International Journal of Mathematical Archive-5(9), 2014, 65-72
IJMA Available online through www.ijma.info ISSN 2229 – 5046

SEMIPARAMETRIC ESTIMATION OF VOLATILITY

Ismail B[1] and Ashwini Kumari*[2]

[1]Department of Statistics, Mangalore University, Mangalagangothri, Mangalore-574199, India.

[2]Department of Statistics, Alva's College, Moodbidri- 574227, D.K. Karnataka, India.

(Received On: 24-08-14; Revised & Accepted On: 29-09-14)

ABSTRACT

Stochastic volatility models are considerable interest in empirical finance. We investigate the use of a semiparametric model for estimating volatility. ARCH models are commonly used to estimate volatility. But there are situations where influence of some exogenous factors on volatility is seen in practice in addition to the ARCH component. In this paper a new model for volatility is presented which includes a regressors (exogenous variable) in addition to ARCH component. The regression part is estimated using nonparametric Kernel smoothing technique and ARCH component is estimated by parametric approach. Further two methods are connected to build a combination forecasting model by combining nonparametric estimator of the regression function and parametric estimator of the ARCH effect. The practical application of the proposed model for forecasting volatility is examined for a sample of gold price returns. The proposed model shows minimum mean square error compared with existing models.

Keywords: *Nonparametric regression, Local linear Kernel smoothing, Volatility, GARCH, ARCH testing.*

1. INTRODUCTION

Financial econometrics is an active field of research which combines finance, economics, probability and statistics. Most financial data is available in time series form. Financial time series data provides the information about the prices of financial assets over a period of time. There are two main objectives of investigating financial time series. First it is important to understand how prices behave and the second objective is to use our knowledge of price behaviour to reduce risk or take better decisions.

Most of the financial studies involve returns instead of prices of assets. Returns from financial market variables measured over short time intervals like intra daily, daily, weekly etc. are uncorrelated, but not independent. Although the signs of successive price movements seem to be independent, their magnitude as represented by the absolute value or square of the price increments is correlated in time. This phenomena is denoted volatility clustering and indicates that the volatility of the series is time varying.

Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models (Bollerslev, 1986) and Stochastic volatility models (Aquilar and West 2000, Kim et all 1998) are the two main types of techniques which have been widely used for analyzing time varying volatility in financial econometrics. The success of the GARCH – models at capturing volatility clustering in financial market is extensively documented in the literature.

A time series is said to be heteroscedastic if its variance changes over time. Since Engle (1982) proposed the seminal work of Autoregressive Conditional Heteroscedasticity (ARCH) model, many researchers have been working in the field of modeling the time varying volatility of economic data. Modeling and forecasting time varying financial market volatility are important for investors.

A nonparametric method offer flexible ways of estimating volatility functions. In nonparametric context local linear kernel smoothing technique provide estimators which satisfies desirable asymptotic properties. Loader C (1999), Silverman, J.D., Hart (1997), Wand and Jones(1995) and Fan and Gijbels(1996) give a comprehensive coverage of these techniques.

Corresponding author: Ashwini Kumari[2]*
[2]Department of Statistics, Alva's College, Moodbidri- 574227, D.K. Karnataka, India.

Hardle and Tsybakov (1997) proposed a new idea regarding the specific estimation of variance functions. They separately estimated $E\left(Y^2/X=x\right)$ and $E\left(Y/X=x\right)$ and then combined them through $\mathrm{var}\left(Y/X=x\right)=E\left(Y^2/X=x\right)-E\left(Y/X=x\right)^2$. Fan and Yao (1998) suggested an approach which is asymptotically fully adaptive to the unknown conditional mean. They used the local linear method to estimate conditional mean and conditional variance. Flavio A Ziegelmann (2002) used local linear method to estimate conditional mean and in a second step, conditional variance is estimated by using local exponential estimator.

In this paper we have developed a model for volatility by introducing a function g(x) which summerises the influences of exogenous factors on volatility in addition to the parametric GARCH component. We used local linear regression method to estimate the function g(x) and GARCH component is estimated by parametric method. Finally a linear combination of these two estimators by choosing the weights which minimises the mean square error is taken as new estimator for volatility.

