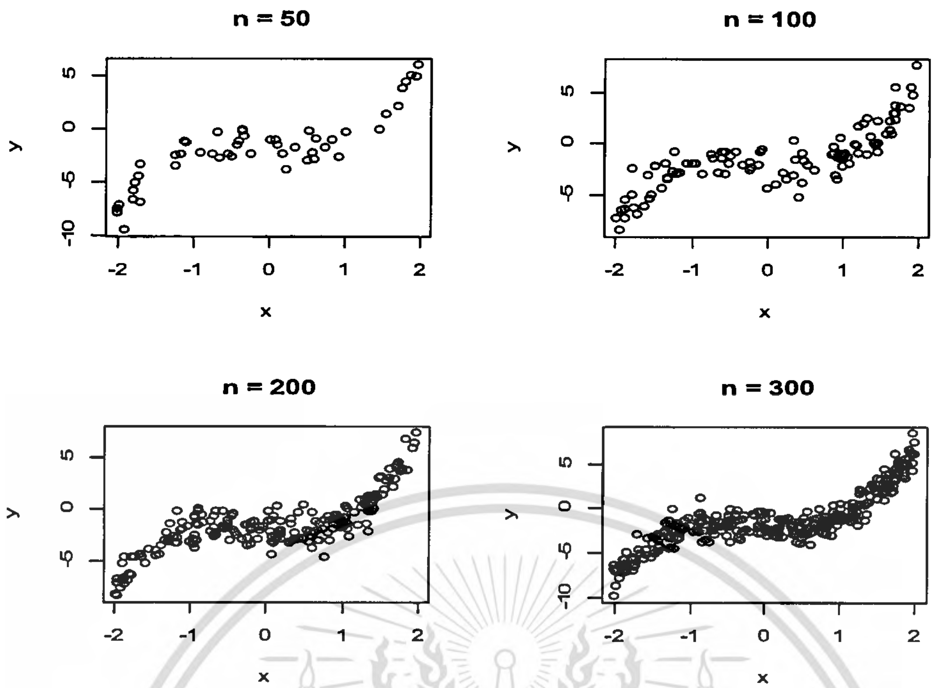
1.4 Simulation Study

The nonlinear data of this study is simulated in 2 models for estimating the performance of smoothing techniques base on independent variables which considered in class of uniform distribution. Figure 1 and 2 show the plot of x_t and y_t on model 1and 2 with 50, 100, 200, 300 sample sizes.

Model 1

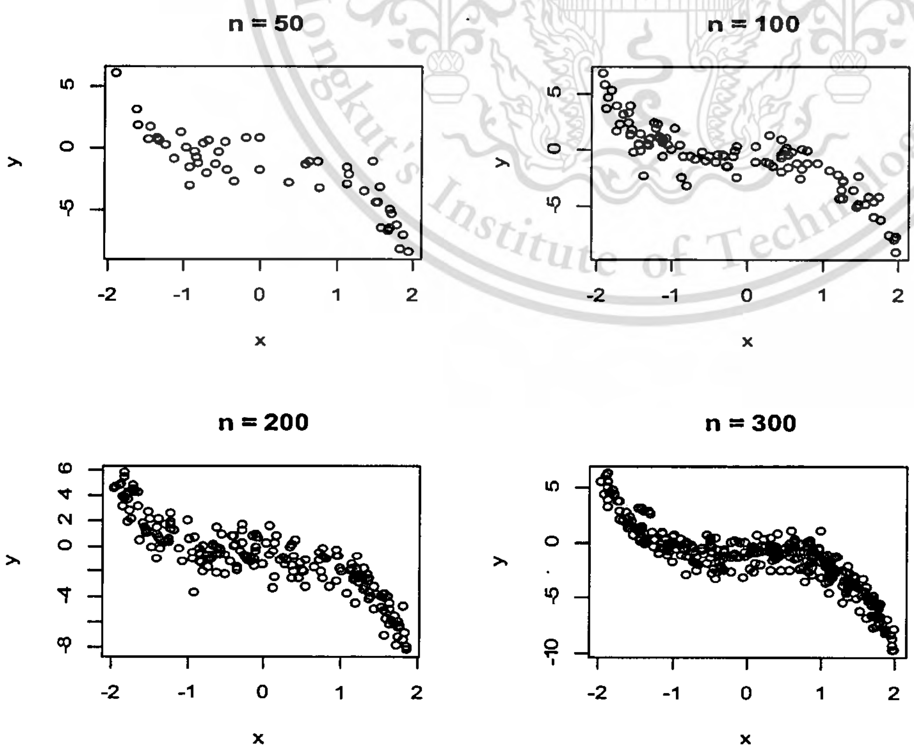
$$S(x_t) = (x_t^3) - \cos(x_t) - \exp\left\{\frac{x_t}{1+|x_t|}\right\}, \quad x_t \sim Uniform(-2,2) \ , t = 1,2,3,\dots,n$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.



Model 2

$$S(x_t) = \sin(x_t) - \exp\left\{\frac{x_t}{1+|x_t|}\right\} - x_t^3 \ , \ x_t \sim Uniform(-2,2) \ , \ t = 1,2,3,\ldots,n$$



This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

The next step, the estimates of $S(x_t)$ or called $\hat{S}(x_t)$ are approximated from SS, PS, and BS that used to compute the bias and MSE of $S(x_t)$ following

$$S_{bias} = \sum_{t=1}^{n}\left(\frac{\hat{S}(x_t)-S(x_t)}{S(x_t)}\right),$$

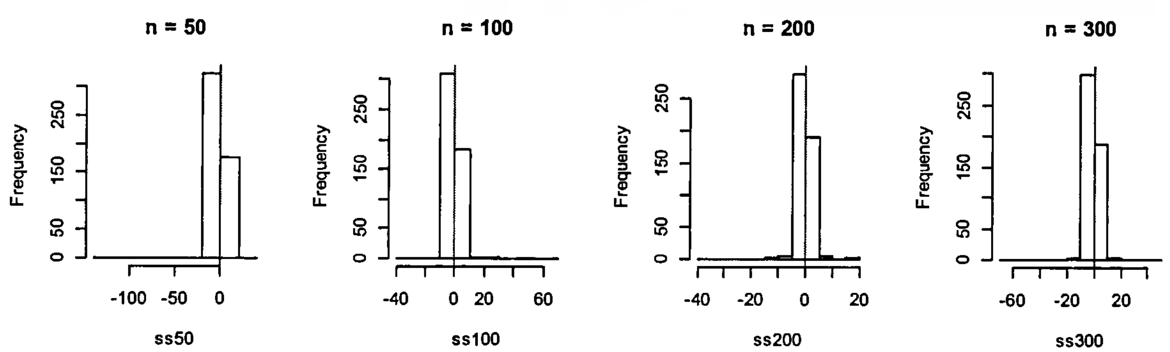
$$MSE = \frac{\sum_{t=1}^{n}\left[\hat{S}(x_t)-S(x_t)\right]^2}{n}.$$

The data is generated and repeated for fitting model 500 times. A t-statistic is adopted to test the mean of bias is equal the zero or called unbiased estimator. Table1 and 2 present the various summary statistics for the smoothing estimator that obtained from 3 methods. The third and the fourth columns of these tables represent the sample mean and standard deviation of biases. The sample mean for the lower and upper bounds of 95% confidence interval are given in next two columns. The last two columns of these tables list the t-statistic, and p-values for hypothesis testing ($H_0 : \mu_{S_{bias}} = 0$ versus $H_0 : \mu_{S_{bias}} \neq 0$). The histogram of the biases on model 1 are presented in Figure 1-3, and model 2 are presented in Figure 4-6

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 1 The summary statistics of simulation studies with model 1 based on smoothing spline (SS), penalized spline (PS), and B-spline (BS).

