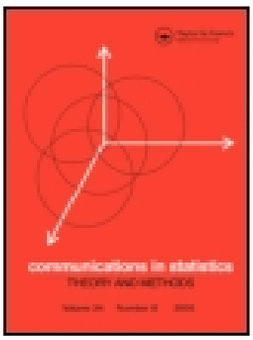
Made available by Hasselt University Library in https://documentserver.uhasselt.be

Comparison of spline estimator at various levels of autocorrelation in smoothing spline non parametric regression for longitudinal data

Peer-reviewed author version

Fernandes, Adji Achmad Rinaldo; JANSSEN, Paul; Sa'adah, Umu; Solimun; Effendi, Achmad; Nurjanna & Amaliana, Luthfatul (2018) Comparison of spline estimator at various levels of autocorrelation in smoothing spline non parametric regression for longitudinal data. In: COMMUNICATIONS IN STATISTICS-THEORY AND METHODS, 47(21), p. 5265-5285.

DOI: 10.1080/03610926.2017.1388404
Handle: http://hdl.handle.net/1942/28461



Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

Comparison of Spline Estimator at Various Levels of Autocorrelation in Smoothing Spline Nonparametric Regression For Longitudinal Data

Adji Achmad Rinaldo Fernandes, Paul Janssen, Umu Sa'adah, Solimun, Achmad Effendi, Nurjannah & Luthfatul Amaliana

To cite this article: Adji Achmad Rinaldo Fernandes, Paul Janssen, Umu Sa'adah, Solimun, Achmad Effendi, Nurjannah & Luthfatul Amaliana (2017): Comparison of Spline Estimator at Various Levels of Autocorrelation in Smoothing Spline Nonparametric Regression For Longitudinal Data, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2017.1388404

To link to this article: http://dx.doi.org/10.1080/03610926.2017.1388404

Accepted author version posted online: 09 Oct 2017.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=lsta20

Download by: [University of Virginia, Charlottesville] **Date:** 09 October 2017, At: 10:55

ACCEPTED MANUSCRIPT

Check for updates

Comparison of Spline Estimator at Various Levels of Autocorrelation in Smoothing Spline Nonparametric Regression For Longitudinal Data

Adji Achmad Rinaldo Fernandes[1], Paul Janssen[2], Umu Sa'adah[1], Solimun[1], Achmad Effendi[1], Nurjannah[1], and Luthfatul Amaliana[1]

[1]Statistics Department, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Indonesia

[2]Center for Statistics, Universiteit Hasselt, Belgium

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

ABSTRACT

The purpose of this research are: (1) to obtain spline function estimation in nonparametric regression for longitudinal data with and without considering the autocorrelation between data of observation within subject, (2) to develop the algorithm that generates simulation data with certain autocorrelation level based on size of sample (N) and error variance (EV), (3) to establish shape of spline estimator in nonparametric regression for longitudinal data to simulation with various level of autocorrelation, as well as compare DM and TM approaches in predicting spline estimator in the data simulation with different of autocorrelation observational data on within subject. The results of the application as follows: (a) Implementation of smoothing spline with PWLS approach with or without consideration of autocorrelation in general (in all sizes and all

ACCEPTED MANUSCRIPT

error variances levels) provides significantly different spline estimator when the autocorrelation level > 0.8. (b) spline estimator in nonparametric regression smoothing spline with PWLS approach with or without consideration of autocorrelation in all sizes observations showed significantly different results when the autocorrelation level > 0.8, whereas the size of a small observation when the level of autocorrelation > 0.7, the size of the observation was and the size of a large observation when the level of autocorrelation> 0.8. (c) spline estimator in nonparametric regression smoothing spline with PWLS approach with or without consideration of autocorrelation in all the error variance give significantly different results when the autocorrelation level> 0.8, whereas the small variance error when autocorrelation level> 0.7, error variance was and the error variance is greater when the level of autocorrelation of 0.8.

Keywords

Smoothing Spline, Longitudinal, Autocorrelation, Nonparametric Regression

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

1. Introduction

In the statistical method, the pattern of response (*y*) relation on the predictors (*x*) can be determined using regression analysis. The pattern of these relationships can be displayed in graphical form, including linear, quadratic, cubic, exponential, and other form. Patterns of response (y) with the predictor (x) relation can be determined using two approaches, namely parametric and nonparametric approach. Parametric approach assumes that the shape of the curve $\underset{\sim}{f}$ is known shape and assuming the form of a linear relationship between the response and predictor must be met. If the assumptions are not met then the nonparametric approach can be used as an alternative in which the shape of the curve $\underset{\sim}{f}$ is unknown. Nonparametric regression is flexible because the shape of the curve is estimated adjust the pattern of the data, while the data to adjust the parametric regression curve shape is already known [1], [2], [15], [4], [5].

Nonparametric regression can be approached by spline. Spline is an approach that could follow the pattern of changing relationships in subgroups specified interval and has a very flexible properties [1], [6]. Spline is divided into two type such as truncated spline and smoothing spline. Basically, truncated spline consider their point of knots in determining the optimal point while Smothing spline not need to specify a point knots. Estimates of smoothing spline based on criteria curve smoothing model and size that has been set by the smoothing parameter (λ).

Data obtained from observations which is an essential component needed in research. Much of the data obtained from the study is longitudinal data, derived from observations made on some subjects which are independent, observed repeatedly (repeated measurement) within a certain time. The nature of the data that is independent between subject, and dependent on the

 ACCEPTED MANUSCRIPT

observation within subject [2], [6], [3]. In longitudinal data often autocorrelation between observations within subject because of the time difference of observations. In general smoothing spline approach Penalized Least Square (PLS). However, the autocorrelation cause PLS unusable so that alternatives can accommodate autocorrelation is Penalized Weighted Least Square (PWLS). PWLS nonparametric regression approach required the addition of such weighting inverse of variance-covariance matrix which is symbolized $\mathbf{\Sigma}^{-1}$ ([8], [9], 8], [11]).

Application of smoothing spline on longitudinal data which examines the growing weight of the baby for 24 months ([2],[6]). However, these studies do not take into consideration the level of autocorrelation between observations within subject so that in this study will examine the autocorrelation between observations. The level of autocorrelation in each data are not always the same. P value is used as a reference level of autocorrelation varying ranges. Variations in the value of autocorrelation investigated using PWLS approach by considering the autocorrelation and without considering the autocorrelation. To obtain the autocorrelation of different designs used data simulation because it is very difficult to obtain data in accordance with the above conditions.

The estimation of nonparametric smoothing spline regression curves at various levels will be studied further autocorrelation using PWLS approach by considering the autocorrelation and without considering the autocorrelation. Comparison of nonparametric regression curve through PWLS approach by considering the autocorrelation (DM) and PWLS without considering autocorrelation (TM) to determine the appropriate approach based on the degree of autocorrelation in the data [18], [19]. P value close to 0 indicate insignificant autocorrelation between subject while the value of autocorrelation, close to 1 indicates a significance

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

autocorrelation between observations. If the level of autocorrelation smaller then expected longitudinal data can be analyzed as a cross section so that the analysis of the smoothing spline more easily applied mathematically.

For purposes of a more complex, this study uses simulation data that is able to give a complete picture of the overall level of autocorrelation at various levels. Based on this background, the purpose of this research are as follows: (1) to obtain spline function estimation in nonparametric regression for longitudinal data with and without considering the autocorrelation between data of observation within subject, (2) to develop the algorithm that generates simulation data with certain autocorrelation level based on size of sample (N) and error variance (EV), (3) to establish shape of spline estimator in nonparametric regression for longitudinal data to simulation with various level of autocorrelation, as well as compare DM and TM approaches in predicting spline estimator in the data simulation with different of autocorrelation observational data on within subject.