In the next section different models for volatility is presented and estimation procedure is highlighted. A new estimation procedure for estimation of volatility is presented in Section 3. Finally in the last section the performance of the proposed method is compared with the other existing methods on a sample of gold prices.

2. REGRESSION SETTINGS AND ESTIMATOR

2.1: Let $\left\{\left(Y_t,\ X_t\right)\right\}$ be a two dimensional strictly stationary process having the same marginal distributions as (Y, X).

Let $m(x)=E\left(Y/X=x\right)$ be the conditional mean function and $\sigma^2\left(x\right)=Var\left(Y/X=x\right)\neq 0$ the conditional volatility function. We can write the regression model as

$$Y_t=m\left(X_t\right)+\sigma\left(X_t\right)\varepsilon_t \tag{1}$$

where $E\left(\varepsilon_t/X_t\right)=0$ and $Var\left(\varepsilon_t/X_t\right)=1$

Autoregressive conditional heteroscedasticity(ARCH(p)) model for volatility has the representation

$$\sigma_t^2=\alpha_0+\alpha_1 y_{t-1}^2+\dots\dots\dots\dots\dots\dots\dots\dots\dots+\alpha_p y_{t-p}^2 \tag{2}$$

The generalized autoregressive conditional heteroscedasticity(GARCH(p, q)) model for volatility has the representation

$$\sigma_t^2=\alpha_0+\alpha_1 y_{t-1}^2+\dots+\alpha_p y_{t-p}^2+\quad\beta_1\sigma_{t-1}^2+\quad\dots\dots+\beta_q\sigma_{t-q}^2 \tag{3}$$

Bollerslev (1986) proposed GARCH (1, 1) model, which depends on long run average variance rate V_L as well as σ_{t-1} and ε_{t-1}.

The equation for this model is

$$\sigma_t^2=\gamma V_L+\alpha\varepsilon_{t-1}^2+\beta\sigma_{t-1}^2$$

Where γ is the weight assigned to V_L, α is the weight assigned to ε_{t-1}^2 and β is the weight assigned to σ_{t-1}^2.

The proposed model for volatility has the representation

$$\sigma_t^2=w_1\,g\left(x_t\right)+w_2\,\left(\alpha_0+\alpha_1 y_{t-1}^2+\dots+\alpha_p y_{t-p}^2+\beta_1\sigma_{t-1}^2+\dots\dots+\beta_q\sigma_{t-q}^2\right) \tag{4}$$

where w_1 and w_2 are the weights to be chosen optimally under mean squared error criteria satisfying $w_1+w_2=1$

2.2 SEMIPARAMETRIC ESTIMATION OF VOLATILITY

Our main interest in this paper is estimating the conditional volatility function $\sigma^2(.)$ in (1). In the proposed new method, first m(x) is estimated by local linear regression method and compute the residuals. Using this residual g(x) in (4) is estimated by local linear regression and the GARCH component is estimated using the parametric method. Finally the estimator of volatility is obtained as a linear combination of these two estimators.

2.21Estimation of $m(x_t)$ by using Local Linear Method

The regression estimator of m(x) in equation (1) is based on the local least square fitting of Kernel weighted linear regression function. The locally weighted linear regression estimator of m(x) is $\hat{\alpha}_0$. Where $\hat{\alpha}_0$ is the solution to the problem.

$$\min \sum_{t=1}^{n} \left[y_t - \alpha_0 - \alpha_1 (x - x_t) \right]^2 K_h (x - x_t) \tag{5}$$

K is the Kernel function and h is the band width.

The problem (5) is the weighted least square problem and solution for α_0 is $\hat{\alpha}_0 = \hat{m}(x).$,

where

$$\hat{m}(x) = \frac{\sum_{t=1}^{T} w_t y_t}{\sum_{t=1}^{T} w_t + \frac{1}{T^2}} \tag{6}$$

2.22 Estimation of g(x) by smoothing the squared residuals.

In the proposed new model (4) for estimating volatility, g(x) is the additional term introduced to represent the exogenous factors influencing volatility .First we obtain preliminary estimate for $\sigma^2(x)$ by smoothing the squared residuals r_t^2 , where

$$r_t^2 = \left[y_t - \hat{m}(x_t) \right]^2 = \sigma^2(x_t)\varepsilon_t^2 \tag{7}$$

By taking the conditional expectation of the squared residuals, we get $E\left[r_t^2 / X_t\right] = \sigma^2(X_t)$. The local liner regression estimator of α and β is the solution to the problem minimize L(α, β) w.r.t to α and β, where

$$L(\alpha,\beta) = \sum_{t=1}^{T} \left[r_t^2 - \alpha - \beta(x - x_t) \right]^2 K_h (x - x_t) \tag{8}$$

and $K_h(.)$ is a kernel function and h is a band width.