Sample Sizes	Methods	Mean	S.D.	LCI	UCI	t-stat	p-values
n=50	SS	-0.3344	6.5444	-0.9095	0.2405	-1.1428	0.2537
	PS	-0.3484	6.7454	-0.9411	0.2442	-1.1552	0.2468
	BS	0.3666	6.6935	-0.2214	0.9548	1.2249	0.2212
n=100	SS	0.0719	4.2680	-0.3031	0.4469	0.3767	0.7065
	PS	0.0837	4.2029	-0.2855	0.4530	0.4454	0.6562
	BS	-0.0895	4.2152	-0.4599	0.2807	-0.4751	0.6349
n=200	SS	-0.0954	2.8932	-0.3504	0.1595	-0.7353	0.4625
	PS	-0.0921	2.9123	-0.3480	0.1637	-0.7072	0.4798
	BS	0.0149	3.2685	-0.2299	0.2598	0.1200	0.9045
n=300	SS	-0.1574	4.2673	-0.5347	0.2197	-0.8201	0.4125
	PS	-0.1441	4.1982	-0.5130	0.2247	-0.7676	0.4431
	BS	0.1475	1.8514	-0.2786	0.5738	0.6802	0.4966



This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Figure 1. Histogram of bias for fitting data of smoothing spline method with model 1.

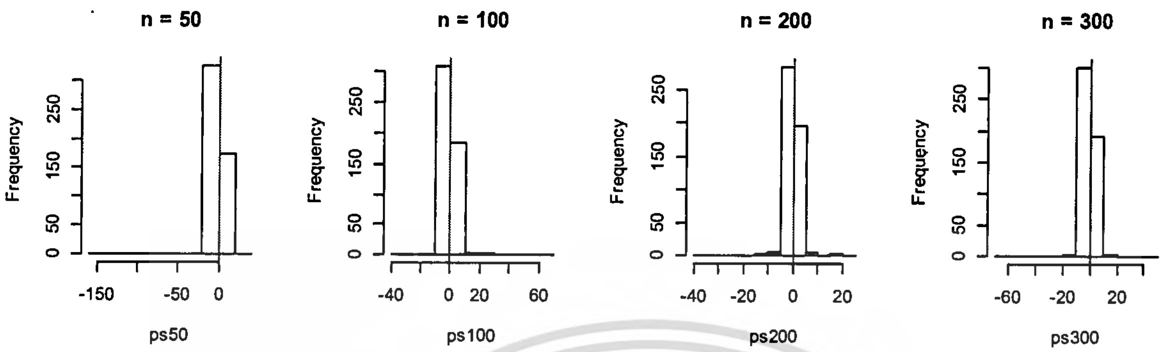


Figure 2. Histogram of bias for fitting data of penalized spline method with model 1.

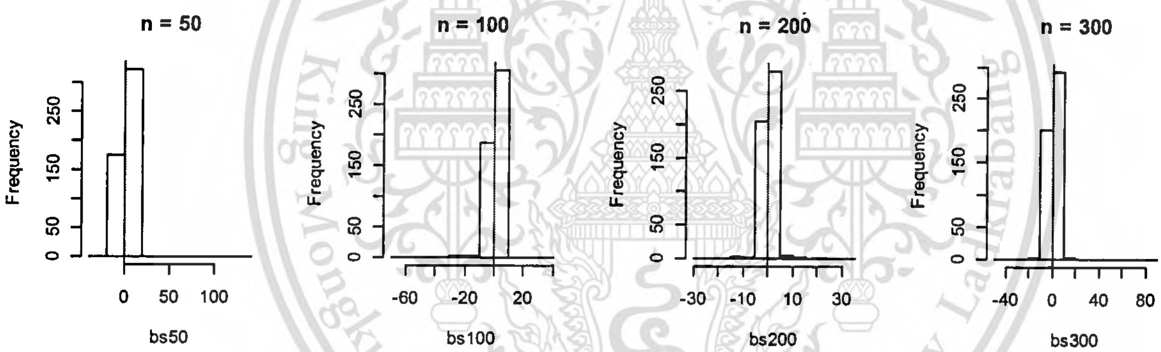


Figure 3. Histogram of bias for fitting data of B-spline method with model 1.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 2 The summary statistics of simulation studies with model 3 based on smoothing spline (SS), penalized spline (PS), and B-spline (BS).

Sample Sizes	Methods	Mean	S.D.	LCI	UCI	t	p-values
n=50	SS	-0.0037	24.2710	-2.1363	2.1288	-0.0034	0.9972
	PS	-0.2582	23.8736	-2.3559	1.9893	-0.2419	0.8089
	BS	-0.3230	27.1906	-2.7121	2.0660	-0.2656	0.7906
n=100	SS	-0.3966	15.8261	-1.7871	0.9939	-0.5603	0.5755
	PS	-0.2489	16.4557	-1.6947	1.1969	-0.3382	0.7353
	BS	0.5461	16.2381	-0.8806	1.9729	0.7520	0.4524
n=200	SS	-0.3704	8.1031	-1.0838	0.3430	-1.0201	0.3082
	PS	-0.3447	7.8716	-1.0363	0.3468	-0.9792	0.3279
	BS	0.4035	8.5116	-0.3443	1.1514	1.0602	0.2896
n=300	SS	0.0411	7.3081	-0.6055	0.6878	0.1250	0.9005
	PS	0.0408	7.0664	-0.5800	0.6617	0.1291	0.8973
	BS	0.0344	7.2205	-0.5999	0.6688	0.1066	0.9151

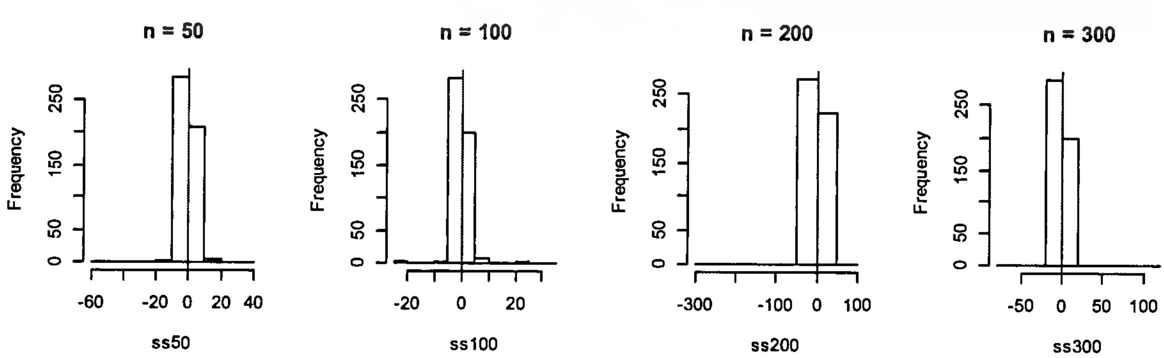


Figure 4. Histogram of bias for fitting data of smoothing spline method with model 3.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