2. Spline in Nonparametric Regression for Longitudinal Data

Nonparametric regression model in longitudinal data, developed by N subjects were observed repeatedly (repeated measurement) in the T period. Nonparametric regression model for longitudinal data have a difference with a cross-section, which is located on the observation between subjects assumed to be independent of each other, but between observations within subject is dependent [2], [6], [12]. Many researchers have developed a spline estimator in nonparametric regression model for longitudinal data ([13], [14], [17], [15], [16]). The relationship between predictors of response to longitudinal data involving N subject to the T observations of each subject, following the regression model as follows:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

 ACCEPTED MANUSCRIPT

$$y_{it} = f_i(x_{it}) + \varepsilon_{it}; \;\; i = 1, 2, ..., N; \;\; t = 1, 2, ..., T. \tag{1}$$

with

y_{it} : Response on the tth-time i^{th}-subject

x_{it} : Predictors on the tth-time i^{th}-subject

f_i : Regression curve predictor relationship with the response on on the ith-subject

N: number of subjects,

T: The number of observations of each subject,

ε_{it} : Random error on the tth-time i^{th}-subject,

The regression model in equation (1) to include as a regression curve f_i that accommodates autocorrelation within subject. Random error $\underset{\sim}{\varepsilon} = (\varepsilon_{11}, \varepsilon_{12} ..., \varepsilon_{1T}, \varepsilon_{21}, \varepsilon_{22} ..., \varepsilon_{2T}, ..., \varepsilon_{N1}, \varepsilon_{N2} ..., \varepsilon_{NT})'$ is assumed *NT*-variat normal distribution, with mean E ($\underset{\sim}{\varepsilon}$) = (NT-vector) and the variance-covariance matrix Var ($\underset{\sim}{\varepsilon}$) = $\mathbf{\Sigma}$ (matrix measuring NT × NT) as follows ([8], [9], 8], [11]):

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \sigma_{1(1,2)} & \cdots & \sigma_{1(1,T)} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \sigma_{1(2,1)} & \sigma_{12}^2 & \cdots & \sigma_{1(2,T)} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1(T,1)} & \sigma_{1(T,1)} & \cdots & \sigma_{1T}^2 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{21}^2 & \sigma_{2(1,2)} & \cdots & \sigma_{2(1,T)} & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{2(1,2)} & \sigma_{22}^2 & \cdots & \sigma_{2(2,T)} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \sigma_{2(T,1)} & \sigma_{2(T,1)} & \cdots & \sigma_{2T}^2 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & \sigma_{N1}^2 & \sigma_{N(1,2)} & \cdots & \sigma_{N(1,T)} \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & \sigma_{N(1,2)} & \sigma_{N2}^2 & \cdots & \sigma_{N(1,T)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 & \cdots & \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & \sigma_{N(T,2)} & \sigma_{N(T,2)} & \cdots & \sigma_{NT}^2 \end{bmatrix}_{(NT)\times(NT)} . \tag{2}$$

The matrix $\mathbf{\Sigma}$ can be simplified into sub-matrices $\mathbf{\Sigma}_i$ and $\mathbf{0}$.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Sigma}_N \end{bmatrix}_{(NT)\times(NT)}$$

Sub-matrix $\mathbf{\Sigma}_i$ and $\mathbf{0}$ measuring T × T, are presented as follows:

$$\mathbf{\Sigma}_i = \begin{bmatrix} \sigma_{i1}^2 & \sigma_{i(1,2)} & \cdots & \sigma_{i(1,T)} \\ \sigma_{i(2,1)} & \sigma_{i2}^2 & \cdots & \sigma_{i(2,T)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{i(T,1)} & \sigma_{i(T,2)} & \cdots & \sigma_{iT}^2 \end{bmatrix}_{T\times T} \quad \text{dan } \mathbf{0} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{T\times T}$$

Elements outside the diagonal $\sigma_{i(1,2)}, \sigma_{i(2,1)}, ..., \sigma_{i(T-1,T)}, \sigma_{i(T,T-1)}$, namely the sub-matrix $\mathbf{\Sigma}_i$ is a random error covariance between observations within subject. This covariance can be not 0, which accommodates the correlation within subject. On the other hand, the sub-matrix that $\mathbf{0}$ is the matrix of all elements of value 0 states that the covariance between observations between subjects are independent.

Spline approach generally specify f_i in equation (1) in the form of regression curve shape is unknown, but f_i it's assumed smooth, in the sense of space is contained in a particular function, especially Sobolev spaces or written ([8], [9], 8], [11]), $f_i \in \mathrm{W}_2^m[a_i, b_i]$ where:

$$\mathrm{W}_2^m[a_i, b_i] = \left\{ f_i : f_i,\ f_i^{(1)}, ..., f_i^{(m-1)} \text{ absolute continue } ; \int_{a_i}^{b_i} [\, f_i^{(m)}(x_{it})]^2 dx_{it} < \infty \right\}, \tag{3}$$

with m stating order polynomial spline. Completion curve estimation regression f_i for longitudinal data in equation (1) using Penalized Weighted Least Square (PWLS) involving weights in the form of inverse variance-covariance matrix of random errors symbolized $\mathbf{\Sigma}$ as has

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

been described in equation (2). To obtain the estimates of the regression curve f_i using the optimization PWLS namely the completion of optimization as follows [1]:

$$\underset{f_i \in W_2^m[a_i,b_i],i=1,2,...,N}{\text{Min}} \left\{ M^{-1}(\underset{\sim}{y}-\underset{\sim}{f})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})+\sum_{i=1}^{N}\lambda_i\int_{a_i}^{b_i}(f_i^{(m)}(x_{it}))^2 dx_{it} \right\}, \tag{4}$$

where $\underset{\sim}{y}=(y_{11},y_{12},...,y_{1T},y_{21},y_{22},...,y_{2T},...,y_{N1},y_{N2},...,y_{NT})^{\mathrm{T}}$, and

$$\underset{\sim}{f}=(f_1(x_{11}),f_1(x_{12}),...,f_1(x_{1T}),f_2(x_{21}),f_2(x_{22}),...,$$
$$f_2(x_{2T}),...,f_N(x_{N1}),f_N(x_{N2}),...,f_N(x_{NT}).$$

PWLS optimization in equation (4) using the smoothing parameter λ_i, as a controller between the goodness of fit (first segment) and a roughness penalty (second segment)

3. Results and Discussion

3.1. Spline Function Estimation in Nonparametric Regression for Longitudinal Data

The first goal of the study is obtain spline function estimation in nonparametric regression for longitudinal data with and without considering the autocorrelation between data of observation within subject, it's presented in Theorem 1 as follow:

Theorem 1:

When given the data pairs following the nonparametric regression model involves a single predictor on longitudinal data that meets the form of nonparametric regression functions for longitudinal data as presented in equation (1), assuming

$$\mathrm{E}(\underset{\sim}{\varepsilon})=\underset{\sim}{0},\ \mathrm{Var}(\underset{\sim}{\varepsilon})=\mathbf{\Sigma},$$

then the spline estimator that minimizes PWLS

is $\hat{\underset{\sim}{f}}_\lambda=\mathbf{A}_\lambda \underset{\sim}{y}$, (5)

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

with:

$$\mathbf{A}_{\lambda} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}]. \tag{6}$$

$$\hat{\mathbf{U}} = \hat{\mathbf{\Sigma}}^{-1}\mathbf{V} + M\mathbf{\Lambda}. \tag{7}$$

Proof:

Considering the equation that is function $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c}$, then the nonparametric regression model (1) can be stated as[8]:

$$\underset{\sim}{y} = \underset{\sim}{f} + \underset{\sim}{\varepsilon} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c} + \underset{\sim}{\varepsilon}.$$

where **T** is $(NT) \times (Nm)$ matrix as follow:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{T}_N \end{bmatrix}, \tag{8}$$

where

$$\mathbf{T}_i = \begin{pmatrix} \langle \eta_{i1}, \phi_{i1} \rangle & \langle \eta_{i1}, \phi_{i2} \rangle & \cdots & \langle \eta_{i1}, \phi_{im} \rangle \\ \langle \eta_{i2}, \phi_{i1} \rangle & \langle \eta_{i2}, \phi_{i2} \rangle & \cdots & \langle \eta_{i2}, \phi_{im} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{iT}, \phi_{i1} \rangle & \langle \eta_{iT}, \phi_{i2} \rangle & \cdots & \langle \eta_{iT}, \phi_{im} \rangle \end{pmatrix}, \tag{9}$$

$$\langle \eta_{it}, \phi_{ij} \rangle = \frac{x_{it}^{\;j-1}}{(j-1)!}, with\, t = 1, 2, ..., T;\; j = 1, 2, ..., m$$

$$\underset{\sim}{d} = (\underset{\sim}{d}_1', \underset{\sim}{d}_2', \cdots, \underset{\sim}{d}_N')', \quad \underset{\sim}{d}_i' = (d_{i1}, d_{i2}, ..., d_{im}),$$

V is *(NT)* $\times$ *(NT)*-sized matrix as follow:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_N \end{bmatrix}, \tag{10}$$

$$\text{where } \mathbf{V}_i = \begin{pmatrix} \langle \xi_{i1}, \xi_{i1} \rangle & \langle \xi_{i1}, \xi_{i2} \rangle & \cdots & \langle \xi_{i1}, \xi_{iT} \rangle \\ \langle \xi_{i2}, \xi_{i1} \rangle & \langle \xi_{i2}, \xi_{i2} \rangle & \cdots & \langle \xi_{i2}, \xi_{iT} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \xi_{iT}, \xi_{i1} \rangle & \langle \xi_{iT}, \xi_{i2} \rangle & \cdots & \langle \xi_{iT}, \xi_{iT} \rangle \end{pmatrix}, \tag{11}$$

$$\langle \xi_{it}, \xi_{is} \rangle = \langle \xi_{is}, \xi_{it} \rangle = \int_a^b \frac{(x_{it} - u)_+^{m-1} (x_{is} - u)_+^{m-1}}{\left((m-1)!\right)^2} du, t = 1, 2, ..., T; s = 1, 2, ..., T$$

$\underset{\sim}{c}$ is vector with NT-sized

$\underset{\sim}{c} = (\underset{\sim}{c}_1', \underset{\sim}{c}_2', \cdots, \underset{\sim}{c}_N')'$, where $\underset{\sim}{c}_i' = (c_{i1}, c_{i2}, ..., c_{iT})$.