In the above problem $\hat{\alpha}.$ is taken as preliminary estimate of $\sigma^2(.x)$ and $\hat{\beta}$ is the estimate of the first derivative of $\sigma^2(.x)$ evaluated at x.

Taking the partial derivatives of $L(\alpha,\beta)$ with respect to both α and β and equating the derivatives to zero gives,

$$\sum_{t=1}^{T} K_h(x - x_t) r_t^2 = \alpha \sum_{t=1}^{T} K_h(x - x_t) + \beta \sum_{t=1}^{T} (x - x_t) K_h(x - x_t)$$

$$\sum_{t=1}^{T} r_t^2 (x - x_t) K_h(x - x_t) = \alpha \sum_{t=1}^{T} (x - x_t) K_h(x - x_t) + \beta \sum_{t=1}^{T} (x - x_t)^2 K_h(x - x_t)$$

Consider

$$S_{T,l} = \sum_{t=1}^{T} K_h(x - x_t)(x - x_t)^l, \qquad l = 0,1,2$$

The Prior system of equation becomes

$$\begin{bmatrix} S_{T,0} & S_{T,1} \\ S_{T,1} & S_{T,2} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T} K_h(x - x_t) r_t^2 \\ \sum_{t=1}^{T} (x - x_t) K_h(x - x_t) r_t^2 \end{bmatrix}$$

Consequently we have

$$\hat{\alpha} = \frac{S_{T,2} \sum_{t=1}^{T} K_h(x - x_t) r_t^2 - S_{T,1} \sum_{t=1}^{T} (x - x_t) K_h(x - x_t) r_t^2}{S_{T,0}\ S_{T,2} - S_{T,1}^2}$$

The numerator and denominator of the prior function can be further simplified as

$$S_{T,2} \sum_{t=1}^{T} K_h(x - x_t) r_t^2 - S_{T,1} \sum_{t=1}^{T} (x - x_t)\, K_h(x - x_t) r_t^2 = \sum_{t=1}^{T} \left[K_h(x - x_t) \left(S_{T,2} - (x - x_t)\, S_{T,1} \right] r_t^2 \right.$$

$$S_{T,0}\ S_{T,2} - S_{T,1}^2 = \sum_{t=1}^{T} K_h(x - x_t)\, S_{T,2} - \sum_{t=1}^{T} (x - x_t)\, K_h(x - x_t)\, S_{T,1}$$

$$= \sum_{t=1}^{T} K_h(x - x_t) \left[S_{T,2} - (x - x_t)\, S_{T,1} \right]$$

$$\hat{\alpha} = \frac{\sum_{t=1}^{T} w_t r_t^2}{\sum_{t=1}^{T} w_t}$$

where $w_t = K_h(x - x_t) \left[S_{T,2} - (x - x_t)\, S_{T,1} \right]$

$$K_h(x) = \frac{1}{h\sqrt{\pi}} \exp\left(-\frac{x^2}{2h^2} \right)$$

To avoid possible zero in the denominator, we use

$$\sigma^2(x) = \frac{\sum_{t=1}^{T} w_t r_t^2}{\sum_{t=1}^{T} w_t + \frac{1}{T^2}} \qquad (9)$$

This preliminary estimate of $\sigma^2(.x)$ is used to obtain the estimate of g(x). The local linear regression estimate of the regression function g(x) in equation (4) is to find a and b that minimize

$$L(a,b) = \sum_{t=1}^{T} \left[\sigma_t - a - b(x - x_t) \right]^2 K_h(x - x_t)$$

where $K_h(.)$ is a kernel function and h is a bandwidth.

