รายงานการวิจัยฉบับสมบูรณ์

การวิเคราะห์การถดถอยแบบไม่ใช้พารามิเตอร์ของตัวแบบการถดถอยผันผวนไม่ใช้พารามิเตอร์ในข้อมูลอนุกรมเวลา

Nonparametric Regression Analysis of Nonparametric Heteroscedastic regression model for Time Series Data

นางสาวอัชฌา อระวีพร

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2560
คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

An Estimating Parameter of Local Polynomial Regression on Economic Time Series Data

Autcha Araveeporn

Department of Statistics, Faculty of Science,
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
kaautcha@hotmail.com

Abstract. The local linear estimation, Nadaraya-Watson estimation, and locally weighted scatterplot smoothing (lowess) estimation are the way for estimating unknown parameter of local polynomial regression based on nonparametric regression model, fitting function of independent variable by the weighted least squares is computed for economic time series data. With the local polynomial regression fitting we can estimate the local linear estimator based on kernel function as a Gausian density function, and the Nadaraya-Watson estimation which is a part of local linear estimation or called a local constant estimation. Furthermore the lowess is used a weighted least squares to estimate parameter that modified the kernel function as a tricube kernel. The goal of this article is to show, through application, which three estimations can be used the best for fitting economic time series data. The Average Mean Square Errors (AMSE) is considered as a criterion to check a bias of fitted estimator. The data of this paper is studied from two cases following the simulation data and the actual data on economic time series data. For simulation data, the data is generated by autoregressive process by several coefficients. The Nadaraya-Watson estimation outperforms other methods by showing the minimum of AMSE values while the results of actual data are similar the simulation data.

Keywords: local linear estimation; local polynomial regression; locally weighted scatter plot smoothing; Nadaraya-Watson estimation

1 Introduction

Regression analysis is a statistical tool for the investigation of relationship between independent and dependent variables in term of regression model. An estimated parameter of regression model requires an assumption such as the linearity, statistical independence, homoscedasticity, and normality in the form of the underlying regression analysis. If an inappropriate regression model is used, it is possible to produce misleading conclusions. To overcome the difficulty caused by the restrictive assumption of the regression function, this approach leads to so-called nonparametric regression.

Typically, the nonparametric regression methods are overcome this problem, because the fitted model interpolates on the curve of data base on bandwidth values. The local polynomial regression is a class of nonparametric regression method which has a long history in the smoothing of time series. Nadaraya [2] and Watson [3] studied the simplest and most useful of local polynomial regression are called the local constant and linear smoothers. Stone [4] examined the consistency properties of many nonparametric regression estimators in local polynomial regression. Cleveland [5] was a catalyst for renewed interest in local polynomial regression, introducing lowess as using a tricube kernel function with bandwidth based on neighbor distances. Müller [6] studied certain equivalences between local polynomial kernel estimators and kernel estimators.

In this paper, we consider the nonparametric regression model in Section 2 and apply in local polynomial regression methods in Section 3. In the Section 4, we show the fitting estimator of simulation data and actual data in Section 5, and conclude in Section 6.

2 The Nonparametric Regression

The simple of nonparametric regression model written as

$$y_t = f(x_t) + \varepsilon_t, \qquad t = 1, 2, \ldots, n \tag{1}$$

where x_t are know the independent variable at the time points, y_t are the dependent variable at the time points, ε_t denote the measurement errors, and $f(x_t)$ are the nonparametric regression function that we want to estimate.

3 Methods of Parameter Estimation

The local polynomial regression is a method based on a class of kernel function that can be used to estimate the nonparametric regression model at a particular point by locally fitting a pth degree polynomial to the data via weighted least squares. This class includes the Nadaraya-Watson estimators since it corresponds to fitting degree zero polynomials. The importance and simplicity is the local polynomial regression corresponding to $p = 1$ or called the local linear regression.

The main concept of local polynomial regression estimator is to approximate parameter by a polynomial of some degree. The Taylor expansion can be approximated by a polynomial of degree p and assumed that $f(x_t)$, as

$$f(x_t) = f(x_0) + (x_t - x_0)f^{(1)}(x_0) + \frac{1}{2}(x_t - x_0)^2 f^{(2)}(x_0) + \ldots, t = 1, 2, \ldots, n \tag{2}$$

where we define $\beta_j = f^{(j)}(x_0)/j\,!$ for $j = 0, 1, 2, \ldots$, then the local approximation is

$$f(x_t) = \sum_{j=0}^{p} (x_t - x_0)^j \beta_j.$$

The local polynomial regression estimator is used the weight least-squares to minimize of

$$(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p) = \arg\min \sum_{t=1}^{n} \left(y_t - \sum_{j=0}^{p} (x_t - x_0)^j \beta_j \right)^2 K_h(x_t - x_0), \qquad (3)$$

where $K(\cdot)$ is a kernel function, $h > 0$ is called the bandwidth and $\boldsymbol{\beta}$ denotes the vector of coefficient $(\beta_0, \beta_1, \ldots, \beta_p)^\top$ evaluated at x_t.