Equation (11) shows that the value of t is depend on t, t-1, t-2, …., and not depend on t+1, t+2, …. Thats why the $\mathbf{V}_i$ as lower-triangle matrix. Nonparametric regression analysis is conducted to get estimator of regression curve f. To get the estimation, Reproducing Kernel Hilbert Space (RKHS) is used. The purpose is to get estimasion of f that meets PWLS optimization[9]:

$$\min_{\substack{f_i \in \mathcal{H} \\ i=1,2,..,N}} \left\{ \left\| \mathbf{\Sigma}^{-\frac{1}{2}} \underset{\sim}{\varepsilon} \right\|^2 \right\} = \min_{\substack{f_i \in \mathcal{H} \\ i=1,2,..,N}} \left\{ \left\| \mathbf{\Sigma}^{-\frac{1}{2}} (\underset{\sim}{y} - \underset{\sim}{f}) \right\|^2 \right\}, \tag{12}$$

with restricted:

$$\| f_i \|^2 < \gamma_i, \quad \gamma_i \geq 0. \tag{13}$$

Then, space function $\mathcal{H} = \mathrm{W}_2^m[a_i, b_i]$ used is order-2 Sobolev space defined as follow:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\mathrm{W}_2^m[a_i,b_i]=\left\{ f_i\colon \int_{a_i}^{b_i}[\,f_i^{(m)}(x_{it})]^2\,dx_{it}<\infty \right\},$$

Where $a_i \le x_{it} \le b_i$ and i = 1,2,…,N. Based on the space, norm of each

$f_i \in \mathrm{W}_2^m[a_i,b_i]$

is described as follow:

$$\|\,f_i\,\|^2=\int_{a_i}^{b_i}[\,f_i^{(m)}(x_{it})]^2\,dx_{it}.$$

Optimization with restricted in equation (12) can be stated as:

$$\min_{\substack{f_i\in \mathrm{W}_2^m[a_i,b_i]\\ i=1,2,..,N}}\left\{\left\|\,\mathbf{\Sigma}^{-\frac{1}{2}}\underset{\sim}{\varepsilon}\,\right\|^2\right\}=\min_{\substack{f_i\in \mathrm{W}_2^m[a_i,b_i]\\ i=1,2,..,N}}\left\{\left\|\,\mathbf{\Sigma}^{-\frac{1}{2}}(\underset{\sim}{y}-\underset{\sim}{f})\,\right\|^2\right\} \tag{14}$$

With restricted in equation (15) into:

$$\int_{a_i}^{b_i}[\,f_i^{(m)}(x_{it})]^2\,dx_{it}<\gamma_i,\quad \gamma_i\ge 0. \tag{15}$$

Weighting optimization (14) with equivalent restricted (15) by solving Penalized Weighted Least Square (PWLS) optimization:

$$\min_{\substack{f_i\in \mathrm{W}_2^m[a_i,b_i],\\ i=1,2,...,N}}\left\{M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)+\sum_{i=1}^{N}\lambda_i\int_{a_i}^{b_i}\left[\,f_i^{(m)}\left(x_{it}\right)\right]^2dx_{it}\right\}, \tag{16}$$

where $M=NT$ and λ_i is smoothing parameter controlling between Goodness of fit:

$$M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right), \tag{17}$$

and penalty:

 ACCEPTED MANUSCRIPT

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\sum_{i=1}^{N}\lambda_i\int_{a_i}^{b_i}\left[f_i^{(m)}\left(x_{it}\right)\right]^2 dx_{it} \quad . \tag{18}$$

To solve optimization in equation (16) with penalty component:

$$\sum_{i=1}^{N}\lambda_i\int_{a_i}^{b_i}\left[f_i^{(m)}\left(x_{it}\right)\right]^2 dx_{it} = \underset{\sim}{c}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{V}\underset{\sim}{c} \quad , \tag{19}$$

The equation (19) proof by Fernandes [17]

where $\mathbf{\Lambda} = \begin{bmatrix} \lambda_1\mathbf{I}_T & \mathbf{0}_T & \cdots & \mathbf{0}_T \\ \mathbf{0}_T & \lambda_2\mathbf{I}_T & \cdots & \mathbf{0}_T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_T & \mathbf{0}_T & \cdots & \lambda_N\mathbf{I}_T \end{bmatrix}$.

Using $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c}$ as reference, Goodness of fit in PWLS optimization (15) can be stated as:

$$M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right) = M^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right). \tag{20}$$

Solving PWLS optimization by combining goodness of fit (20) and penalty (19), can be described as:

$$\min_{\substack{f_i\in \mathrm{W}_2^m[a_i,b_i],\\ i=1,2,\ldots,N}}\left\{M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)+\sum_{i=1}^{N}\lambda_i\int_{a_i}^{b_i}\left[f_i^{(m)}\left(x_{it}\right)\right]^2 dx_{it}\right\}$$

$$= \min_{\substack{\underset{\sim}{c}\in\Re^{NT}\\ \underset{\sim}{d}\in\Re^{Nm}}}\left\{M^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)+\underset{\sim}{c}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{V}\underset{\sim}{c}\right\}$$

$$= \min_{\substack{\underset{\sim}{c}\in\Re^{NT}\\ \underset{\sim}{d}\in\Re^{Nm}}}\left\{\left(\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)+\underset{\sim}{c}^{\mathrm{T}}M\mathbf{\Lambda}\mathbf{V}\underset{\sim}{c}\right)M^{-1}\right\}$$

$$= \min_{\substack{\underset{\sim}{c}\in\Re^{NT}\\ \underset{\sim}{d}\in\Re^{Nm}}}\left\{\left[(\underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}-\underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}-\underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}-\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\right.\right.$$

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}-\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}+\underset{\sim}{c}^{\mathrm{T}}M\mathbf{\Lambda}\mathbf{V}\underset{\sim}{c}\Big]M^{-1}\Big\}$$

$$=\min_{\substack{\underset{\sim}{c}\in\Re^{NT}\\ \underset{\sim}{d}\in\Re^{Nm}}}\Big\{\Big[(\underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}-2\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}-2\underset{\sim}{c}^{\mathrm{T}}\,\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}+$$

$$+\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\underset{\sim}{c}^{\mathrm{T}}\left(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}+M\mathbf{\Lambda}\mathbf{V}\right)\underset{\sim}{c}\Big]M^{-1}\Big\}$$

$$=\min_{\substack{\underset{\sim}{c}\in\Re^{NT}\\ \underset{\sim}{d}\in\Re^{Nm}}}\{\ \mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})\ \}. \tag{21}$$

Solving optimization (21) is obtained by conducting partial derivative $\mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})$ by $\underset{\sim}{c}$ and $\underset{\sim}{d}$, then the result equals to zero. The partial derivative is presented as follow:

$$\frac{\partial\,\mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})}{\partial\underset{\sim}{c}}=\underset{\sim}{0},$$

and the result is:

$$-2\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+2\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+2\left(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}+M\mathbf{\Lambda}\mathbf{V}\right)\hat{\underset{\sim}{c}}=\underset{\sim}{0}.$$

$$\mathbf{V}^{\mathrm{T}}\left\{-\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+[\mathbf{\Sigma}^{-1}\mathbf{V}+M\mathbf{\Lambda}\mathbf{I}]\hat{\underset{\sim}{c}}\right\}=\underset{\sim}{0}.$$

$$-\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+[\mathbf{\Sigma}^{-1}\mathbf{V}+M\mathbf{\Lambda}\mathbf{I}]\hat{\underset{\sim}{c}}=\underset{\sim}{0}. \tag{22}$$

When matrix $\mathbf{U}$ is presented as:

$$\mathbf{U}=\mathbf{\Sigma}^{-1}\mathbf{V}+M\mathbf{\Lambda}.$$

equation (23) can be stated as:

$$-\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\mathbf{U}\hat{\underset{\sim}{c}}=\underset{\sim}{0}.$$

$$\mathbf{U}\hat{\underset{\sim}{c}}=\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}) \tag{23}$$

Equation (24) is doubled from the left with $\mathbf{U}^{-1}$ and the following equation is ontained:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\hat{\underset{\sim}{c}} = \mathbf{U}^{-1}\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}\underset{\sim}{d}) \tag{24}$$