The estimate of g(x) is defined as

$$\hat{a} = \frac{\sum_{t=1}^{T} \left[K_h(x - x_t) \left(S_{T,2} - (x - x_t)\, S_{T,1} \right) \right] \sigma_t}{\sum_{t=1}^{T} K_h(x - x_t) \left[S_{T,2} - (x - x_t)\, S_{T,1} \right]}$$

Let $w_t = K_h(x - x_t) \left[S_{T,2} - (x - x_t)\, S_{T,1} \right]$

where $K_h(x) = \frac{1}{h\sqrt{\pi}} \exp\left(-\frac{x^2}{2h^2} \right)$

$$\therefore \quad \hat{a} = \frac{\sum_{t=1}^{T} w_t \sigma_t}{\sum_{t=1}^{T} w_t}$$

To avoid possible zero in the denominator, we use

$$\hat{g}(x) = \frac{\sum_{t=1}^{T} w_t \sigma_t}{\sum_{t=1}^{T} w_t + \frac{1}{T^2}} \tag{10}$$

For estimating the GARCH component in model (4), we use the likelihood function.

3. COMBINED ESTIMATOR FOR FORECASTING VOLATILITY

Using the local linear kernel weighted estimator of the regression function g(x) and the parametric estimate of GARCH component a combined estimator is proposed for volatility as a weighted average of these two estimates, were weights are proportional to the size of the variance of each estimator.

The combination model of the two prediction method is

$$f = w_1\ f_1 + w_2\ f_2, \quad with \quad w_1 + w_2 = 1$$

Setting $f = \sigma(x_t),\ \ f_1 = \hat{g}(x_t),\ f_2 = \hat{\alpha}_0 + \hat{\alpha}_1\ y_{t-1}^2 + \hat{\beta}_1\ \sigma_{t-1}^2$ ~ GARCH (1 1) We get

$$\hat{\sigma}(x_t) = w1\ \hat{g}(x_t) + w2GARCH(1,1) \quad , \tag{11}$$

$$W1 + W2 = 1$$

$$W_i = \left(\sum_{t=1}^{n} e_{it}^2\right)^{-1}\left[\sum_{j=1}^{2}\left(\sum_{t=1}^{n} e_{jt}^2\right)^{-1}\right]^{-1} = 1, \quad i = 1,2$$

where

$$e_{1t} = s_t^2 - f_1$$
$$e_{2t} = s_t^2 - f_2$$

Here s_t is the standard deviation calculated from return series.

f_1, f_2 are two methods of prediction, e_{it} is absolute error when the method 'i' is used to predict the sample .

4. NUMERICAL STUDIES

The computation is carried out in 3 different stages by writing program using software package MATLAB 6.5

In the first stage the GARCH model is fitted using the return series of a gold price data from 10/08/2010 to 15/April/2013.

In the second stage tested for the presence of GARCH effect by selecting the hypothesis
H_0: No ARCH effect exists.
H_1: ARCH effect exists.

Since GARCH is Generalize ARCH, existence ARCH implies the existence of GARCH effect. Thus rejection of null hypothesis implies that the existence of GARCH effect.

Finally volatility is estimated by using the proposed method. This estimator is compared with the existing estimators and also with the method proposed by Fan and Yao (1998).

Using the function GARCHFIT gives the estimates of the coefficients of the GARCH model as follows:

Coefficients	values
α_0	3.3756e+011
α_1	0.8129
β_1	0.0396

Mean squared error based on GARCH model is 4.28e+010 and for the model (7) which is proposed by Fan and Yao (1998) is 1.004. Mean squared error obtained from the proposed model (4) is 0.0338 which shows the superiority of the proposed estimator.