In matrix form, the weight least squares problems of (3) is to minimize

$$(\boldsymbol{y} - \boldsymbol{X}\beta)^\top \boldsymbol{W} (\boldsymbol{y} - \boldsymbol{X}\beta),$$

where

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 - x_0 & \ldots & (x_1 - x_0)^p \\ 1 & x_2 - x_0 & \ldots & (x_2 - x_0)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_0 & \ldots & (x_n - x_0)^p \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

and $\boldsymbol{W}$ is a diagonal matrix with n^{th} diagonal element $K\left(\frac{x_t - x_0}{h}\right)$ as

$$\boldsymbol{W} = \begin{pmatrix} K_h(x_1 - x_0) & 0 & \ldots & 0 \\ 0 & K_h(x_2 - x_0) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & K_h(x_n - x_0) \end{pmatrix}.$$

We get the weighted least squares estimator as

$$\hat{\beta} = \left(\boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X} \right)^{-1} \boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{y}. \qquad (4)$$

The estimators of local polynomial regression are depended on the process of weight least squares and the bandwidth selection. The following methods are shown the parameter estimation on each method.

3.1 The Local Linear Estimation

The nonparametric regression model can also be estimated by using the local linear regression model to fit local linear regression introduced by Fan [7] and Fan [8]. For the weighted least squares in (4), the local linear estimator can be approximated by setting $p = 1$, and the matrix form of $\boldsymbol{X}$, $\boldsymbol{y}$, and $\boldsymbol{W}$ can be written as

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_1 - x_0 \\ 1 & x_2 - x_0 \\ \vdots & \vdots \\ 1 & x_n - x_0 \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

and

$$\boldsymbol{W} = \begin{pmatrix} K_h(x_1 - x_0) & 0 & \dots & 0 \\ 0 & K_h(x_2 - x_0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & K_h(x_n - x_0) \end{pmatrix}.$$

The weighted least squares can approximate estimators by

$$\hat{\beta}(x;h) = (\hat{\beta}_0(x;h), \hat{\beta}_1(x;h)) = \left(\boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \boldsymbol{W} \boldsymbol{y}$$
$$= \begin{pmatrix} S_0(x;h) & S_1(x;h) \\ S_1(x;h) & S_2(x;h) \end{pmatrix}^{-1} \begin{pmatrix} T_0(x;h) \\ T_1(x;h) \end{pmatrix}. \quad (5)$$

The local linear estimator is

$$\hat{\beta}_0(x;h) = \frac{T_0(x;h)\ S_2(x;h) - T_1(x;h)\ S_1(x;h)}{S_0(x;h)\ S_2(x;h) - S_1^2(x;h)}, \quad (6)$$

and

$$\hat{\beta}_1(x;h) = \frac{T_1(x;h)\ S_0(x;h) - T_0(x;h)\ S_1(x;h)}{S_0(x;h)\ S_2(x;h) - S_1^2(x;h)}, \quad (7)$$

where

$$S_j(x;h) = \sum_{t=1}^{n} K_h(x_t - x_0)(x_t - x_0)^j,$$

$$T_j(x;h) = \sum_{t=1}^{n} K_h(x_t - x_0)(x_t - x_0)^j y_t,$$

for $j = 0, 1, 2$.

The quality of the local linear regression depends on the bandwidth selection method. Ruppert, et al.[9] studied the ideas of plug-in bandwidth selection for local linear regression of kernel estimators. The plug-in bandwidth selection is based on the asymptotic mean squared error as

$$h^5 = \frac{\sigma^2 R(K)}{n(\mu_2(K))^2 (f''(x))^2},$$

where $R(K) = \int K^2(z)dz$, $K = K_h(x - x_0)$, and $f''(x)$ is the second derivatives of $f(x)$.

3.2 Nadaraya-Watson Estimation

Yoa and Tong [10] suggested the simple decomposition in estimating smooth trend function and smooth volatility function evaluated by using the Nadaraya-Watson kernel estimator (Nadaraya [2] and Watson [3]). The trend estimator is written as

$$\hat{\beta}_0(x;h) = \frac{\sum_{t=1}^{n} K_h(x_t - x_0) y_t}{\sum_{t=1}^{n} K_h(x_t - x_0)}, \quad (8)$$

where h is known as the bandwidth parameter which controls the smoothness of the estimated curve and the kernel weights.The bandwidth can be chosen using a cross-validation criteria or information criteria. Allen [11] and Stone [12] proposed the cross-validation(CV) method to select the smoothing parameter(λ) by the cross-validation criterion

$$CV(\lambda) = n^{-1}\sum_{t=1}^{n}\{y_t - \hat{f}_\lambda^{(-t)}(x_t)\}^2 \tag{9}$$

where $\hat{f}_\lambda^{(-t)}(x_t)$ is the fitted value at time point x_t, using all data except (y_t, x_t), with the smoothing parameter λ for $t = 1, \ldots, n$. The basic idea of cross-validation is to choose the value of λ that minimizes $CV(\lambda)$.

The kernel functions K can be chosen as the Gaussian density function given by,

$$K(u) = (2\pi)^{-1/2}e^{-u^2/2}.$$

For the weighted least squares in (4), the local constant estimator can be approximated by setting $p = 0$, the local constant estimator is rewritten the same as the trend estimator of Nadaraya-Watson estimator.

3.3 lowess Estimation

Cleveland [5] introduced an alternative form of the local linear regression which is called lowess ,locally weighted scatter plot smoothing. The basic idea is to start the weighted least square as the kernel function in term of the tricube kernel function

$$K(t) = \frac{70}{81}(1-\mid t\mid^3)^3, \text{if} \mid t \mid \leqslant 1.$$

Cleveland [5] proposed a nearest neighbor bandwidth by setting $r = nf + 0.5$. For each predictor x_t, let

$$h_k = \mid x_k - x_t \mid_{(r)},$$

where h_k is the rth order statistic of the sample $\mid x_k - x_1 \mid, \mid x_k - x_2 \mid, \ldots, \mid x_k - x_n \mid$.