Furthermore, partial derivative:

$$\frac{\partial \mathrm{Q}(\underset{\sim}{c}, \underset{\sim}{d})}{\partial \underset{\sim}{d}} = \underset{\sim}{0}$$

results in:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\hat{\underset{\sim}{c}} = \underset{\sim}{0}$$

Elaboration of eqaution (17) results in the following equations:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\{\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}\hat{\underset{\sim}{d}})\} = \underset{\sim}{0}$$

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{T}^{\mathrm{T}}[\mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{U}^{-1}]\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}\hat{\underset{\sim}{d}}) = \underset{\sim}{0} \tag{25}$$

Considering $\mathbf{U} = \mathbf{\Sigma}^{-1}\mathbf{V} + M\mathbf{\Lambda}\mathbf{I}$, then $\mathbf{V} = \mathbf{\Sigma}(\mathbf{U} - M\mathbf{\Lambda}\mathbf{I})$, as the consequence, the result is the following equations:

$$\mathbf{V}\mathbf{U}^{-1} = \mathbf{\Sigma}(\mathbf{U} - M\mathbf{\Lambda}\mathbf{I})\mathbf{U}^{-1}$$

$$\mathbf{V}\mathbf{U}^{-1} = \mathbf{\Sigma}(\mathbf{I} - M\mathbf{\Lambda}\mathbf{U}^{-1}).$$

Reduplicating the equation above with $\mathbf{\Sigma}^{-1}$ resulting in:

$$\mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{U}^{-1} = \mathbf{I} - M\mathbf{\Lambda}\mathbf{U}^{-1}.$$

The equation is substituted in equation (24) resulting in:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{T}^{\mathrm{T}}[\mathbf{I} - M\mathbf{\Lambda}\mathbf{U}^{-1}]\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}\hat{\underset{\sim}{d}}) = \underset{\sim}{0}$$

When the equation above is elaborated further, the result is:

$$-M\mathbf{\Lambda}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + M\mathbf{\Lambda}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} = \underset{\sim}{0}$$

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$M\mathbf{\Lambda}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} = M\mathbf{\Lambda}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y}.$$

Both segments of the equation are reduplicated with $(M\mathbf{\Lambda})^{-1}$ and then simplified resulting in:

$$\hat{\underset{\sim}{d}} = \left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y}. \tag{26}$$

Equation (24) is substituted into equation (26) resulting in:

$$\hat{\underset{\sim}{c}} = \mathbf{U}^{-1}\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}[\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y}])$$

$$= \mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}]\underset{\sim}{y}. \tag{27}$$

Based on equation (26) and (27), estimator for nonparametric regression curve for longitudinal data involving single predictor as follows:

$$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \begin{bmatrix} \hat{\underset{\sim}{f}}_{1,\lambda_1} \\ \hat{\underset{\sim}{f}}_{2,\lambda_2} \\ \vdots \\ \hat{\underset{\sim}{f}}_{N,\lambda_N} \end{bmatrix} = \mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{V}\hat{\underset{\sim}{c}}$$

$$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{V}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}]\underset{\sim}{y}$$

$$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \{\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1} + \mathbf{V}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}]\}\underset{\sim}{y}$$

$$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}, \tag{28}$$

where

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1} + \mathbf{V}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}].$$

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

 ACCEPTED MANUSCRIPT

Error variance-covariance matrix $\hat{\mathbf{\Sigma}}$ will be presented in the next section (Theorem 2), so Theorem 1 uses $\hat{\mathbf{\Sigma}}$ as well as $\hat{\mathbf{U}} = \hat{\mathbf{\Sigma}}^{-1}\mathbf{V} + M\mathbf{\Lambda}$. resulting in:

$$\hat{\underset{\sim}{d}} = \left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\underset{\sim}{y}.$$

$$\hat{\underset{\sim}{c}} = \hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}]\underset{\sim}{y}.$$

$$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y},$$

where

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}].$$

Based on the the theorem above, the equation $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c}$ with spline function estimation considering the autocorrelation (DM) is $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ where

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}],$$

And to estimate the spline function without considering the autocorrelation (TM) is equivalent $\hat{\mathbf{\Sigma}} = \mathbf{I}$ to $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ where

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}]. \qquad \blacksquare$$

Theorem 2

The weighted using Error variance-covariance matrix for nonparametric regression longitudinal data model using maximum likelihood is as follow:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Sigma}}_{NN} \end{bmatrix} \tag{29}$$

With

$$\hat{\mathbf{\Sigma}}_{ij} = \frac{(\underset{\sim}{y}_i - \hat{\underset{\sim}{f}}_i)(\underset{\sim}{y}_j - \hat{\underset{\sim}{f}}_j)'}{T},$$

Proof:

The studies related to single-response nonparametric regression model have been conducted extensively. The researchers in general assumed variance-covariance matrix from the random error is unknown/ unidentified. As the effect, one should conduct estimation for the variance-covariance matrix from the random error in single-response nonparametric regression model. In order to do so, Maximum Likelihood Estimator (MLE) method is used.

When it is assumed that $\hat{\underset{\sim}{\varepsilon}} = \underset{\sim}{y} - \hat{\underset{\sim}{f}}$ is the result of normally distributed random sample of M-variat (M = 3T), and mean of E($\underset{\sim}{\varepsilon}$) = $\underset{\sim}{0}$ (M-sized vector) and variance-covariance matrix of Var($\underset{\sim}{\varepsilon}$) = $\mathbf{\Sigma}$ (M × M-sized matrix), combined density function from each observation is obtained from normal marginal density. It is as follow:

$$\left\{ \begin{matrix} \text{Joint } density \\ \text{of } \underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2, \underset{\sim}{\varepsilon}_3 \end{matrix} \right\} = L(\underset{\sim}{f}, \mathbf{\Sigma} | \underset{\sim}{y}) = \prod_{t=1}^{T} \left\{ \frac{1}{(2\pi)^{M/2} |\mathbf{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} (\underset{\sim}{y} - \underset{\sim}{f})' \mathbf{\Sigma}^{-1} (\underset{\sim}{y} - \underset{\sim}{f}) \right) \right\}$$

$$L(\underset{\sim}{f}, \mathbf{\Sigma} | \underset{\sim}{y}) = \frac{1}{(2\pi)^{TM/2} |\mathbf{\Sigma}|^{T/2}} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} (\underset{\sim}{y} - \underset{\sim}{f})' \mathbf{\Sigma}^{-1} (\underset{\sim}{y} - \underset{\sim}{f}) \right) \tag{30}$$

$(\underset{\sim}{y} - \underset{\sim}{f})' \mathbf{\Sigma}^{-1} (\underset{\sim}{y} - \underset{\sim}{f})$ in $L(\underset{\sim}{f}, \mathbf{\Sigma} | \underset{\sim}{y})$ can be elaborated as follow:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$(\underset{\sim}{y}-\underset{\sim}{f})'\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}) = \mathrm{tr}\left[(\underset{\sim}{y}-\underset{\sim}{f})'\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})\right]$$
$$= \mathrm{tr}\left[\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right]$$

$$\sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})'\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}) = \sum_{t=1}^{T}\mathrm{tr}\left[\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right]$$
$$= \mathrm{tr}\left[\mathbf{\Sigma}^{-1}\sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right] \tag{31}$$

Thus, equation (30) is substituted using equation (31) and the result is as follow:

$$L(\underset{\sim}{f},\mathbf{\Sigma}\,|\,\underset{\sim}{y}) = \frac{1}{(2\pi)^{TM/2}\left|\mathbf{\Sigma}\right|^{T/2}}\exp\left(-\frac{1}{2}\mathrm{tr}\left[\mathbf{\Sigma}^{-1}\sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right]\right) \tag{32}$$

Estimator for variance-covariance matrix $\hat{\mathbf{\Sigma}}$ is obtained by maximizing function of $L(\underset{\sim}{f},\mathbf{\Sigma}\,|\,\underset{\sim}{y})$, through $\frac{\partial L(\underset{\sim}{f},\mathbf{\Sigma}\,|\,\underset{\sim}{y})}{\partial\mathbf{\Sigma}} = \mathbf{0}$. As mentioned in Fernandes [17], likelihood function in equation (16) will meet the maximum condition if $\hat{\mathbf{\Sigma}} = \frac{1}{2b}\mathbf{B}$, with $b = 2T$, and $\mathbf{B} = \sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'$, or can be reformulated as follow:

$$\hat{\mathbf{\Sigma}} = \frac{1}{2b}\mathbf{B}$$
$$\hat{\mathbf{\Sigma}} = \frac{\sum_{t=1}^{T}(\underset{\sim}{y}-\hat{\underset{\sim}{f}})(\underset{\sim}{y}-\hat{\underset{\sim}{f}})'}{T} \tag{33}$$

Random error variance-covariance matrix in the study is similar to equation (33) or can be reformulated as:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Sigma}}_{NN} \end{bmatrix}$$