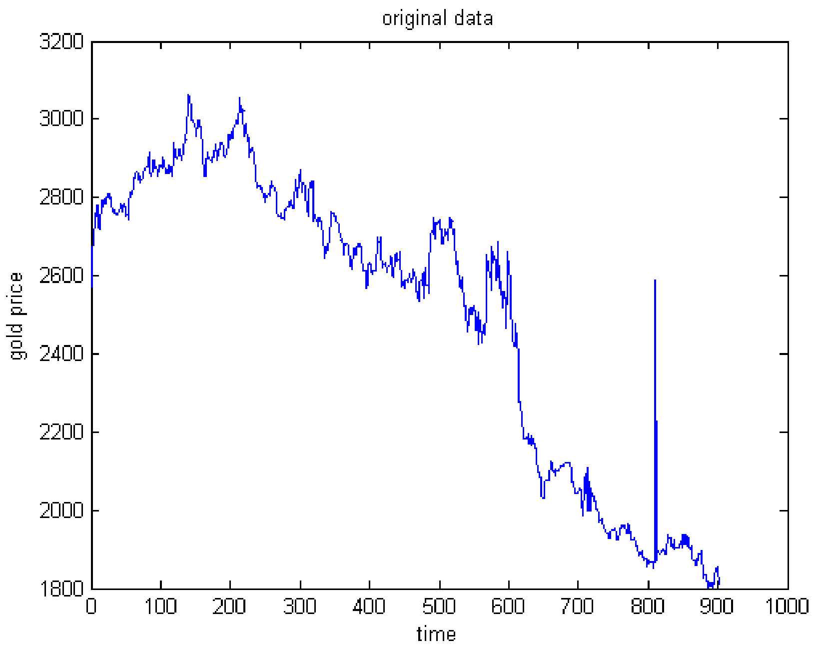

Fig1: Gold price data from 10/August/2010 to 15/April/2013

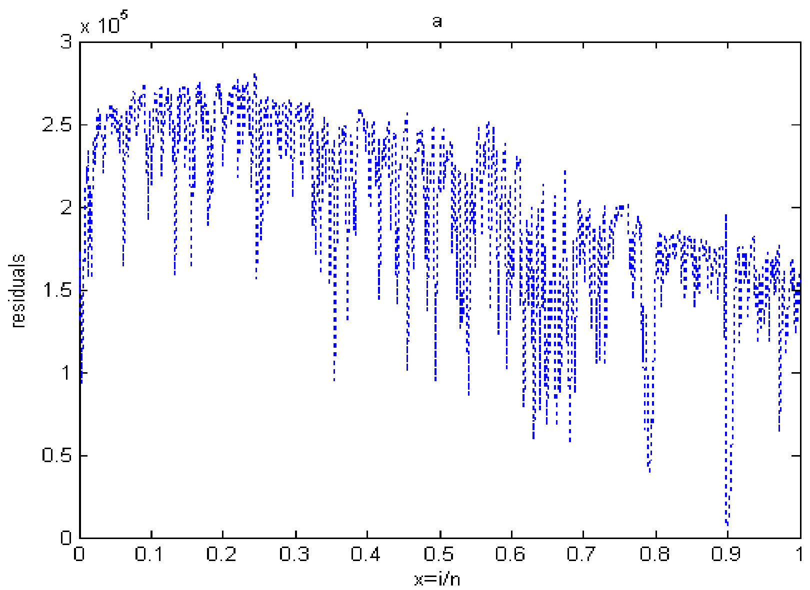

Fig 2: Plot of residuals obtained from the existing model

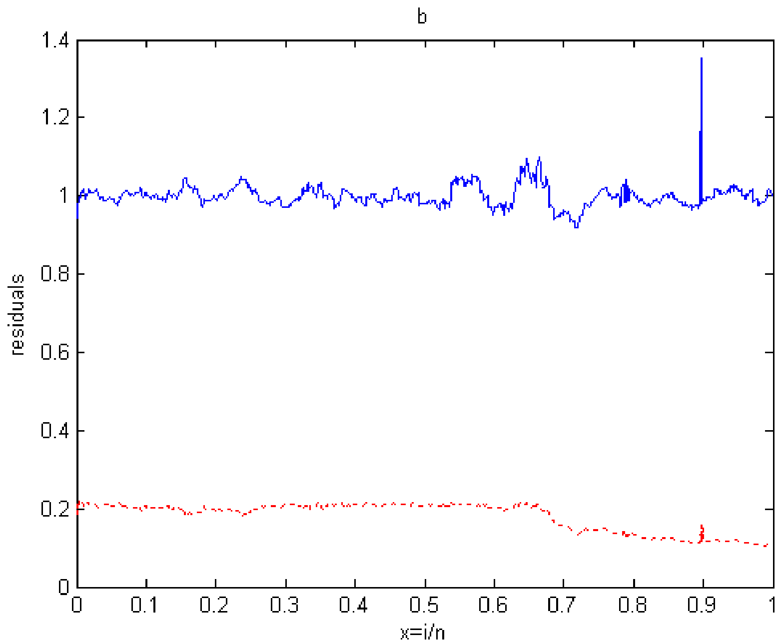

Fig 3: comparing the plot of residuals obtained from the proposed model (red) with the model which is proposed by Fan and Yao (1998) (blue)