The local linear estimators of lowess are to minimize

$$(\hat{\beta}_0(x; h_k), \hat{\beta}_1(x; h_k)) = \arg\min \sum_{t=1}^{n}\left(y_t - \sum_{j=0}^{1}(x_t - x_0)^j\beta_j\right)^2 K_{h_k}(x_t - x_k), \tag{10}$$

The lowess estimators $(\hat{\beta}_0(x; h_k), \hat{\beta}_1(x; h_k))$ are similar the local linear estimator as (6) and (7), but kernel function and bandwidth selection are modified by tricube kernel function and nearest neighbor bandwidth.

4 Application of Simulation Data

The data is generated in term of autoregressive model in order 1 (AR1) following

$$y_t = \rho y_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim N(0,1)$, the coefficient ($\rho$) of AR(1) is defined by $0.1, 0.5, 0.7$, and 0.99, and the sample sizes $n = 100, 200, 300$, and 400. The Figure 1 shows the sample of simulation data when $n = 200$ and $\rho = 0.1, 0.5, 0.7$, and 0.99.

The accuracy of fitting function is consider by the Mean Square Errors (MSE) as follows:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \tag{11}$$

where y_t denoted the simulated data and $\hat{y}_t$ denoted the fitting data of 3 estimated methods From Table 1, the results show that the average MSE of Nadaraya-Watson estimator is the minimum values in all cases.

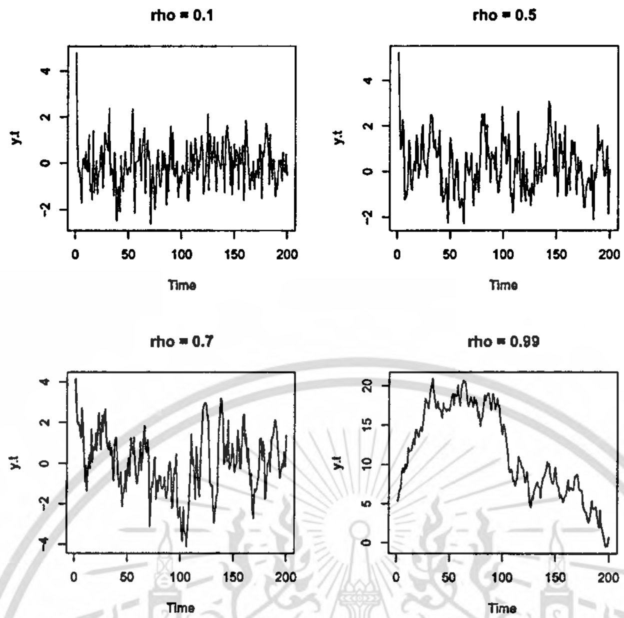

Fig. 1. The time series plot of simulation data when $n = 200$ and $\rho = 0.1, 0.5, 0.7$, and 0.99

Table 1. The average MSE on simulated data of local linear (LL), Nadaraya-Watson (NW), and locally weighted scatter plot smoothing (lowess) estimations.

Sample sizes	Method	ρ =0.1	ρ =0.5	$\rho = 0.7$	$\rho = 0.99$
	LL	1.0513	0.7593	0.6099	0.4822
n=100	NW	0.5438	0.3604	0.2966	0.2197
	lowess	1.2167	1.4810	1.8635	3.3187
	LL	1.0438	0.9299	0.8489	0.6327
n=200	NW	0.5137	0.3494	0.2918	0.2228
	lowess	1.1113	1.4033	1.9588	6.5061
	LL	1.0337	1.0671	1.0527	0.7893
n=300	NW	0.5011	0.3501	0.2910	0.2190
	lowess	1.0709	1.3906	1.9561	9.9320
	LL	1.0377	1.1317	1.2328	0.9651
n=400	NW	0.5042	0.3468	0.2908	0.2227
	lowess	1.0672	1.3736	1.9885	12.3403

5 Application of Actual Data

The actual data is the monthly average price of crude oil in Dollars per Barrel from January 1987 to December 2016 that consisted of 348 records. Let y_t denote the Dollars per Barrel of month t, where $x_t = 1$ represents January of 1987 and $x_t = 348$ represents December of 2016. The crude oil price are fitted function based on local linear, Nadaraya-Watson, and lowess estimations shown in Figure 2.

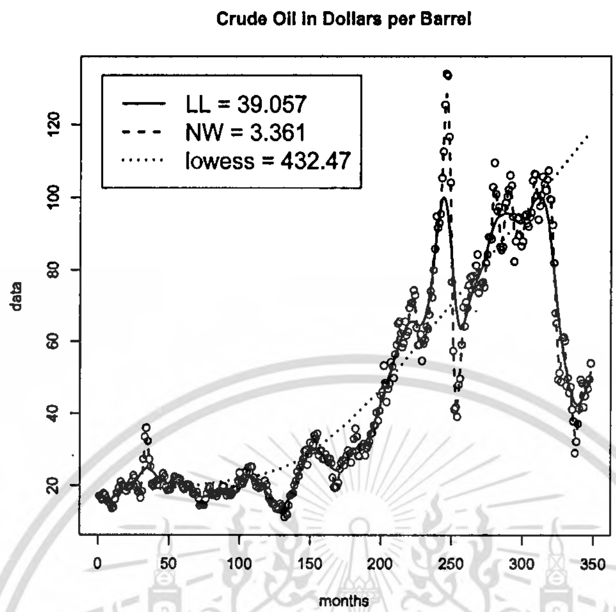

Fig. 2. The time series plot of crude oil in Dollars per Barrel and fitting value of local linear (LL), Nadaraya-Watson (NW), and locally weighted scatter plot smoothing (lowess) estimation

From Figure 2, it can be seen that the Nadaraya-Watson estimation is a good fitted function which is closed the plot of actual data, so the MSE of Nadaraya-Watson method is 3.361 which is the smallest value.