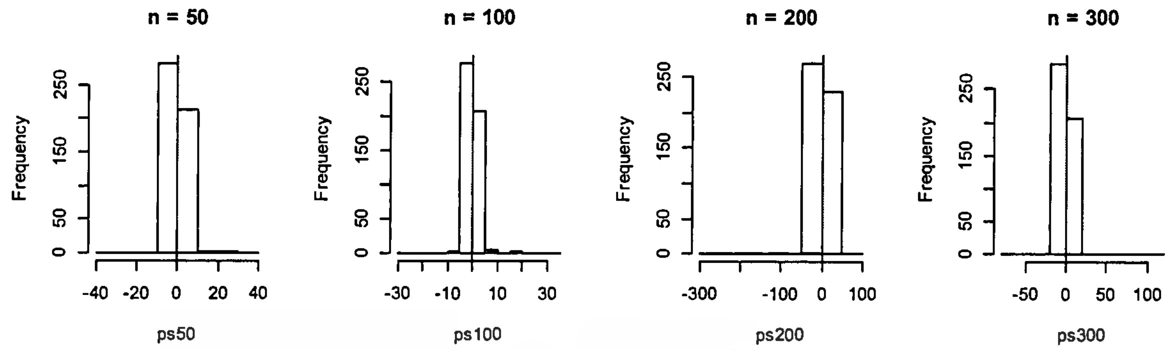


Figure 5. Histogram of bias for fitting data of penalized spline method with model 3.

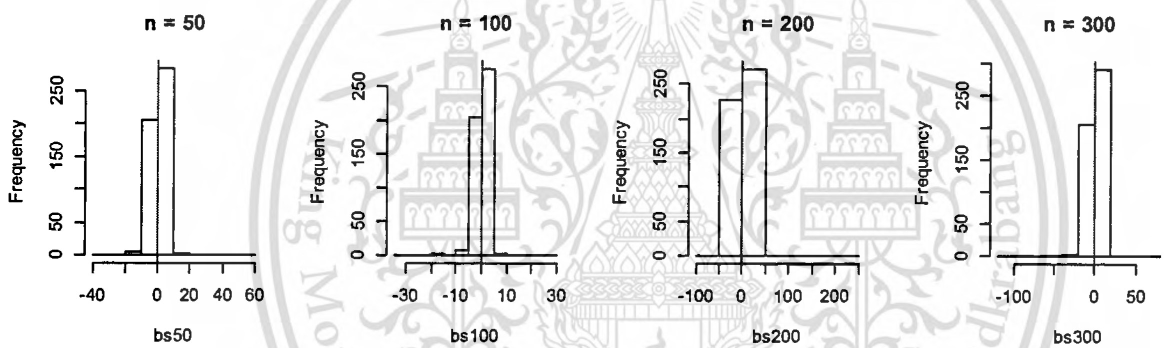


Figure 6. Histogram of bias for fitting data of B-spline method with model 3.

From Table1 and 2, by observing the p-values, the SS, PS, and BS provide asymptotically unbiased estimates for $S(x_t)$ nearly for all sample sizes on 2 models. The p-values form 2 tables are seen that the SS, PS, and BS of smoothing method have a good performance to fit data in a class of nonlinear data. From the histogram is apparent that standard deviation of relative biases increase with increasing sample sizes, so it makes the histogram to be leptokurtic distribution. The average of MSE can answer the final question which smoothing method can be the best estimator. Table 3 shows the average MSE for fitting 500 times on 2 model, and it can see that the PS method shows the minimum of average MSE for all sample sizes and models.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 3 The average MSE of simulation studies with 3 models based on smoothing spline (SS), penalized spline (PS), and B-spline (BS).

Sample sizes	Methods	Model 1	Model2
n=50	SS	0.8225	0.8348
	PS	**0.7712**	**0.7846**
	BS	0.9174	0.9364
n=100	SS	0.9087	0.9029
	PS	**0.8810**	**0.8719**
	BS	0.9632	0.9508
n=200	SS	0.9528	0.9521
	PS	**0.9387**	**0.9439**
	BS	0.9784	0.9840
n=300	SS	0.9666	0.9648
	PS	**0.9640**	**0.9615**
	BS	0.9899	0.9877

1.5 Application of Real Data

In this section, we consider the application of smoothing method based on SS, PS, and BS methods that we developed in previous section. As real data, we use the lidar data frame which consisted 221 observations from a light detection and ranging (LIDAR) experiment. This data frame contains the range distance travelled before the light is reflected back to its source and logratio logarithm of the ratio of received light from two laser sources that shown the plot in Figure 7.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

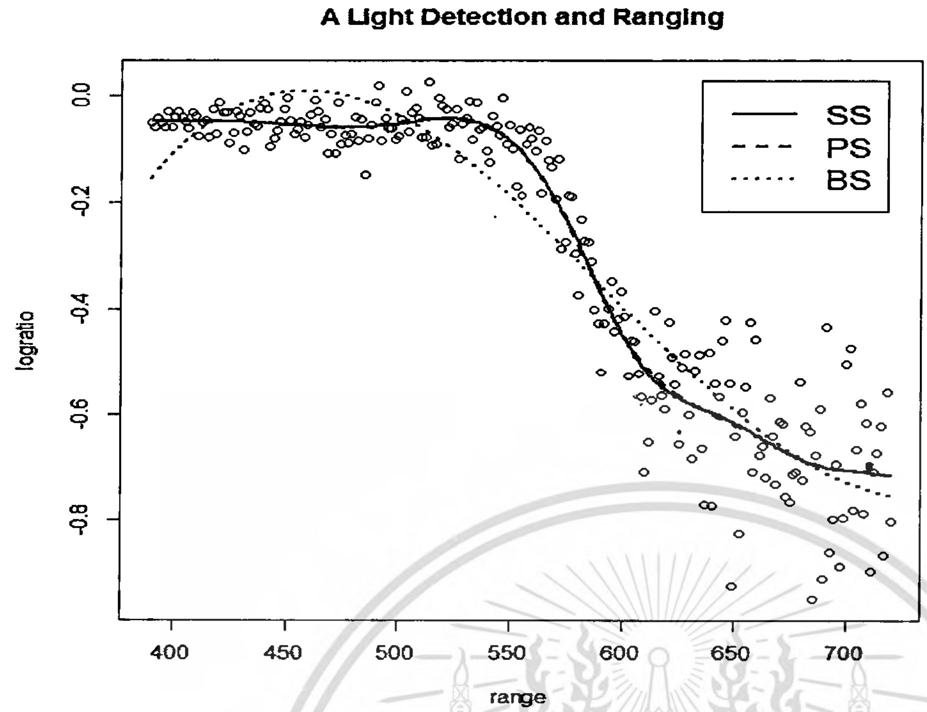


Figure 7. The plot of lidar data frame and model fitting of SS, PS, and BS methods.

After fitting model, the estimating values play on a plot of light detection of ranging. It can be seen that the SS and PS interpolate in mass data more than BS method that followed the MSE values such as SS =0.006016, PS = 0.006010, and BS = 0.009288. The minimum of MSE is PS which is closed the SS as the result on Table 3.

1.6 Conclusion

In this section, we use the smoothing techniques of SS, PS, and BS methods based on nonparametric regression modell. Through a Monte Carlo simulation study, we evaluated the smoothing estimator of SS, PS, and BS methods. For hypothesis testing based on the *p*-value, the fitting values supports the null value, and showed that the smoothing estimators work reasonably well for all methods, but the BS shows the minimum of average MSE.

For application in real data, we are also interested the power of fitting values by considering the mean square error (MSE). We can see that the SS and PS method

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

performs slightly difference in terms of plot data, but the PS shows the best of smoothing techniques as the simulation data.