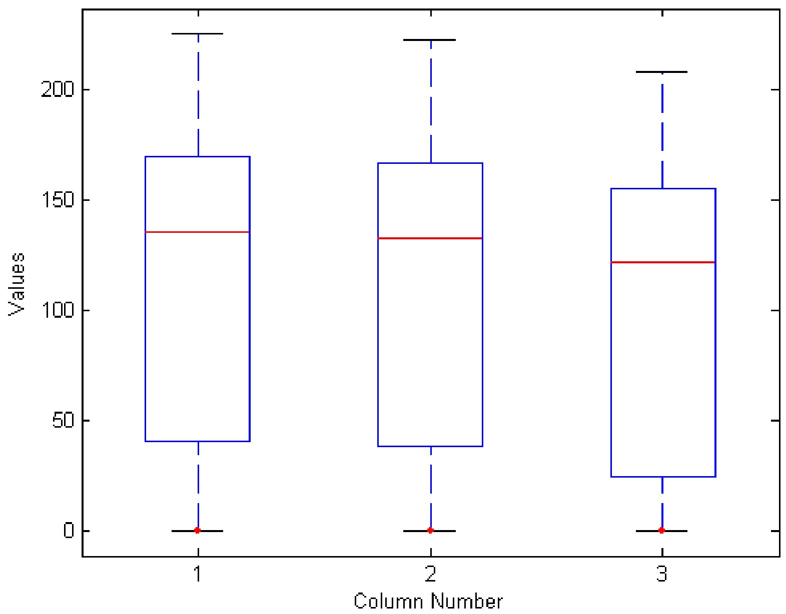

Fig 4*:* Box plots of the mean absolute deviations 1) existing model 2) Fan and Yao's (1998) model 3) Proposed model

5. CONCLUSION

A semiparametric approach is proposed to estimate volatility and it is observed that this procedure provides an estimator for the return series with minimum mean square error compared with the estimator based on GARCH model. Residual plot clearly shows remarkable superiority of the proposed estimator. The residuals are much smaller compare to the residuals based on estimators obtained using the existing method. The box plots of the mean absolute deviations show superiority of the proposed estimator compared with the existing methods.

REFERENCES

1. Fan, J & I.Gijbels, Local polynomial modeling and its applications, London: Chapman and Halll (1996).
2. Bollerslev, Generalised autoregressive conditional heteroskedasticity, Journal of Econometrics. Volume 31 (1986), 307-327.
3. Fan, J &Q. Yao, Efficient estimation of Conditional Variance functions in stochastic regression, Biometrika. 85 (1998), 645-660.

4. Flavio A Ziegelmann, Nonparametric estimation of volatility functions: The local exponential estimator, Econometric Theory. Volume18 (2002), pp. 985-991.
5. Hardle , W.&A. Tsybakov, Local polynomial estimators of the volatility function in nonparametric autoregression, Journal of Econometrics . 81 (1997), 233-242.
6. Hart, J.D, Nonparametric smoothing and lack-of -fit tests, Springer, New York, (1997).
7. Hideaki Shimazaki. Shigera Shinomoto, Kernel bandwidth optimization in spike rate estimation, J.Compute Neurosci .Volume 29 (2010), 171-182.
8. Jibo Chen, Xiaorong Huang, Jie LI, Guizhi Wang, Prediction of energy consumption in Jiangsu based on combination model, 2010 Third International Conference on Information and Computing.
9. Lijian Yang, A semiparametric GARCH model for foreign exchange volatility, Journal of Econometrics. Volume 130(2005), pp. 365-384.
10. Loader C, Local regression and likelihood, Springer, New York, (1999).
11. S Ruey Tsay, Analysis of financial time Series, 2nd Edition, John Wiley & Sons, (2005).
12. Silverman, B.W, Density estimation for statistics and data analysis, Lonson: Chapmann and Hall, (1986).
13. Walter Enders, Applied Econometric Time Series, 2nd Edition, John Wiley & Sons, (2004).
14. Wand, M &M. Jones, Kernel smoothing, London: Chapmann and Hall, (1995).

Source of support: Nil, Conflict of interest: None Declared