6 Conclusion

We have been discussed the method to estimate parameter of local polynomial regression function based on nonparametric regression function. It is concluded that the estimation of Nadaraya-Watson method has indicated a good performance more than local linear and lowess methods in case of simulated data and actual data. Hence it can be say that Nadaraya-Watson method leads to a nearly enough class of local polynomial regression to estimate adequately fitted economic time series data. The cross-validation criteria is a good performance for fitting model.

References

1. Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
2. Nadaraya, E. A., : On Estimating Regression. Theory of Probability and Its Application. 10, 186–196 (1964)
3. Watson, G.S., : Smooth regression analysis. Sankhyā, Series A. 26, 359–372 (1964)
4. Stone, C.G., : Consistent nonparametric regression. The Annals of Statistics. 5, 959–620 (1977)
5. Cleveland, W.S., : Robust Locally Weight Regression and Smoothing Scatter Plots. Journal of American Statistics Association. 74, 829–836 (1979)
6. Müller, H. G., : Weighted Local Regression and Kernel Methods for Nonparametric Curve Fitting. Journal of American Statistics Association. 82, 231–238 (1987)
7. Fan, J., : Design-adaptive nonparametric regression. Journal of American Statistics Association. 87, 998–1004 (1992)
8. Fan, J., : Local linear regression smoothers and their minimax efficiency. The Annals of Statistics. 21, 196–226 (1993)
9. Ruppert, D., Sheather, S.J., Wand, M.P., : An Effective Bandwidth Selector for Local Least Squares Regression. Journal of the American Statistical Association. 90, 1257–1270 (1995)
10. Yao, Q., Tong, H., : Quantifying the influence of initial values on nonlinear prediction. Journal of the Royal Statistical Society, Series B. 56, 701–725 (1994)
11. Allen, D.M., The relationship between variable and data augmentation and a method of prediction. Technometrics. 16, 125–127 (1974)
12. Stone, M., Cross-validatory choice and assessment of statistical predictions(with discussion). Journal of the Royal Statistical Society, Series B. 36, 111–147 (1974)

Iteration of Smoothing Spline Method by Nonparametric Regression Time Series Model

Autcha Araveeporn *

Abstract

The goal of this work is to develop a nonparametric regression model for forecasting nonstationary time series data. The trend and the heteroscedasticity are modeled using a class of smoothing splines method. The model is then simulated to investigate the behavior of these trend ($\mu(\cdot)$), heteroscedasticity or volatility ($\sigma(\cdot)$) with the several sample sizes. Simulation results have shown that the proposed trend and volatility estimator provide asymptotically unbiased estimates nearly for all two data generating models

Keywords: Nonparametric regression; Smoothing spline; Volatility;

*Lecturer, Department of Applied Statistics, Faculty of Science, King Monkut's Institute of Technology Ladkrabang, Bangkok, Thailand, 10520. *E-mail : kaautcha@kmitl.ac.th*

1 Introduction

The modeling of available explanatory variables has a variety of applications in regression model. The parametric and nonparametric method are the choice for estimating regression function between two sets of variables that consist of a vector of predictors and a response variable. A parametric regression model requires an assumption in the form of the underlying regression function. To overcome the difficulty caused by the restrictive assumption of the parametric form of the regression function, one may remove the restriction that the regression function belongs to a parametric family. This approach leads to so-called nonparametric regression.

Typically, the nonparametric regression methods are based on a smoothing technique which produces a smoother. A smoother is a tool for summarizing the trend of a response variable as a function of one or more predictor variables. The single predictor case is called scatterplot smoothing that can be used to enhance the visual appearance of the scatterplot of response versus predictor variable, to help our eyes pick out the trend in the plot(Hastie and Tibshirani,1990). There are many smoothing techniques, E.g., a local polynomial regression (Wand and Jones, 1995, Fan and Gijbels, 1996 , regression splines (Eubank, 1988, 1999), smoothing splines (Wahba, 1990, Green and Silverman, 1994), and penalized splines (Ruppert, Wand and Carroll, 2003). These smoothing techniques are generally based on the assumption of homoscedastic variance model which may not be suitable when the data involves high volatility.

There are several methods to model volatility in time series, such as the autoregressive conditional heteroscedastic model (ARCH) by Engle (1982), who was the first to introduce the ARCH model to obtain the predictive variance for U.K. inflation rate. Gouriéroux and Monfort (1992) and Masry and Tjøstheim (1999a) have proposed the conditional heteroscedastic autoregressive nonlinear (CHARN) model in financial time series. For simplicity, the case is one lag of the CHARN model were studied to model the foreign exchange rates (Bossaerts et al., 1996). Nonparametric smoothing techniques can be applied for the estimation of CHARN model by considering the response and predictor variables in terms of nonparametric regression by using the nonparametric conditional heteroscedastic autoregressive nonlinear (NCHARN) model.