Hence, equation (32) can be reformulated as:

$$L(\underset{\sim}{f}, \mathbf{\Sigma} \big| \underset{\sim}{y}) = \frac{1}{(2\pi)^{TM/2} \left|\mathbf{\Sigma}\right|^{T/2}} \exp\{-\frac{1}{2}\sum_{t=1}^{T}\{(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{11}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t}) + (\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{22}(\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t}) +$$
$$(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{33}(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t}) + (\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{12}(\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t}) + (\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{12}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t}) +$$
$$(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{13}(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t}) + (\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{13}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t}) + (\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{23}(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t}) +$$
$$(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{23}(\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})\}$$

$$L(\underset{\sim}{f}, \mathbf{\Sigma} \big| \underset{\sim}{y}) = \frac{1}{\prod_{t=1}^{T}\left\{(2\pi)^{M/2}\left|\mathbf{\Sigma}\right|^{1/2}\right\}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\left\{(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{11}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t}) + \right.\right.$$

$$+ (\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{22}(\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t}) + (\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{33}(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t}) +$$

$$+ (\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{12}(\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t}) + (\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{21}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t}) +$$

$$+ (\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{13}(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t}) + (\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{31}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t}) +$$

$$+ (\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{23}(\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t}) + (\underset{\sim}{y}_{3t} - \underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{23}(\underset{\sim}{y}_{2t} - \underset{\sim}{f}_{2t})\}$$

$$= \left\{\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{11.1}\right|^{T/2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{11}(\underset{\sim}{y}_{1t} - \underset{\sim}{f}_{1t})\right]\right)\times\right.$$

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{22.2}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{2t}-\underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{22}(\underset{\sim}{y}_{2t}-\underset{\sim}{f}_{2t})\right]\right)\times$$
$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{33.2}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{3t}-\underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{33}(\underset{\sim}{y}_{3t}-\underset{\sim}{f}_{3t})\right]\right)\times$$
$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{12.N}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{12}(\underset{\sim}{y}_{2t}-\underset{\sim}{f}_{2t})\right]\right)\times$$

$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{12.N}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{2t}-\underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{12}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})\right]\right)\times$$

$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{13.N}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{13}(\underset{\sim}{y}_{3t}-\underset{\sim}{f}_{3t})\right]\right)\times$$
$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{13.N}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{3t}-\underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{13}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})\right]\right)\times$$

$$\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{23.N}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{2t}-\underset{\sim}{f}_{2t})'\mathbf{\Sigma}_{23}(\underset{\sim}{y}_{3t}-\underset{\sim}{f}_{3t})\right]\right)\times$$
$$\left.\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{23.N}\right|^{T/2}}\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{3t}-\underset{\sim}{f}_{3t})'\mathbf{\Sigma}_{23}(\underset{\sim}{y}_{2t}-\underset{\sim}{f}_{2t})\right]\right)\right\}$$

Estimator for variance-covariance matrix $\hat{\mathbf{\Sigma}}$ is obtained by maximizing function of $L(\underset{\sim}{f},\mathbf{\Sigma}\,|\,\underset{\sim}{y})$, through $\dfrac{\partial L(\underset{\sim}{f},\mathbf{\Sigma}_{ij}\,|\,\underset{\sim}{y})}{\partial\mathbf{\Sigma}_{ij}}=\mathbf{0}$. Elaboration of each sub-matrix of $\hat{\mathbf{\Sigma}}_{ij}$ is as follow:

For $\hat{\mathbf{\Sigma}}_{11}$, it is obtained that:

$$\frac{\partial L(\underset{\sim}{f},\mathbf{\Sigma}_{11}\,|\,\underset{\sim}{y})}{\partial\mathbf{\Sigma}_{11}}=\frac{\partial\left(\dfrac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{11}\right|^{T/2}}\exp\{-\dfrac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})'\mathbf{\Sigma}_{11}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})\}\right)}{\partial\mathbf{\Sigma}_{11}}$$

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$= \frac{\partial\left(\frac{1}{(2\pi)^{TT/2}\left|\mathbf{\Sigma}_{11}\right|^{T/2}}\exp\{-\frac{1}{2}\mathrm{tr}\left[\mathbf{\Sigma}_{11}^{-1}\left(\sum_{t=1}^{T}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})'\}\right)\right]\right)}{\partial\mathbf{\Sigma}_{11}}$$

$$= -\frac{T}{2}\frac{\partial\ln\left|\mathbf{\Sigma}_{11}\right|}{\partial\mathbf{\Sigma}_{11}} - \frac{1}{2}\frac{\partial\left(\mathrm{tr}\left[\mathbf{\Sigma}_{11}^{-1}\left(\sum_{t=1}^{T}(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})(\underset{\sim}{y}_{1t}-\underset{\sim}{f}_{1t})'\}\right)\right]\right)}{\partial\mathbf{\Sigma}_{11}}$$

$$\frac{\partial L(\underset{\sim}{f},\mathbf{\Sigma}\left|\underset{\sim}{y}\right.)}{\partial\mathbf{\Sigma}_{11}} = \mathbf{0},$$

Based on the elaboration of equation (33), estimation of $\hat{\mathbf{\Sigma}}_{11}$ is as follow:

$$\hat{\mathbf{\Sigma}}_{11} = \frac{(\underset{\sim}{y}_1-\hat{\underset{\sim}{f}}_1)(\underset{\sim}{y}_1-\hat{\underset{\sim}{f}}_1)'}{T}.$$

Using the same method, $\hat{\mathbf{\Sigma}}_{22},\hat{\mathbf{\Sigma}}_{33},\hat{\mathbf{\Sigma}}_{12},\hat{\mathbf{\Sigma}}_{13},\hat{\mathbf{\Sigma}}_{23}$ is:

$$\hat{\mathbf{\Sigma}}_{22} = \frac{(\underset{\sim}{y}_2-\hat{\underset{\sim}{f}}_2)(\underset{\sim}{y}_2-\hat{\underset{\sim}{f}}_2)'}{T}.$$

$$\hat{\mathbf{\Sigma}}_{33} = \frac{(\underset{\sim}{y}_3-\hat{\underset{\sim}{f}}_3)(\underset{\sim}{y}_3-\hat{\underset{\sim}{f}}_3)'}{T}.$$

$$\hat{\mathbf{\Sigma}}_{12} = \frac{(\underset{\sim}{y}_1-\hat{\underset{\sim}{f}}_1)(\underset{\sim}{y}_2-\hat{\underset{\sim}{f}}_2)'}{T}.$$

$$\hat{\mathbf{\Sigma}}_{13} = \frac{(\underset{\sim}{y}_1-\hat{\underset{\sim}{f}}_1)(\underset{\sim}{y}_3-\hat{\underset{\sim}{f}}_3)'}{T}.$$

$$\hat{\mathbf{\Sigma}}_{23} = \frac{(\underset{\sim}{y}_2-\hat{\underset{\sim}{f}}_2)(\underset{\sim}{y}_3-\hat{\underset{\sim}{f}}_3)'}{T}.$$

Or it can be formulated that:

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Sigma}}_{NN} \end{bmatrix} \tag{34}$$

with

$$\hat{\mathbf{\Sigma}}_{ij} = \frac{(\underset{\sim}{y}_i - \hat{\underset{\sim}{f}}_i)(\underset{\sim}{y}_j - \hat{\underset{\sim}{f}}_j)'}{T}, \qquad \blacksquare$$

Estimation of variance-covariance matrix in equation 34) can be used to predict regression curve equation (1).

3.2. Algorithm For Generating Data Simulation

The second goal of the study is to develop algorithm that generates simulation data with certain autocorrelation level based on size of sample (*T*) and error variance (*EV*). Simulations were obtained by the following steps.

Algorithm 1:

Step 1

Calculate the value of $x_i, i = 1, 2, ..., N$ as fix variable with $x_i \in [0,1]$ by design subject points define by $x_i = \frac{2i-1}{2N}$

Step 2

Non-linear function is obtained using the following function: Exponential function: $f(x_i) = 4.26(e^{-3.25x_i} - 4e^{-6.5x_i} + 3e^{-9.75x_i})$, with x_i from step 1.

Step 3

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

Generate the error vector (ε_i) by the following steps.

Step 3.1

Generating autocorrelation in 10 condition, namely: 0.01 – 0.09; 0.11 – 0.19; 0.21 – 0.29; 0.31 – 0.39; 0.41 – 0.49; 0.51 – 0.59; 0.61 – 0.69; 0.71 – 0.79; 0.81 – 0.89; 0.91 – 0.99.

Step 3.2

The ten conditions of autocorrelation is generated randomly with uniform distribution

Step 3.3

Establishing autocorrelation by deciding EV of (0.1,1,5)

Step 3.4

Establishing Variance Matrix ($\hat{\mathbf{\Sigma}}$) accomodating autocorrelation

Step 4

Obtain *y* from the function: $y_i = f\left(x_i\right) + \varepsilon_i$, from step 2 and 3.