Reference

Craven, P., Wahba, G., Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. Numerische Mathematik, 31, 377-403, 1979.

De Boor, C., A Practical Guide to Splines. Springer, Berlin, 1978.

Eilers P. H. C. and Marx B.D., Flexible Smoothing with B-splies and Penalties. Statistical Science, 11(2), p89-102, 1996.

Eubank, R.L., Spline Smoothing and Nonparametric Regression. Marcel Dekker, New York,1988.

Eubank, R.L., Nonparametric Regression and Spline Smoothing. Marcel Dekker, New York, 1999.

Green P.J. and Silverman, B. W., *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall, London, 1994.

Ruppert, D., Carroll, R.J., Spatial-adaptive penalties for spline fitting. Australian and New Zealand Journal of Statistics, 42, 205-224, 2000.

Ruppert, D., Wand, M.P., and Carroll. R.J. *Semiparametric Regression*, Cambridge University Press, 2003.

Wahba G., *Spline Models for Observational Data*, SIAM, Philadelphia, PA, 1990.

Wahba, G., A survey of some smoothing problems and the method of generalized cross-validation for solving them. In Proceeding of the Conference on the Application of Statistics , 507-523, 1976.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Chapter 2

Comparing Random Coefficient Autoregressive Model With and Without Autocorrelated Errors

2.1 Introduction

In finance, it is very common to see data in the form of time series that often show in the term of volatility when volume suddenly changes on the economics system, especially if the data are collected over a long period of time. Researchers have long been interested in modeling time series data to extract a time dependent feature of conditional variance or volatility. There are several volatility models for financial time series such as an AutoRegressive Conditional Heteroscedastic (ARCH) model introduced by Engle [1]. After it was introduced, the ARCH model has been extended by Bollerslev [2], creating a Generalized ARCH (GARCH) model. The GARCH model allows incorporation of past observations and past volatility into the model.

There are volatility models that use random coefficients such as the Conditional Heteroscedastic Autoregressive Moving Average (CHARMA) model by Tsay [3]. This CHARMA model uses the coefficients of an Autoregressive Moving Average (ARMA) model and the second-order properties of the ARCH model. A special variant of the CHARMA model with a reduced number of parameters called a Random Coefficient Autoregressive (RCA) model was studied by Nicholls and Quinn[4].

To estimate the parameters in an RCA model, Nicholls and Quinn [4] employed a least square method and the maximum likelihood method. Hwang and Basawa [5] used conditional least square and weighted least square estimators, then Chandra and Taniguchi [6] used a conditional least square and classical moment estimators to estimate unknown parameters in a RCA model. Wang and Ghosh [7] used a Bayesian approach to estimate the first order parameters of the RCA model when the time series data do not have a constant variance and highly volatile periods.

Parameter estimation of time series models is also difficult to estimate and might mislead the results because the error terms are the systematic relationships, the so-called autocorrelated errors. Several autoregressive models have appeared in the literature. Haggan and Ozaki [8] modeled nonlinear vibration by using an amplitude-dependent autoregressive time series model called an Exponential Autoregressive (EXPER) model. Tong [9] introduced a Threshold AutoRegressive (TAR) model for nonlinear time series. This problem is most noticeable in time series data, since the each of the time series data element has a correlation depended on period of time. However, we thought that these errors affected the

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

estimation of unknown parameters, so we applied an autoregressive process on the error terms.

We organize this paper as follows: Section 2 describes the RCA model and the autoregressive process; Section 3 describes the Bayesian approach to estimating parameters; Section 4 shows the prior distribution of the parameters; and Section 5 presents the Gibbs sampling algorithm implementation in the WinBUGS Program. We apply our proposed method to simulated data and real data of monthly averaged SET indexes in Section 6 and 7. Finally, in Section 8, we conclude with the results from simulated and real data.

2.2 The RCA model with Autocorrelated Errors

Consider the following Random Coefficient Autoregressive (RCA) model of order p called RCA(p) that is given by

$$x_t = \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \varepsilon_t \qquad ,t = 2,3,\ldots,n. \tag{2.1}$$

Wang and Ghosh [7] suggested

$$\underline{\beta}_t = \underline{\mu}_\beta + \Sigma_\beta^{1/2}\, \underline{u}_t,$$

where α is the scalar constant , $\underline{\beta}_t = (\beta_{t1},\ldots,\beta_{tp})^{\mathrm{T}}$,and $\underline{\mu}_\beta = (\mu_{\beta 1},\ldots,\mu_{\beta p})^{\mathrm{T}}$. It is assumed that ε_t's and $\underline{u}$'s are the sequence of iid (independent and identically distributed random variables) from a distribution with mean zero and unit variance. The modelling of time series data is a good performance when the errors are shown the randomization, but the most of error terms are correlated between members of a series of number arrange in time. In this case, we focus the errors of time series data which refers to correlation in a class AutoRegressive (AR) process following,

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

where e_t is assumed a white noise that are independently and identically distributed.

The RCA model based on autocorrelated errors was defined by RCA(p)-AR(q), we will write as

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

$$x_t = \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \sum_{j=1}^{q} \rho_j \varepsilon_{t-j} + e_t \qquad ,t = 2,3,\ldots,n. \tag{2.3}$$

$$\underline{\beta}_t = \underline{\mu}_\beta + \Sigma_\beta^{1/2}\, \underline{u}_t,$$

where $\underline{\rho} = (\rho_1,\ldots,\rho_q)^{\mathsf{T}}$ is the $q \times 1$ vector of constant, and e_t's are the errors of AR(q) model.

In this paper, we focus on the simplicity case study of first order on RCA(1) model as the RCA model Wang and Ghosh [7] and conditional heteroscedastic autoregressive model (CHARN) model that studied to model the foreign exchange rate Bossaerts, Hardle, and Hafner [10] .

$$x_t = \alpha + \beta_t x_{t-1} + \varepsilon_t \qquad ,t = 2,3,\ldots,n. \tag{2.4}$$

$$\beta_t = \mu_\beta + \sigma_\beta u_t.$$

Next, we obtain

$$\hat{\varepsilon}_t = x_t - \hat{\alpha} + \hat{\mu}_\beta x_{t-1} \qquad ,t = 2,3,\ldots,n. \tag{2.5}$$

In this case, RCA(1) model with autocorrelated errors in the first order is denoted by RCA(1)-AR(1) model that can be shown as

$$x_t = \hat{\alpha} + \hat{\mu}_\beta x_{t-1} + \rho \hat{\varepsilon}_{t-1} + e_t \qquad ,t = 2,3,\ldots,n. \tag{2.6}$$

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) model, 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_ε^2 of the ε_t, or defined as $\theta_1 = (\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2)^{\mathsf{T}}$. The RCA(1)-AR(1) model gives the parameter as the variance σ_e^2 of the e_t, and the coefficient of AR process, or defined as $\theta_2 = (\rho, \sigma_e^2)^{\mathsf{T}}$.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