This paper explores various nonparametric models for trend and conditional variance functions under NCHARN model. Section 2 describes the method for estimation of smoothing spline estimator. Section 3 applies the methodology from Section 2 to iteration of smoothing spline method. Section 4 shows the results of simulation study and we discuss the results and present future research ideas in Section 5.

2 Estimation of Smoothing Spline Estimator

Wahba (1990) defined the natural polynomial spline $s(x) = s_n^m(x)$ is a real-valued function on $[a, b]$ with the aid of n so-called knots $-\infty \leq a < x_1 < x_2 < \ldots < x_n < b \leq \infty$. 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 $\mu \in W^m[a, b]$ that used to measure the roughness of curve which is called the quadratic penalty function given by

$$\int_b^a \{\mu^{(m)}(x)\}^2 dx \tag{2.1}$$

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

Consider the simple nonparametric regression model where the observation y_t at design points $x_t, t = 1, 2, \ldots, n$ assumed to satisfy

$$y_t = \mu(x_t) + \varepsilon_t, \qquad t = 1, 2, \ldots, n \tag{2.2}$$

where $\mu(\cdot)$ denotes a smooth function. To estimate $\hat{\mu}(\cdot)$ minimizes $s_n^{(m)}(\mu)$ over the class of function $\mu(\cdot)$ following

$$s_n^{(m)}(\mu) = \sum_{t=1}^{n} \{y_t - \mu(x_t)\}^2 + \lambda \int_a^b \left\{\mu^{(m)}(x)\right\}^2 dx \tag{2.3}$$

where $\lambda > 0$ denotes a smoothing parameter to be determined by a suitable cross-validation criteria or information criteria. The smoothing parameter controls the trade-off between fidelity to the data and roughness of function, if $\lambda \to 0$, the $\hat{\mu}_\lambda(\cdot)$ converges to linear function, if $\lambda \to \infty$, the $\hat{\mu}_\lambda(\cdot)$ converges to interpolating spline.

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). We use this class of cubic smoothing splines to fit $\hat{\mu}(\cdot)$ by starting with the simple nonparametric regression model.

The first procedure of smoothing spline is considered a least square problem to fit a function $\mu(\cdot)$ that minimizes the residuals sum of squares

$$RSS = \sum_{t=1}^{n} \{y_t - \mu(x_t)\}^2 \tag{2.4}$$

Assuming the range of $\mu(\cdot)$ in (2.2) is finite interval, $[a, b] = [x_1, x_n]$, where x_1 denotes the ith order statistic and the roughness penalty of $\mu(\cdot)$ is measured by

$$\int_a^b \{\mu''(x_t)\}^2 \, dx \tag{2.5}$$

This leads of the following penalized least squares regression to find $\hat{\mu}_\lambda(x_t)$ called the smoothing spline estimator by minimizing

$$\hat{\mu}_\lambda(x_t) = \arg\min_{\mu} S_\lambda(\mu) \tag{2.6}$$

and

$$S_\lambda(\mu) = \sum_{t=1}^{n} \{y_t - \mu(x_t)\}^2 + \lambda \int_a^b \{\mu''(x)\}^2 \, dx \tag{2.7}$$

where $\lambda > 0$ is a smoothing parameter controlling the size of the roughness penalty and used to trade-off the goodness of fit.

3 Proposed Trend and Volatility Estimation

In this section we explain how to fit $\hat{\mu}(.)$ and $\hat{\sigma}(.)$ from the NCHARN model following

$$y_t = \mu(x_t) + \sigma(x_t)\ \varepsilon_t\ , \quad t = 1, 2, \ldots \tag{3.1}$$

by starting in the simple nonparametric regression model written as

$$y_t = \mu(x_t) + \varepsilon_t, \qquad t = 1, 2, \ldots \tag{3.2}$$

where (y_t, x_t) are a set of predictor and response variables, $\mu(.)$ is a smooth unknown trend (conditional mean) function or $\mu(x_t)$ is the conditional mean of y_t given to x.

We estimate the trend $\mu(x_t)$ and volatility $\sigma^2(x_t)$ using an iterative two step-method. First, assuming initially that $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$ we obtain the initial estimate of $\mu(\cdot)$ as

$$\hat{\mu}^{(0)}(x) = SS\left\{(y_t, x_t); t = 1, \ldots, n\right\} \tag{3.3}$$

where $SS\left\{(y_t, x_t); t = 1, \ldots, n\right\}$ denotes the smoothing spline estimate of $\mu(\cdot)$. Using the data $\{(y_t, x_t); t = 1, \ldots, n\}$ we also select the smoothing parameter using the method of generalized cross-validation(GCV) suggested by Wahba (1977) and Craven and Wahba (1979) . In practice, this step can be implemented by using the function "smooth.spline" in the software R.