Step 5

Obtain estimation of function ($\hat{f}_i$) from smoothing spline nonparametric regression.

Step 5.1

Establishing matrix **T** with m is spline order of which value is 2

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{T}_N \end{bmatrix}$$

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

$$\mathbf{T}_i = \begin{pmatrix} \langle \eta_{i1}, \phi_{i1} \rangle & \langle \eta_{i1}, \phi_{i2} \rangle & \cdots & \langle \eta_{i1}, \phi_{im} \rangle \\ \langle \eta_{i2}, \phi_{i1} \rangle & \langle \eta_{i2}, \phi_{i2} \rangle & \cdots & \langle \eta_{i2}, \phi_{im} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{iT}, \phi_{i1} \rangle & \langle \eta_{iT}, \phi_{i2} \rangle & \cdots & \langle \eta_{iT}, \phi_{im} \rangle \end{pmatrix},$$

$$\langle \eta_{it}, \phi_{ij} \rangle = \frac{x_i^{j-1}}{(j-1)!}, with\, t = 1, 2, ..., T;\ j = 1, 2, ..., m$$

$$\underset{\sim}{d} = (\underset{\sim}{d}_1', \underset{\sim}{d}_2', \cdots, \underset{\sim}{d}_N')',\ \underset{\sim}{d}_i' = (d_{i1}, d_{i2}, ..., d_{im}),$$

Step 5.2

Establishing matrix **V**

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_N \end{bmatrix},$$

$$\mathbf{V}_i = \begin{pmatrix} \langle \xi_{i1}, \xi_{i1} \rangle & 0 & \cdots & 0 \\ \langle \xi_{i2}, \xi_{i1} \rangle & \langle \xi_{i2}, \xi_{i2} \rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \langle \xi_{iT}, \xi_{i1} \rangle & \langle \xi_{iT}, \xi_{i2} \rangle & \cdots & \langle \xi_{iT}, \xi_{iT} \rangle \end{pmatrix},$$

$$\langle \xi_{it}, \xi_{is} \rangle = \int_a^b \frac{(x_{it} - u)_+^{m-1} (x_{is} - u)_+^{m-1}}{\left((m-1)!\right)^2} du, t = 1, 2, ..., T; s = 1, 2, ..., T$$

$$\underset{\sim}{c} = (\underset{\sim}{c}_1', \underset{\sim}{c}_2', \cdots, \underset{\sim}{c}_N')',\ \underset{\sim}{c}_i' = (c_{i1}, c_{i2}, ..., c_{iT}).$$

Step 5.3

Using λ equals zero to get $\hat{c}$ and $\hat{d}$ as initial

Step 5.4

 ACCEPTED MANUSCRIPT

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

Evaluate the λ by minimize the GCV function by using iteration $0 < \lambda < \infty$, the iteration stopped when $\left|GCV_t - GCV_{t-1}\right| < 10^{-6}$ or maximum iteration of 1000.

3.3. Application of Smoothing Spline the autocorrelation Level Group

The third objective of this study is to apply the form of spline estimator in nonparametric regression on longitudinal data to simulate various levels of autocorrelation between data of observation within subject, as well as compare the approach of DM and TM in predicting estimator spline to data simulation of various levels of autocorrelation between data of observation on the subject that same.

A simulation study was conducted with statistical software R. spline nonparametric regression estimator using 100 iterations and 10 repetitions and apply a smoothing spline models with two approaches: (1) Approach PWLS by considering the autocorrelation (DM) and (2) without considering the approach PWLS autocorrelation (TM). Both approaches are applied to the variations in the size of the observation, which is a measure minor observations (T = 20), the size of the observation medium (T = 50), the size of the observation of large (T = 100) and the level of error variance, the error variance is small (EV = 0.1), error variance being (EV = 1), a large error variance (EV = 5). Autocorrelation level are grouped into ten categories: starting from a very small autocorrelation (| 0.01 to 0.09 |), up to the level of autocorrelation is very large (| 0.91 to 0.99 |).

Visually, the scatter diagram of the various levels of autocorrelation (0.2; 0.4; 0.6; 0.8) indicates how closely the model is able to explain the relationship between predictors and response. In the scatter diagram in Figure 1 shows the autocorrelation level of 0.2 that the actual curve and prediction ($\hat{f}_{TM}$ and $\hat{f}_{DM}$) formed a pattern that follows the pattern of distribution data.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

 ACCEPTED MANUSCRIPT

In Figure 1 shows the graph $\hat{f}_{TM}$ and $\hat{f}_{DM}$ the various levels of autocorrelation, where the autocorrelation level is low (0.2) shows that the $\hat{f}_{TM}$ curve is quite identical to the prediction $\hat{f}_{DM}$ curve. It is seen from the difference in the coefficient of determination (DR^2) is quite small at 5.69%. Unlike the autocorrelation level of 0.4, smoothing spline curve by considering the autocorrelation ($\hat{f}_{DM}$) better (DR^2 = 9:53%) compared smoothing spline curves without considering autocorrelation ($\hat{f}_{TM}$). The higher the level of autocorrelation, the better the smoothing spline curve by considering the correlation ($\hat{f}_{DM}$) compared without considering the correlation ($\hat{f}_{TM}$).

3.3.1. Application of General Smoothing Spline

Estimator spline smoothing spline regression in general (all sizes observations and error variance) is obtained curve R2 difference DM and TM in Figure 2 below. Figure 2 shows there is a trend (increase) / R2 is not constant from the difference between DM and TM, the greater the value, the greater the difference autocorrelation R2, so the spline estimator indicated their differences with PWLS approach by considering the autocorrelation and without considering the autocorrelation. Thus, to find a point of difference (cut off) performed statistical tests to determine the differences estimator F spline at a certain level of autocorrelation.

Based on Table 1 it is known that the F test spline estimator at various levels of autocorrelation have the p-value of 0.000 is smaller than the value α (0.05) so that there is a significant difference spline estimator at various levels of autocorrelation. To find more differences at each level different autocorrelation continued Tukey test. Tukey test results

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

showed that there were significant differences in the level of autocorrelation 0.81 to 0.90 so that the distribution of two groups: group 1 (autocorrelation <0.8) and group 2 (autocorrelation ≥ 0.8).

T test results in Table 3 shows the p-value of 0.000 is smaller than the value α (0.05) so that the significant difference between the groups. It shows the spline estimator in group 1 was significantly different with group 2. Thus, the autocorrelation> 0.81 is a point of difference (as cut off value) the results of spline estimator with PWLS approach considering the autocorrelation and without considering the autocorrelation.

3.3.2. Application of Smoothing Spline by Size Observations

R^2 difference DM and TM spline estimator in nonparametric smoothing spline regression in size variation explained by the observation of Figure 2. Estimator spline on the size of the observation showed no significant differences in the various levels of autocorrelation indicated by F test are presented in Table 2 below:

Based on Table 2 that the F test spline estimator at various levels of autocorrelation have the p-value of 0.000 is smaller than the value α (0.05) so that the spline estimator at various levels of autocorrelation differed significantly in all sizes observations. To find more differences at each level different autocorrelation continued Tukey test. Tukey test results showed that there were significant differences in the level of autocorrelation 0.81 to 0.90 so that the distribution of two groups: group 1 (level autocorrelation <0.81) and group 2 (autocorrelation levels ≥ 0.81).

T-test results in Table 2 shows that the p-value of 0.000 is smaller than the value α (0.05) so that the spline estimator first group differ significantly from group 2. Thus, the autocorrelation level>

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

0.8 is a point of difference (cut off) spline estimator results on all measures considering the observations with PWLS approach without considering the autocorrelation and autocorrelation. With the same test conditions as the test differences in all sizes observations, produced the following test.

Based on the above test results it can be seen that there are significant differences estimator spline through PWLS approach by considering the autocorrelation and without considering the autocorrelation at small observation size (T = 20) when the level of autocorrelation> 0.7, the size of the observation medium (T = 50) at autocorrelation level> 0.8, and the size of a large observation (T = 100) when the level of autocorrelation> 0.8.

3.3.3. Application of Smoothing Spline based Error Variance

R2 difference DM and TM spline estimator in nonparametric regression smoothing spline on the variation of error variance explained by Figure 4 below:

Based on Table 4 is known that the F test spline estimator at various levels of autocorrelation have the p-value of 0.000 is smaller than the value α (0.05) so that the spline estimator at various levels of autocorrelation significantly different at all error variance. To find more differences at each level different autocorrelation continued Tukey test. Tukey test results showed that there were significant differences in the level of autocorrelation 0.81 to 0.90 so that the distribution of two groups: group 1 (level autocorrelation <0.8) and group 2 (autocorrelation levels ≥ 0.8).