2.3 Parameter Estimation

2.3.1 RCA(1) model

In Bayesian estimation for the RCA(1) model, we propose a three-level hierarchical model. At the first level is the conditional distribution of the data x_t's given the observed random variables x_{t-1}, coefficient α, β_t, and σ_ε^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 $\theta_1 = \left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right)^{\mathrm{T}}$. Consequently, given the sample variables $x_1, x_2, \ldots, x_n$, we are able to express the RCA(1) model in the following hierarchical structure,

$$\begin{aligned} x_t \mid x_{t-1}, \alpha, \beta_t, \sigma_\varepsilon^2 &\sim N\left(\alpha + \beta_t x_{t-1}, \sigma_\varepsilon^2\right), \\ \beta_t \mid \mu_\beta, \sigma_\beta^2 &\sim N\left(\mu_\beta, \sigma_\beta^2\right), \\ \left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right) &\sim p\left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right), \end{aligned} \tag{3.1}$$

where $p(\cdot)$ is the prior density of θ_1 which reflects our prior about the unknown parameters. Following (3.1), we can express the likelihood function of θ_1 as,

$$L\left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2 \mid x_1, x_2, \ldots, x_n, \varepsilon_1, \ldots, \varepsilon_n\right)$$

$$= \phi\left(x_1; \alpha, \sigma_\varepsilon\right) \prod_{i=2}^{n} \phi\left(x_t; \alpha + \beta_t x_{t-1}, \sqrt{\sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2}\right), \tag{3.2}$$

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

$$f\left(\theta_1 \mid x_1, x_2, \ldots, x_n\right) \propto L\left(\theta_1 \mid x_1, x_2, \ldots, x_n\right) p\left(\theta_1\right),$$

where $p(\theta)$ is a prior density of θ. From the hierarchical structure in (3.2), the joint posterior density can be written as

$$f\left(\mu_\beta \mid x_1, x_2, \ldots, x_n\right) \propto \int f\left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2 \mid x_1, x_2, \ldots, x_n\right) d\alpha\, d\sigma_\beta^2\, d\sigma_\varepsilon^2.$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

To deal with the complicated likelihood function, we use the Markov Chain Monte Carlo (MCMC) method to generate samples from the posterior distribution of $\theta_1 = \left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right)^{\mathrm{T}}$. We will carry out the Gibbs sampler see [11], a widely used MCMC method, to obtain the parameter from the posterior distribution using the software WinBUGS.

2.3.2 RCA(1)-AR(1) model

At this point, we only know the Bayesian estimation data analysis of RCA(1) model, and we get $\hat{\theta}_1 = \left(\hat{\alpha}, \hat{\mu}_\beta, \hat{\sigma}_\beta^2, \hat{\sigma}_\varepsilon^2\right)^{\mathrm{T}}$. We propose the hierarchical model, for which the conditional distribution of the data x_t's given the unobserved random variables x_{t-1}, coefficient ρ and σ_e^2, and the prior distribution of $\theta_2 = \left(\rho, \sigma_e^2\right)^{\mathrm{T}}$. Therefore the sample variables $x_1, x_2, \ldots, x_n$, we can make the RCA(1)-AR(1) model in the following hierarchical structure ;

$$\begin{aligned} x_t \mid x_{t-1}, \rho, \sigma_e^2 &\sim N\left(\hat{\alpha} + \hat{\mu}_\beta x_{t-1} + \rho\hat{\varepsilon}_{t-1}, \sigma_e^2\right), \\ \left(\rho, \sigma_e^2\right) &\sim p\left(\rho, \sigma_e^2\right), \end{aligned} \tag{3.3}$$

where $p(\cdot)$ is the prior density of θ_2 which unknown parameter. Following (3.3) , we can express the likelihood function of θ_2 as

$$L\left(\rho, \sigma_e^2 \mid x_2, \ldots, x_n, e_2, \ldots, \varepsilon_n\right)$$

$$= \phi\left(x_2; \rho, \sigma_e^2\right) \prod_{i=3}^{n} \phi\left(x_t; \hat{\alpha} + \hat{\mu}_\beta x_{t-1} + \rho\hat{\varepsilon}_{t-1}, \sqrt{\sigma_e^2 + \hat{\mu}_\beta^2 \sigma_\beta^2 + \rho^2 \hat{\sigma}_\varepsilon^2}\right), \tag{3.4}$$

Hence, under certain RCA(1) model, the Markov Chain Monte Carlo (MCMC) methods can be thought of as approximate posterior distribution of $\theta_2 = \left(\rho, \sigma_e^2\right)^{\mathrm{T}}$.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

2.4 Prior Distribution

Bayesian analysis combines prior information about model parameters with information from observed data, thereby generating a posterior distribution. Bayesian analysis requires prior distribution which is difficult to specify analytically for the model considered.

For the parameter estimation of RCA(1) model, the prior distribution of $\theta_1 = \left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right)^{\mathrm{T}}$ are considered to be a continuous random variable in the set of real numbers following the normal distribution, and $\sigma_\beta^2, \sigma_\varepsilon^2$ are denoted in the set of real positive numbers following the inverse gamma distribution.

An RCA(1)-AR(1) model, the prior distribution of $\theta_2 = \left(\rho, \sigma_e^2\right)^{\mathrm{T}}$ is defined that ρ is the set of real number between interval [-1,1] as a uniform distribution, and σ_e^2 is the set of real positive numbers following the inverse gamma distribution.

2.5 Posterior Distribution via Gibbs Sampling

To manage Bayesian analysis for RCA(1) and RCA(1)-AR(1) model, we are interested in the properties of the density of $\theta_1 = \left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right)^{\mathrm{T}}$ and $\theta_2 = \left(\rho, \sigma_e^2\right)^{\mathrm{T}}$. Deriving the joint posterior density for θ_1 amounts to integrate out the unobserved coefficients $\alpha, \mu_\beta, \sigma_\beta^2,$ and σ_ε^2. The joint posterior density for θ_2 is integrated for coefficients ρ and σ_e^2. However it is difficult to perform both analytically and numerically when the observed data are not normally distributed. We can perform the likelihood function to obtain posterior estimator.

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 Markov Chain can be thought of as points sample from the posterior distribution.

We use the Gibbs sampler [12] which is the most popular MCMC method, to obtain dependent samples from the posterior distribution. Specifically, we derive the condition densities each parameters of θ_1 and θ_2.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

2.5.1 RCA(1) model

The condition densities of parameter in RCA(1) model is $f\left(\alpha \mid \mu_\beta,\sigma_\beta^2,\sigma_\varepsilon^2,\underline{x}\right)$, $f\left(\mu_\beta \mid \alpha,\sigma_\beta^2,\sigma_\varepsilon^2,\underline{x}\right)$, $f\left(\sigma_\beta^2 \mid \alpha,\mu_\beta,\sigma_\varepsilon^2,\underline{x}\right)$, and $f\left(\sigma_\varepsilon^2 \mid \alpha,\mu_\beta,\sigma_\beta^2,\underline{x}\right)$, as the full conditional densities of $\alpha,\mu_\beta,\sigma_\beta^2$, and σ_ε^2, respectively, based on model (3.1).

The Gibbs sampling algorithm is :

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

2. Draw $\alpha^{(k)}$ from $f\left(\alpha \mid \mu_\beta^{(k-1)},\sigma_\beta^{2(k-1)},\sigma_\varepsilon^{2(k-1)},\underline{x}\right)$,

   Draw $\mu_\beta^{(k)}$ from $f\left(\mu_\beta \mid \alpha^{(k)},\sigma_\beta^{2(k-1)},\sigma_\varepsilon^{2(k-1)},\underline{x}\right)$,

   Draw $\sigma_\beta^{2(k)}$ from $f\left(\sigma_\beta^2 \mid \alpha^{(k)},\mu_\beta^{(k)},\sigma_\varepsilon^{2(k-1)},\underline{x}\right)$,

   Draw $\sigma_\varepsilon^{2(k)}$ from $f\left(\sigma_\varepsilon^2 \mid \alpha^{(k)},\mu_\beta^{(k)},\sigma_\beta^{2(k)},\underline{x}\right)$,

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 $\left\{\left(\alpha^{(k)},\mu_\beta^{(k)},\sigma_\beta^{2(k)},\sigma_\varepsilon^{2(k)}\right);k=1,2,\ldots\right\}$ whose stationary distribution is the joint posterior density of the parameters.