Next, we reparametrize $\sigma^2(x) = \exp\{h(x)\}$ and obtain the initial estimate of $h(x)$ using the first stage residual's.

$$\hat{\delta}_t^{(0)} = y_t - \hat{\mu}^{(0)}(x_t) \tag{3.4}$$

Notice that

$$\log \hat{\delta}_t^2 - E[\log \varepsilon_t^2] = h(x_t) + \log \varepsilon_t^2 - E[\log \varepsilon_t^2] \tag{3.5}$$

Now if we require ε_t to be normally distributed with mean 0 and variance 1, then $E[\log \varepsilon_t^2] = -1.2704$ and hence we can apply the smoothing spline method again to obtain

$$\hat{h}^{(0)}(x) = SS\{(\log \hat{\delta}_t^{(0)^2} + 1.2704, x_t) : t = 1, \ldots, n\} \tag{3.6}$$

Next we use the initial estimate $\hat{h}^{(0)}(x)$ to re-estimate $\mu(\cdot)$. Notice that

$$e^{-\frac{h(x_t)}{2}}\ y_t = e^{-\frac{h(x_t)}{2}}\ \mu(x_t) + \varepsilon_t \tag{3.7}$$

Hence we can first obtain the smoothing spline estimate of $e^{-\frac{h(x)}{2}}\ \mu(x) = g(x)$ using the "data" $\{(e^{-\frac{\hat{h}^{(0)}(x_t)}{2}}\ y_t,\ x_t) : t = 1, \ldots, n\}$ and then obtain

$$\hat{\mu}^{(1)}(x) = e^{\frac{\hat{h}^{(0)}(x)}{2}}\ \hat{g}^{(1)}(x) \tag{3.8}$$

where

$$\hat{g}^{(1)}(x) = SS\{(e^{-\frac{\hat{h}^{(0)}(x_t)}{2}}\ y_t,\ x_t) : t = 1, \ldots, n\} \tag{3.9}$$

Now we continue the iterative process with the initial estimate $\hat{\mu}^{(0)}(x)$ and $\hat{h}^{(0)}(x)$ as follows for $l = 1, 2, \ldots$

$$\hat{\mu}^{(l)}(x) = e^{\frac{\hat{h}^{(l-1)}(x)}{2}}\ \hat{g}^{(l)}(x) \tag{3.10}$$

where

$$\hat{g}^{(l)}(x) = SS\{(e^{-\frac{\hat{h}^{(l-1)}(x_t)}{2}}\ y_t,\ x_t) : t = 1, \ldots, n\} \tag{3.11}$$

and

$$\hat{h}^{(l)}(x) = SS\{(\log \hat{\delta}_t^{(l)^2} + 1.2704, x_t) : t = 1, \ldots, n\} \tag{3.12}$$

where

$$\hat{\delta}_t^{(l)} = y_t - \hat{\mu}^{(l)}(x_t) \tag{3.13}$$

We continue the iterative process until we find that

$$\hat{\varepsilon}_t^{(l)} = \frac{y_t - \hat{\mu}^{(l)}(x_t)}{\hat{\sigma}^{(l)}(x_t)} \tag{3.14}$$

where

$$\hat{\sigma}^{(l)}(x) = \exp\left\{\frac{\hat{h}^{(l)}(x)}{2}\right\} \tag{3.15}$$

normally distributed with mean 0 and variance 1. Such a test can be carried out using the Kolmogorov-Smirnov(KS) test.

Let $\hat{\mu}(x)$ and $\hat{\sigma}(x) = \exp\{\frac{\hat{h}(x)}{2}\}$ denote the converged estimates of $\mu(\cdot)$ and $\sigma(\cdot)$ by using a specified size of the KS test. Let

$$\hat{\varepsilon}_t = \frac{y_t - \hat{\mu}(x_t)}{\hat{\sigma}(x_t)}, \qquad t = 1, \ldots, n \tag{3.16}$$

denote the standardized residuals based on the converged values of $\hat{\mu}(x)$ and $\hat{\sigma}(x)$ as described above.

4 Simulation Study

In simulation study to estimate the performance of smoothing spline method is divided into two parts. The first part consists of the study in NCHARN model

$$y_t = \mu(x_t) + \sigma(x_t)\ \varepsilon_t, \qquad t = 1, 2, \ldots, n \tag{4.1}$$

where $\mu(x_t)$ and $\sigma(x_t)$ are generated from 2 model following

Model I : (Fan and Yao, 1998)

$$\begin{aligned} \mu(x_t) &= 0.5(x_t + 2\exp\{-16\ x_t^2\}) \\ \sigma(x_t) &= -0.4\exp\{-2\ x_t^2\} + 0.2 \end{aligned}$$

Model II:

$$\begin{aligned} \mu(x_t) &= x_t^3 - (x_t^2\ \cos(x_t)) - \exp\left\{\frac{x_t}{1+|x_t|}\right\} \\ \sigma(x_t) &= \exp\left\{\sin(x_t) - \frac{x_t}{1+x_t^2}\right\} \end{aligned}$$

where in each of two cases $x_t \sim$ Unif.$(-2, 2)$.

In Figure 1, we present the curve of trend, $\mu(\cdot)$, volatility, $\sigma(\cdot)$, in Model I and II.