The results of the t-test p-value of 0.000 is smaller than the value α (0.05) so that the spline estimator first group differ significantly from group 2. Thus, the autocorrelation> 0.8 is a point of difference (cut off) the results of the spline estimator all error variance with the approach PWLS consideration or without consideration autocorrelation autocorrelation. With the same test as the

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

test differences on all the spline estimator error variance, the results of testing on every variation of error variances described below table.

Based on the above test results it can be seen that there are significant differences estimator spline through PWLS approach by considering the autocorrelation and without considering small autocorrelation in error variance (EV = 0.1) when the autocorrelation> 0.7, error variance being (EV = 1) when autocorrelation> 0.8, and a large error variance (EV = 5) when the autocorrelation> 0.8.

4. Conclusions and Recommendations

Some of the conclusions derived from the results of this simulation study are:

(1) Estimated spline function $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c}$, with an estimated spline functions taking into account the autocorrelation (DM) is $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ to

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}]$$

And to estimate the spline function without considering the autocorrelation (TM) is $\hat{\mathbf{\Sigma}} = \mathbf{I}$ equivalent to $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ the

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}].$$

(2) Algorithm to generate simulated data with a certain degree of autocorrelation based on variations in sample size (T) and the level of error variance (EV), presented in five steps by setting the sample size and the level of error variance (EV) on ten levels of autocorrelation. Simulation data obtained from Algorithm 1 above is implemented to estimate the smoothing spline function as presented in Theorem 1 for the case autocorrelation.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

(3) The results of the application as follows: (a) Implementation of smoothing spline with PWLS approach with or without consideration of autocorrelation in general (in all sizes and all observation error variance) provides significantly different spline estimator when the autocorrelation level> 0.8. (b) spline estimator in nonparametric regression smoothing spline with PWLS approach with or without consideration of autocorrelation in all sizes observations showed significantly different results when the autocorrelation level> 0.8, whereas the size of a small observation when the level of autocorrelation> 0.7, the size of the observation was and the size of a large observation when the level of autocorrelation> 0.8. (c) spline estimator in nonparametric regression smoothing spline with PWLS approach with or without consideration of autocorrelation in all the error variance give significantly different results when the autocorrelation level> 0.8, whereas the small variance error when autocorrelation level> 0.7, error variance was and the error variance is greater when the level of autocorrelation of 0.8.
In this study, the autocorrelation is restricted between so in future studies can be studied more in depth about the value of autocorrelation involving autocorrelation value -1 to 1. In addition to it, can be studied in future research on the optimal level of autocorrelation Truncated Spline-based nonparametric regression because in this study only limited to the level of autocorrelation in nonparametric regression-based Smoothing Spline.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

References

[1] Eubank, R.L., 1999, *Nonparametric Regression and Spline Smoothing*, Second Edition, Marcel Dekker, Inc, New York.

[2] Wu, H., & Zhang, J.T, 2006. *Nonparametric Regression Methods for Longitudinal Data Analysis.* New Jersey: John Wiley and Sons, Inc.

[3] Wahba, G.(1990). Spline *Models for Observational Data*. Pensylvania. SIAM.

[4] Wang, Y. (2011), *Smoothing Spline Methods Applications*, CRC Press, New York.

[5] Weiss, R.E. (2005). *Modelling* Longitudinal *Data*. Springer Texts in Statistic New York. Retrieved, January, 31, 2012. Website: http://www.biostat.ucla. edu/books/mld.

[6] Wang, J.L. (2003). *Nonparametric Regression Analysis of* Longitudinal *Data*. California: University of California Press.

[7] Howell, J.R. (2007). *Analysis Using Smoothing Splines As Implemented In LME() In R*. Thesis. Brigham Young University.

[8] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., and Suhartono. (2015). Spline Estimator for Bi-Responses and Multi-Predictors Nonparametric Regression Model in Case of Longitudinal Data, *Journal of Mathematics and Statistics*, Vol 11, No 2, 2015, pp. 61--69.

[9] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., *and* Suhartono. 2014. Spline estimator for bi-responses nonparametric regression model for longitudinal data. *Applied Mathematical Sciences*, Vol. 8, no. 114, 2014, 5653--5665.

[10] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., and Suhartono. (2014). Reproducing Kernel Hilbert Space and Penalized Weighted Least Square in Nonparametric Regression, *Applied Mathematical Science* Vol 8, 2014, No 146, pp. 7289--7300.

[11] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., and Suhartono. (2014). Reproducing Kernel Hilbert Space for Penalized Regression multi-predictors: case in longitudinal data. *International Journal of Mathemathical Analysis*, Vol. 8, no. 40, 2014, 1951--1961.

[12] Gujarati, D. (2011), *Econometrics By Example*, Palgrave Macmillan, Britania.

[13] Budiantara, I.N. (2000). Method of U, GML, CV and GCV in Nonparametric Regression Spline (In Indonesian Metode U, GML, CV dan GCV Dalam Regresi Nonparametrik Spline). *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6, p.285-290.

[14] Kutner, M.H., Nachtsheim, C.J., Neter, J.& Li W. (2005), *Applied Linear Statistical Models*, Fifth Edition, McGraw-Hill International, Boston.

[15] Kim, Y.J., and Gu, C. (2004). Smoothing Spline Gaussian Regression: More Scalable Computation via Efficient Approximation. *Royal Statistical Society: Series B*, **66**(2), 337--356.

[16] Lee, T.C.M. (2004). Improved Smoothing Spline Regression by Combining Estimates of Different Smoothness. *Statistics and Probability Letters*, **67**(1), 133--140.

[17] Fernandes, A.A.R, Amaliana, L, Samingun, and Solimun. (2017). The Comparison of Spline Estimators in the Smoothing Spline Nonparametric Regression Model Based on Weighted Least Square (WLS) and Penalized Weighted Least Square (PWLS) in Longitudinal Data (A Study on the Baby Growth in Indonesia). *International Journal of Control Theory and Applications*, 10(4), 257--276.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

ACCEPTED MANUSCRIPT

[18] Solimun, Sa'adah, U., Astutik, S., and Fernandes, A.A.R. (2017). The Efficiency of Curve Regression in Semiparametric Spline Regression Models with and without Penalty. *International Journal of Control Theory and Applications*, 10(4), 287--302.

[19] Wardhani, N.W.S, and Fernandes, A.A.R. (2017). The Efficiency of Curve Regression Using Linear and Quadratic Truncated Spline for Estimated Growth Curve of Green Beans. *International Journal of Control Theory and Applications*, 10(4), 339--348.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

 ACCEPTED MANUSCRIPT

Table 1. Tukey test for General Estimator Spline.

Autocorrelation Level	Tukey test				
	a	b	c	d	e
0.00 – 0.10					
0.11 – 0.20					
0.21 – 0.30					
0.31 – 0.40					
0.41 – 0.50					
0.51 – 0.60					
0.61 – 0.70					
0.71 – 0.80					
0.81 – 0.90					
0.91 – 1.00					
F test			p-value		0.000
T test			p-value		0.000

Table 2. Tukey Test for Spline Estimator in All Sizes Level.

Autocorrelation Level	Tukey test				
	a	b	c	d	e
0.00 – 0.10					
0.11 – 0.20					
0.21 – 0.30					
0.31 – 0.40					
0.41 – 0.50					
0.51 – 0.60					
0.61 – 0.70					
0.71 – 0.80					
0.81 – 0.90					
0.91 – 1.00					
F test			p-value		0.000
T test			p-value		0.000

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

Table 3. Tukey Test for Spline Estimator in Size Variation Level.

Autocorrelation Level	Tukey test										
	T = 20					T = 50			T = 100		
	a	b	c	d	e	a	b	c	a	b	c
0.00 – 0.10											
0.11 – 0.20											
0.21 – 0.30											
0.31 – 0.40											
0.41 – 0.50											
0.51 – 0.60											
0.61 – 0.70											
0.71 – 0.80											
0.81 – 0.90											
0.91 – 1.00											
F test	p-value		0.000								
T test	p-value		0.000								

Table 4. Tukey Test for Spline Estimator in All Error Variance Level.

Autocorrelation Level	Tukey test				
	a	b	c	d	e
0,00 – 0,10					
0,11 – 0,20					
0,21 – 0,30					
0,31 – 0,40					
0,41 – 0,50					
0,51 – 0,60					
0,61 – 0,70					
0,71 – 0,80					
0,81 – 0,90					
0,91 – 1,00					
F test			p-value		0,000
T test			p-value		0,000

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

Table 5. Tukey Test for Spline Estimator in Error Variance Variation Level.