Gibbs sampling algorithm is proposed above using a software package known as WinBUGS. In WinBUGS, a Markov chain of parameters of $\alpha,\mu_\beta,\sigma_\beta^2$, and σ_ε^2, are constructed by computing the mean sampling from the joint posterior density as standard distribution. We can perform this procedure using a freeware WinBUGS [13].

The MCMC samples of $\theta_1=\left(\alpha,\mu_\beta,\sigma_\beta^2,\sigma_\varepsilon^2\right)^{\mathrm{T}}$ 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.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

2.5.2 RCA(1)-AR(1) model

The condition densities of parameter in RCA(1)-AR(1) model is $f\left(\rho \mid \sigma_e^2, \underline{x}\right)$, and $f\left(\sigma_e^2 \mid \rho, \underline{x}\right)$, as the full conditional densities of ρ and σ_e^2 based on model (3.3).

The Gibbs sampling algorithm is :

1. Initialize $\rho^{(0)}$, and $\sigma_e^{2(0)}$, for $k = 1, 2, \ldots, m + M$

2. Draw $\rho^{(k)}$ from $f\left(\rho \mid \sigma_e^{2(k-1)}, \underline{x}\right)$,

   Draw $\sigma_e^{2(k)}$ from $f\left(\sigma_e^2 \mid \rho^{(k)}, \underline{x}\right)$,

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 $\left\{\left(\rho^{(k)}, \sigma_e^{2(k)}\right); k = 1, 2, \ldots\right\}$ whose stationary distribution is the joint posterior density of the parameters.

The Gibbs sampling algorithm is performed via WinBUGS, then MCMC sample of $\theta_2 = \left(\rho, \sigma_e^2\right)$ are obtained by computing mean of the joint posterior summaries.

2.6 A Simulation Study

The objective of this study is to estimate parameter $\theta_1 = \left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2\right)^{\mathrm{T}}$ from RCA(1) and $\theta_2 = \left(\rho, \sigma_e^2\right)^{\mathrm{T}}$ from RCA(1)-AR(1) by using the Bayesian analysis. The results have been shown to compare the average estimators in the sample sizes 100 and 500. One of the most common measures is Mean Square Error (MSE), which is evaluated the difference between the estimated values and the real values. MSE measures the average of the square error, while the error becomes to be an amount by which the MSE can be compared to the quality estimated. We also computed the MSE as the criterion defined following:

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

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

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

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)-AR(1) by taking $\alpha = 0, \sigma_\varepsilon^2 = 1, \sigma_e^2 = 1,$ and $\rho =$ 0.1, 0.5 ,and 0.9 following

$$x_t = \alpha + \mu_\beta x_{t-1} + \rho\varepsilon_{t-1} + e_t.$$

To illustrate the implication of RCA(1)-AR(1) model, Figure 1 - 3 show generating data of 500 sample sizes and $\rho = 0.1$, 0.5 ,and 0.9 for each 3 cases as;

1. $\mu_\beta = 0.8, \sigma_\beta^2 = 0$
2. $\mu_\beta = 0.995, \sigma_\beta^2 = 0.1$
3. $\mu_\beta = 0.5, \sigma_\beta^2 = 0.25.$

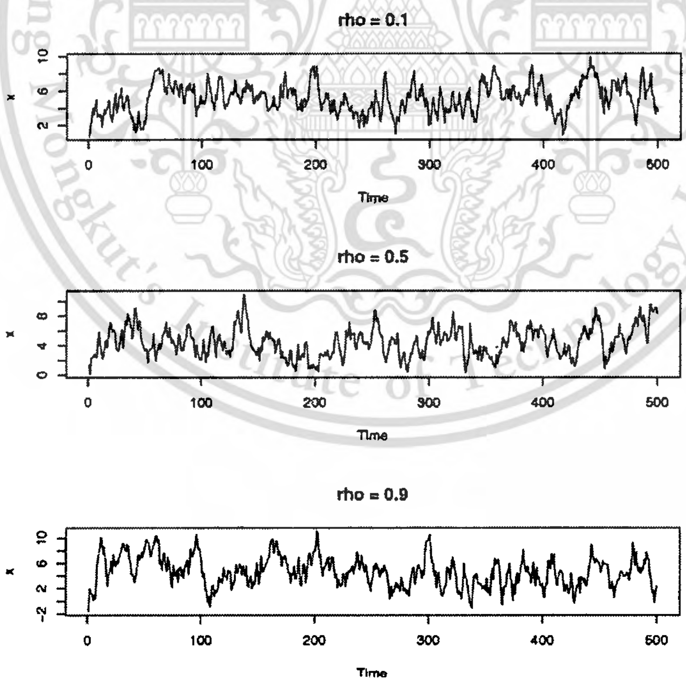


Figure 1. The time series plot for generated data of RCA(1)-AR(1) by fixing $\alpha = 0, \sigma_\varepsilon^2 = 1, \sigma_e^2 = 1, \mu_\beta = 0.8, \sigma_\beta^2 = 0,$ and $\rho = 0.1, 0.5, 0.9$. (500 sample sizes)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

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

In the second, we obtain the estimator $\hat{\theta}_1 = \left(\hat{\alpha}, \hat{\mu}_\beta, \hat{\sigma}_\beta^2, \hat{\sigma}_\varepsilon^2\right)^{\mathsf{T}}$ from MCMC method and we get x_t from

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

Let

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

the MCMC method is used to estimate parameter $\hat{\theta}_2 = \left(\hat{\rho}, \hat{\sigma}_e^2\right)^{\mathsf{T}}$. It's become

$$x_t = \hat{\alpha} + \hat{\mu}_\beta x_{t-1} + \hat{\rho}\hat{\varepsilon}_{t-1} \qquad ,t = 2,3,\ldots,n. \tag{6.1}$$

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

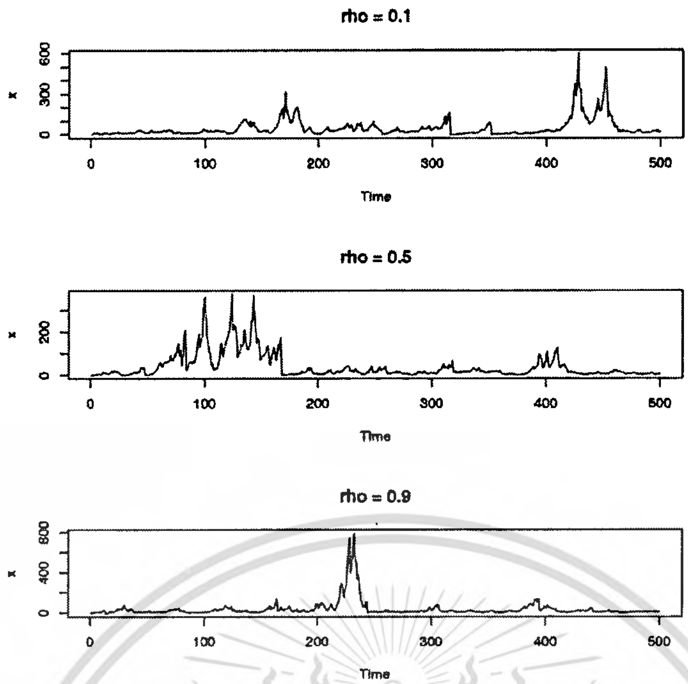


Figure 2. The time series plot for generated data of RCA(1)-AR(1) by fixing $\alpha = 0, \sigma_{\varepsilon}^2 = 1, \sigma_e^2 = 1, \mu_{\beta} = 0.995,\ \sigma_{\beta}^2 = 0.1$, and $\rho = 0.1, 0.5, 0.9$. (500 sample sizes)