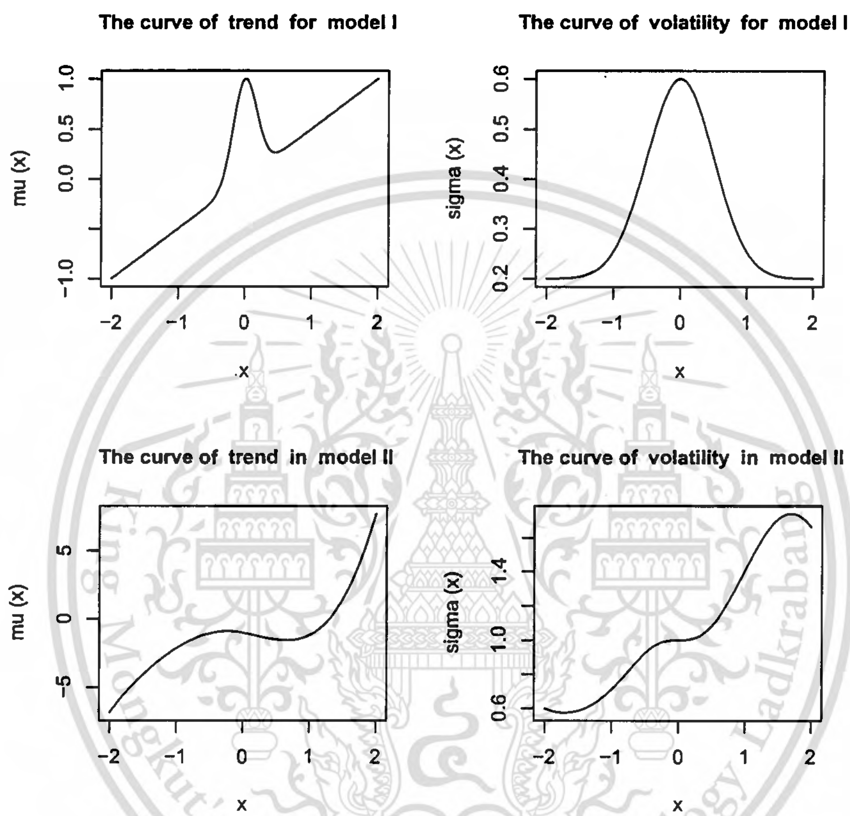

Figure 1: The curve of trend and volatility in Model I and II

The error process $\varepsilon_t, t = 1, 2, \ldots, n$ from in (4.1) which is assumed to follow the normal distribution with mean 0 and variance 1. We consider the sample sizes $n = 100, 200, 400,$ and 800.

The second part, the estimates of $\hat{\mu}(x_t)$ and $\hat{\sigma}(x_t)$, we compute the bias of $\mu(\cdot)$ and $\sigma(\cdot)$ by

$$\begin{aligned}\mu_{bias} &= \frac{1}{n}\sum_{t=1}^{n}\hat{\mu}(x_t) - \mu(x_t))\\ \sigma_{bias} &= \frac{1}{n}\sum_{t=1}^{n}\hat{\sigma}(x_t) - \sigma(x_t))\end{aligned}$$

We simulated data with the sample sizes $n = 100, 200, 400$, and 800, and repeated the data generation and model fitting 500 times. Tables 1 and 2 show various Monte Carlo(MC) summary statistics of the estimates obtained by the smoothing spline method.

The third and the fourth columns of these tables represent the MC 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 : bias $=$ 0 versus H_1 : bias $\neq 0$).

By observing the p-values, the results appear following:

From Tables 1 and 2, the smoothing spline method provides asymptotically unbiased estimates for $\mu(\cdot)$ and $\sigma(\cdot)$ nearly for all two data generating models, and the average relative biases seem to decrease with increasing sample sizes for 2 models.

The histogram of the biases of all parameter estimates are presented in Figures 2-5. From the histogram is apparent that relative biases reduce with increasing sample sizes and the distribution of the biases appear to be more normally distributed for larger sample sizes.

4.1 Smoothing Spline Method

Table 1: The simulation of smoothing spline method in Model I for different sample (500 replications)

bias	sample size	mean	s.d.	lci	uci	t-stat	p-value
μ_{bias}	n=100	0.030	0.690	-0.029	0.091	0.9954	0.320
	n=200	-0.002	0.031	-0.004	0.000	-1.556	0.120
	n=400	-0.000	0.018	-0.002	0.000	-0.966	0.334
	n=800	-0.000	0.013	-0.002	0.000	-1.652	0.099
σ_{bias}	n=100	0.039	0.775	-0.028	0.107	1.131	0.258
	n=200	-0.000	0.036	-0.004	0.002	-0.500	0.616
	n=400	0.001	0.020	-0.000	0.003	1.343	0.179
	n=800	0.002	0.014	0.001	0.004	4.404	0.000*

*indicates significance at 5% level

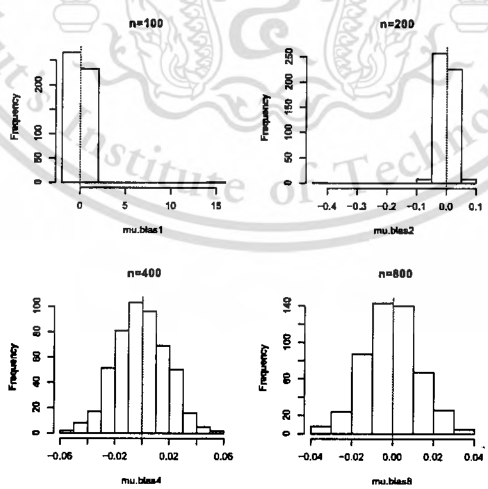

Figure 2: Histogram of Bias for $\mu(.)$ with Smoothing Spline Method in Model I

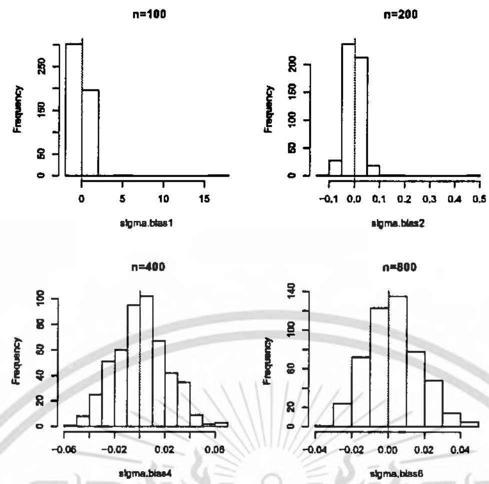

Figure 3: Histogram of Bias for $\sigma(.)$ with Smoothing Spline Method in Model I