Autocorrelation Level	Tukey test											
	EV = 0.1					EV = 1			EV = 5			
	1	2	3	4	5	1	2	3	1	2	3	4
0.00 – 0.10												
0.11 – 0.20												
0.21 – 0.30												
0.31 – 0.40												
0.41 – 0.50												
0.51 – 0.60												
0.61 – 0.70												
0.71 – 0.80												
0.81 – 0.90												
0.91 – 1.00												
F test						p-value				0.000		
T test						p-value				0.000		

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

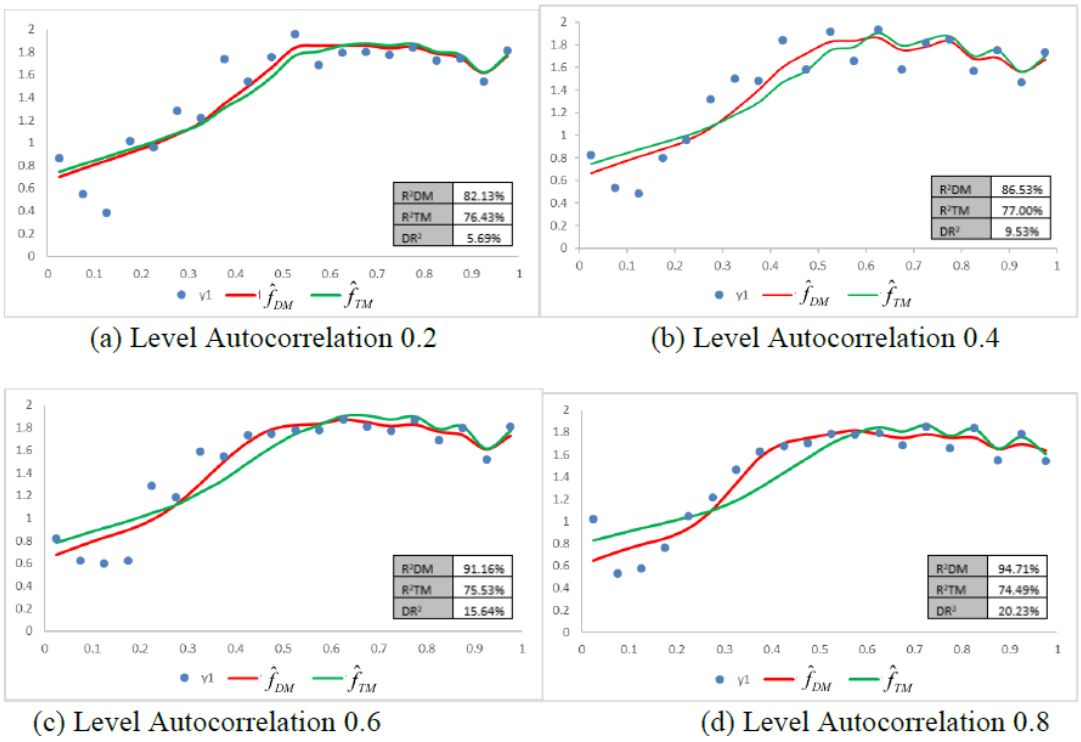


Figure 1. Plot $\hat{f}_{TM}$ and $\hat{f}_{DM}$ On Actual Data with Various levels of autocorrelation.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017

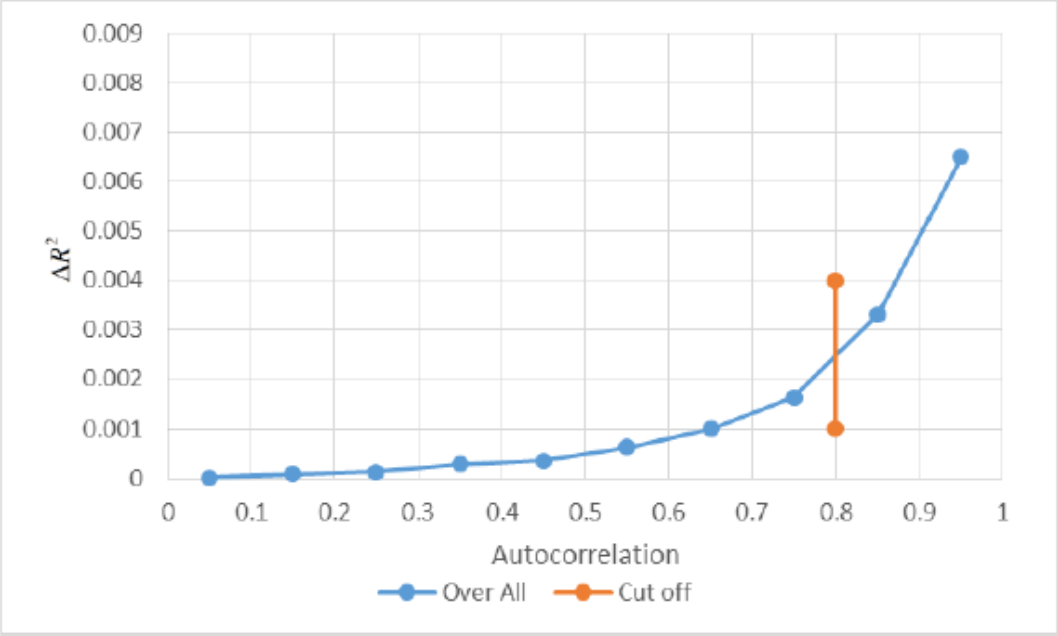
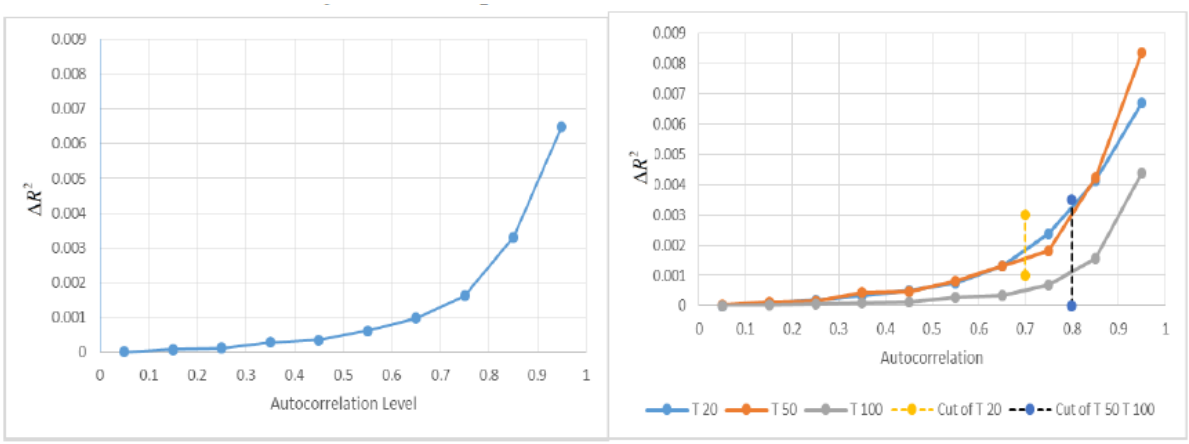


Figure 2. Difference R^2 DM and TM in general in all conditions.

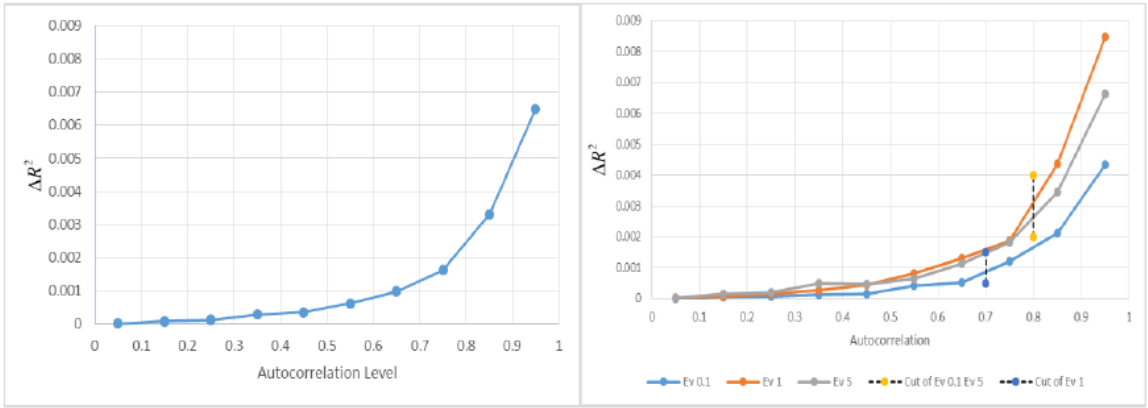
Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017



(a) All Sizes N (b) Variation Sizes N 20 50 100

Figure 3. Difference R^2 DM and TM Based Observation Size.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017



(a) All Ev (b) Varition Ev 0.1, 1, and 5

Figure 4. Difference R^2 DM and TM Based Error Variance (Ev) Spline estimator in the variation of error variance showed a significant difference at every level of autocorrelation is shown with a different test in Table 4.

Downloaded by [University of Virginia, Charlottesville] at 10:55 09 October 2017