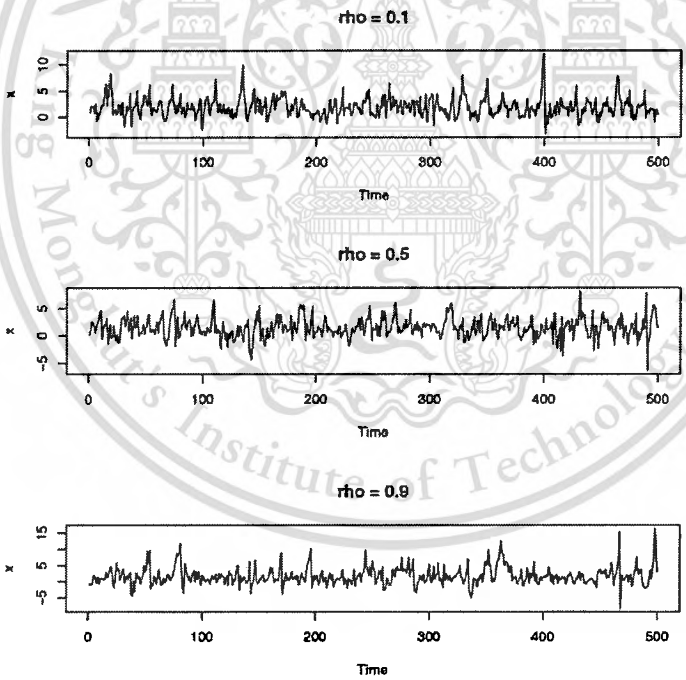


Figure 3. The time series plot for generated data of RCA(1)-AR(1) by fixing $\alpha = 0, \sigma_{\varepsilon}^2 = 1, \sigma_e^2 = 1, \mu_{\beta} = 0.5, \sigma_{\beta}^2 = 0.25$, and $\rho = 0.1, 0.5, 0.9$. (500 sample sizes)

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Finally, in the third part, we simulate data at 500 replications from RCA(1)-AR(1) model. We also compute the Monte Carlo averages and standard deviation of the posterior mean for each parameter based on sample size such as $n = 100$ and $n = 500$. To see this we computed the Average of MSE (AMSE) for comparing the effective estimation between RCA(1) and RCA(1)-AR(1) models. If the AMSEs have a lower value, these models will fit to estimation.

From Table 1-3, we also list the average standard deviation of the posterior mean for fitting RCA(1) and RCA(1)-AR(1) models. The results appear that the average of posterior mean seems to decrease with increasing sample sizes. For the AMSE, the RCA(1)-AR(1) model performs better than RCA(1) model in all cases, since the correlation error process occurs on the generating data, so the correlation is adjusted by RCA(1)-AR(1) model.

King Mongkut's Institute of Technology Ladkrabang

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 1. Monte Carlo average and standard deviation of posterior mean, by fixing $\alpha=0, \sigma_{\varepsilon}^2=1, \sigma_e^2=1, \mu_\beta=0.8, \sigma_\beta^2=0$, and $\rho=0.1, 0.5, 0.9$ of RCA(1) and RCA(1)-AR(1) models

	$n=100$		$n=500$	
true	RCA(1)	RCA(1)-AR(1)	RCA(1)	RCA(1)-AR(1)
$\alpha=1$	1.02(0.44)	1.02(0.44)	1.01(0.29)	1.01(0.29)
$\mu_\beta=0.8$	0.04(0.20)	0.04(0.20)	0.08(0.17)	0.08(0.17)
$\sigma_\beta^2=0$	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
$\sigma_\varepsilon^2=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.1$	-	0.03(0.02)	-	0.01(0.01)
$\sigma_e^2=1$	-	0.99(0.00)	-	0.99(0.00)
AMSE	21.07	19.83	28.51	27.13
$\alpha=1$	1.10(0.31)	1.01(0.31)	1.03(0.14)	1.03(0.14)
$\mu_\beta=0.8$	0.80(0.06)	0.80(0.06)	0.04(0.02)	0.04(0.02)
$\sigma_\beta^2=0$	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
$\sigma_\varepsilon^2=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.5$	-	0.08(0.03)	-	0.03(0.14)
$\sigma_e^2=1$	-	1.00(0.00)	-	1.00(0.00)
AMSE	1.24	1.04	1.27	1.18
$\alpha=1$	1.11(0.34)	1.11(0.34)	1.03(0.15)	1.03(0.15)
$\mu_\beta=0.8$	0.80(0.06)	0.80(0.06)	0.82(0.02)	0.82(0.02)
$\sigma_\beta^2=0$	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
$\sigma_\varepsilon^2=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.9$	-	0.08(0.03)	-	0.03(0.01)
$\sigma_e^2=1$	-	1.00(0.00)	-	1.00(0.00)
AMSE	1.79	1.51	1.82	1.69

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 2. Monte Carlo average and standard deviation of posterior mean, by fixing $\alpha=0,\sigma_{\varepsilon}^{2}=1,\sigma_{e}^{2}=1,\mu_{\beta}=0.995,\ \sigma_{\beta}^{2}=0.1,$ and $\rho=0.1,0.5,0.9$ of RCA(1) and RCA(1)-AR(1) models

	$n=100$		$n=500$	
true	RCA(1)	RCA(1)-AR(1)	RCA(1)	RCA(1)-AR(1)
$\alpha=1$	1.12(0.29)	1.12(0.29)	1.03(0.14)	1.03(0.14)
$\mu_{\beta}=0.995$	0.76(0.05)	0.76(0.05)	0.79(0.02)	0.79(0.02)
$\sigma_{\beta}^{2}\doteq 0.1$	0.01(0.00)	0.01(0.00)	0.01(0.00)	0.01(0.00)
$\sigma_{\varepsilon}^{2}=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.1$	-	0.09(0.03)	-	0.09(0.01)
$\sigma_{e}^{2}=1$	-	0.98(0.00)	-	0.98(0.00)
AMSE	0.98	0.81	1.00	0.92
$\alpha=1$	4.40(7.69)	4.04(7.69)	9.80(41.28)	9.80(41.28)
$\mu_{\beta}=0.995$	0.89(0.08)	0.89(0.08)	0.94(0.04)	0.94(0.04)
$\sigma_{\beta}^{2}=0.1$	0.01(0.00)	0.01(0.00)	0.01(0.00)	0.01(0.00)
$\sigma_{\varepsilon}^{2}=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.5$	-	0.09(0.07)	-	0.05(0.06)
$\sigma_{e}^{2}=1$	-	0.99(0.00)	-	0.99(0.00)
AMSE	2866.80	2504.82	61689.40	49759.63
$\alpha=1$	5.16(14.23)	5.16(14.23)	6.59(14.30)	6.59(14.30)
$\mu_{\beta}=0.995$	0.86(0.09)	0.86(0.09)	0.90(0.04)	0.90(0.04)
$\sigma_{\beta}^{2}=0.1$	0.00(0.00)	0.00(0.00)	0.00(0.00)	0.00(0.00)
$\sigma_{\varepsilon}^{2}=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.9$	-	0.12(0.09)	-	0.05(0.06)
$\sigma_{e}^{2}=1$	-	0.99(0.00)	-	0.99(0.00)
AMSE	12341.83	10821.51	73689.6	60893.54