Table 2: The simulation of smoothing spline method in Model II for different sample (500 replications)

bias	sample size	mean	s.d.	lci	uci	t-stat	p-value
μ_{bias}	n=100	0.030	0.690	-0.029	0.091	0.995	0.320
	n=200	-0.002	0.031	-0.004	0.000	-1.556	0.120
	n=400	-0.000	0.018	-0.002	0.000	-0.966	0.334
	n=800	-0.000	0.013	-0.002	0.000	-1.652	0.099
σ_{bias}	n=100	0.039	0.775	-0.028	0.107	1.131	0.258
	n=200	-0.000	0.036	-0.004	0.002	-0.500	0.616
	n=400	0.001	0.020	-0.000	0.003	1.343	0.179
	n=800	0.002	0.014	0.001	0.004	4.404	0.000*

*indicates significance at 5% level

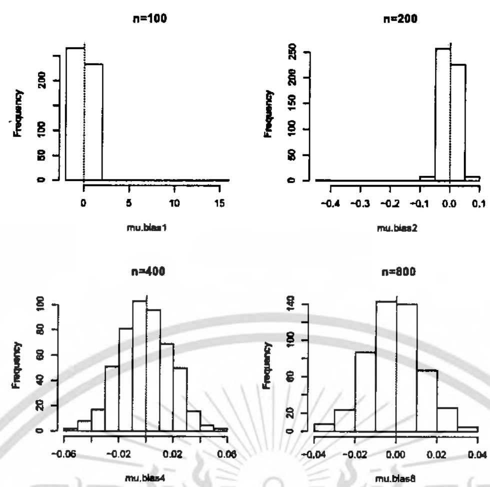

Figure 4: Histogram of Bias for $\mu(.)$ with Smoothing Spline Method in Model II

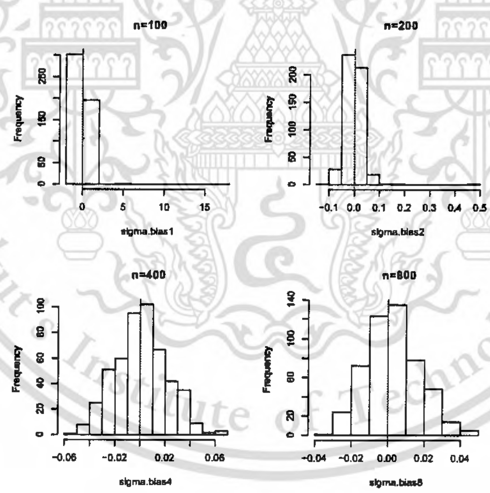

Figure 5: Histogram of Bias for $\sigma(.)$ with Smoothing Spline Method in Model II

5 CONCLUSION

In this section, we have investigated the smoothing spline method to estimate smooth unknown trend and smooth unknown volatility for NCHARN model.

By assuming NCHARN model from the smoothing spline method, we obtained the trend and volatility estimator by using the iterative process and selected the smoothing parameter using the method of generalized cross-validation.

Through a Monte Carlo simulation study, we evaluated the performance of smoothing spline method procedure and showed that the trend estimator ($\hat{\mu}(\cdot)$) and volatility estimator ($\hat{\sigma}(\cdot)$) work reasonably well for data of most to moderate sample sizes. However, the volatility estimator can not be concluded that the large sample sizes perform significantly better than the other.

Based on the nonparametric regression model we have developed the estimation of the trend and volatility using a Bayesian and Maximum Likelihood method approach.

REFERENCES

Bossaerts, P., Härdle, W., & Hafner, C., (1996). Foreign exchange-rates have surprising volatility. In *Athens Conference on Applied Probability and Time Series*, Vol. 2 (ed. P.M. Robinson). Lecture Notes in Statistics 115, New York : Springer, pp. 55-72.

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

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

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

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

Fan, J., & Gijbels, I., (1996). Local polynomial modelling and its applications. Chapman and Hall, London.

Fan, J., & Yao, Q., (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85, 645-660.

Gouriéroux, C., Monfort, A., (1992). Qualitative threshold ARCH models. *Journal of Econometrics*, 52, 159-199.

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

Hastie, T.J., & Tibshirani, R., (1990). Generalized Additive Models. Chapman and Hall, London.

Marsy, E., & Tjøstheim, D., (1995a). Non-parametric estimation and identification of ARCH nonlinear time series: strong convergence and asymptotic normality. *Econometric Theory*, 11, 258-289.

Ruppert, D., Wand, M.P., & Carroll. R.J., (2003). Semiparametric Regression. Cambridge University Press.

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

Wahba, G., (1977). A survey of some smoothing problems and the method of generalized cross-validation for solving them. In *Application of Statistics*, (P. R. Krisnaiah, ed.) North Holland, Amsterdam, 507-523.

Wand, M.P., & Jones, M.C., (1995). Kernel Smoothing. Chapman and Hall, London.