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Table 3. Monte Carlo average and standard deviation of posterior mean, by fixing $\alpha=0, \sigma_\varepsilon^2=1, \sigma_e^2=1, \mu_\beta=0.5$, $\sigma_\beta^2=0.25$, and $\rho=0.1, 0.5, 0.9$ of RCA(1) and RCA(1)-AR(1) models

	$n=100$		$n=500$	
true	RCA(1)	RCA(1)-AR(1)	RCA(1)	RCA(1)-AR(1)
$\alpha=1$	1.07(0.21)	1.07(0.21)	1.02(0.11)	1.02(0.11)
$\mu_\beta=0.5$	0.42(0.11)	0.42(0.11)	0.35(0.06)	0.35(0.06)
$\sigma_\beta^2=0.25$	0.21(0.13)	0.21(0.13)	0.24(0.05)	0.24(0.05)
$\sigma_\varepsilon^2=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.1$	-	0.81(0.05)	-	0.85(0.05)
$\sigma_e^2=1$	-	1.00(0.00)	-	1.00(0.00)
AMSE	109.83	3.90	19.20	0.42
$\alpha=1$	1.07(0.23)	1.07(0.23)	1.03(0.13)	1.03(0.13)
$\mu_\beta=0.5$	0.32(0.20)	0.32(0.20)	0.24(0.06)	0.24(0.06)
$\sigma_\beta^2=0.25$	0.23(0.04)	0.23(0.04)	0.31(0.07)	0.31(0.07)
$\sigma_\varepsilon^2=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.5$	-	0.78(0.06)	-	0.83(0.02)
$\sigma_e^2=1$	-	1.00(0.00)	-	1.00(0.00)
AMSE	128.55	6.301	23.82	0.69
$\alpha=1$	1.08(0.29)	1.08(0.29)	1.05(0.18)	1.05(0.18)
$\mu_\beta=0.5$	0.28(0.19)	0.28(0.19)	0.19(0.14)	0.19(0.14)
$\sigma_\beta^2=0.25$	0.25(0.08)	0.25(0.08)	0.39(0.09)	0.39(0.09)
$\sigma_\varepsilon^2=1$	1.00(0.00)	1.00(0.00)	1.00(0.00)	1.00(0.00)
$\rho=0.9$	-	0.71(0.08)	-	0.79(0.03)
$\sigma_e^2=1$	-	1.00(0.00)	-	1.00(0.00)
AMSE	204.04	16.75	41.20	1.82

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

2.7 Applications in Real Data

In this section, we consider the application of RCA(1) and RCA(1)-AR(1) models using the Bayesian approach that we developed in the previous section and shown the output at Table 4 and Figure 4. 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 1975 to 2012 giving a total of 453 observations which is collected from *http://www.set.or.th/th/market/market statistics.html.*

Table 4. Posterior summaries (average and standard deviation) of RCA(1) and RCA(1)-AR(1) estimator of SET index

Parameters	RCA(1)	RCA(1)-AR(1)
α	6.62(4.50)	6.62(4.50)
μ_β	0.31(434.76)	0.31(434.76)
σ_β^2	102.11(5.78)	102.11(5.78)
σ_ε^2	3058.80(21.30)	3058.80(21.30)
ρ	-	0.99(0.004)
σ_e^2	-	3251.15(220.98)
MSE	198124.8	13.163

From Table 4, the MSE 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.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

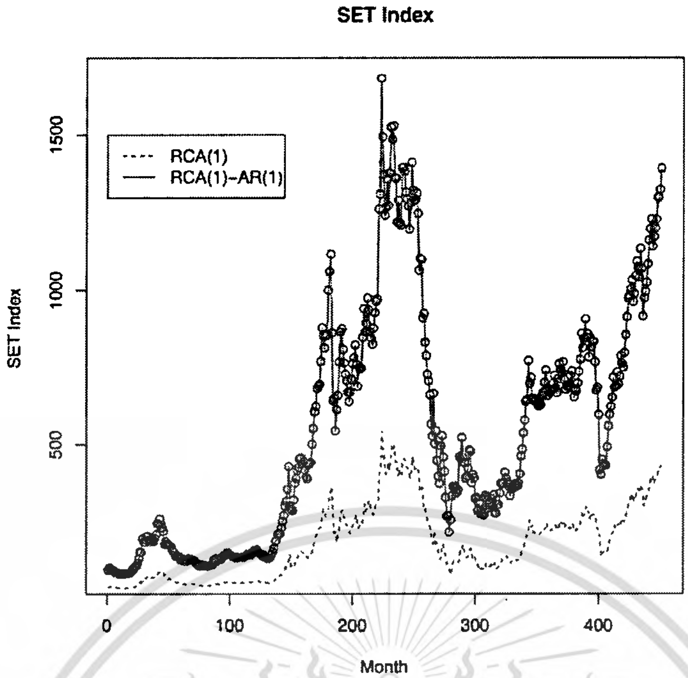


Figure 4. The time series plot for SET index and estimated of RCA(1),AR(1) model

In Figure 4, the bottom panel is the plot of SET index, the dashed line is the RCA(1) estimated model, and the 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.

2.8 Conclusion

In this paper, we have developed the Bayesian analysis for estimating parameters of the first order in RCA and RCA-AR models which are called the RCA(1) and RCA(1)-AR(1) models. For the simulation study, we also investigated the performance of Bayesian analysis and showed the averages and standard deviation of the posterior mean, and AMSEs for different data at the sample size of 100 and 500. It appears that all cases of RCA(1)-AR(1) model performs well in picking up the correct model, to see the AMSE is minimum. It is indicated that it is the time series data based on the errors which are affected the estimation.

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

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

better than the RCA(1) as seen from MSE of the RCA(1)-AR(1) model. It is smaller than the MSE of the RCA(1) model. We would recommend users to fit the RCA(1)-AR(1) models to time series data where stationary and non-stationary data are expected.

References

[1] R. F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50** (1982), 987-1007.

[2] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* **31** (1986), 307-327.

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

[4] D. F. Nicholls, B. G. Quinn, *Random coefficient autoregressive models: An introduction.*, Springer- Verlag Inc, New York, 1982.

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

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

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

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

[9] H. Tong, *Threshold models in nonlinear time series analysis* , Springer-Verlag Inc, Heidelberg, 1982.

[10] T. Bossaerts, W. Hardle, C. Hafner, Foreign exchange-rates have surprising volatility, *Athens Conference on Applied Probability and Time Series* **2** (1996), 55-72.

[11] A. Gelfand , A. F. M. Smith, Sampling-based approaches to calculating marginal densities, *Journal of American Statistical Association* **85** (1990), 398-409.

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

[13] D. J. Lunn, A. Thomas, N. Best, D. Spiegelhalter, WinBUGS-a Bayesian modeling framework: concepts, structure, and extensibility, *Statistics and computing* **10** (2000), 325-337.

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.