International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 4 • 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)

Adji Achmad Rinaldo Fernandes*, Luthfatul Amaliana, Samingun, Solimun

Statistics Department, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Indonesia
**Corresponding Author: fernandes@ub.ac.id*

Abstract: The objectives of this research are; (1) to obtain the estimation of *smoothing spline* nonparametric regression based on PWLS and WLS, (2) to obtain the error variance-covariance matrix based weighted estimation, (3) to examine the efficiency of *Spline Estimator* Curve in Smoothing Spline Nonparametric Regression Model Based on Weighted Least Square (WLS) and Penalized Weighted Least Square (PWLS) on the Longitudinal Data in the data of baby growth. One of the ways to measure the baby growth is by recording the age and weight of the baby monthly and having it written in a card called as *Kartu Menuju Sehat* (KMS) to detect the malnutrition in toddlers. The estimation of *smoothing spline* nonparametric regression PWLS (with penalty) as follow:

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

Without penalty using WLS as follow: $\hat{\underset{\sim}{f}} = \mathbf{A}\underset{\sim}{y},\ \mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}$.

The estimation of error variance-covariance matrix is as follow:

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Sigma}}_{NN} \end{bmatrix} \text{ with } \hat{\mathbf{\Sigma}}_{ij} = \frac{(\underset{\sim}{y}_i - \hat{\underset{\sim}{f}}_i)(\underset{\sim}{y}_j - \hat{\underset{\sim}{f}}_j)'}{T},$$

The estimation of the curve of PWLS based nonparametric regression in the data of baby growth is more efficient than the estimation of the curve of WLS based nonparametric regression. It can be seen from the efficiency of WLS based curve that is only 48.4% or less than 50%, compared to the efficiency of the PWLS based curve. In other word, the use of estimation of the curve of penalty (PWLS) based smoothing spline nonparametric regression has a better efficiency level than the WLS (without penalty) based.

***Keywords**:* Smoothing Spline, Longitudinal Data, WLS, PWLS, Growth Curve

1. INTRODUCTION

Regression Analysis is one of the statistical methods used to find the pattern of the relationship between response *y* on predictor *x*. If the pattern is stated in a graphical form, the linear relationship between *x* and *y* will form be in the form of straight line, quadratic, cubic, exponent, etc. There are two approaches that can be used to estimate the form of curve of regression *f* in describing the pattern of the relationship between the response and predictor, they are parametric and nonparametric regressions. Parametric regression is assumed as the curve of regression *f* is identified for its form. There is an assumption that should be met in parametric regression, that is the assumption of the form of linear relationship between the response and the predictor. If the linearity assumption is not met and the nonlinear form is not or has not been identified, then the nonparametric regression is used as an alternative with an assumption that the curve of regression *f* has not been identified for its form.

There are three types of data in order to conduct regression analysis, they are; cross-section, time-series, and longitudinal. The combination between cross-section and time-series data is called as longitudinal data. Longitudinal data is the data resulted from the observation on *N* subjects that are independent each other and each subject is observed for several times in T period. A research conducted by Verbekke and Molenberghs [1] studied the parametric regression for longitudinal data with Generalized Linear Mixed Model approach. The approach accommodated the correlation of the observations on similar subjects by adding the effect of each subject, the random effect, from the fixed effect. The combination between random effect and fixed effect is called as mixed effect.

One of the approaches in nonparametric regression is smoothing spline that has a specific characteristic by which is able to adjust the changes of data behavior very well. In the smoothing spline, it does not need the knot selection since the estimation of the function is based on the criteria of model accuracy and the smoothness of the curve that have been set by the smoothing parameter. A previous research conducted by Fernandes [2-5] studied the development of bi-response spline estimator specifically for longitudinal data by using smoothing spline with the use of Reproducing Kernel Hilbert Space (RKHS) approach. Another research conducted by Budiantara [6] studied on how to obtain the estimator form of the curve of regression *f* for bi-response longitudinal data using Generalized Penalized Spline approach.

In order to obtain the estimation of nonparametric regression function for longitudinal data, it can be conducted by using the Penalized Weighted Least Square (PWLS) optimization, in which, by adding variance covariance matrix Σ as weighted on the completion of the least quadratic optimization. Lestari [7], Fernandes [2-5] had conducted a research by studying the procedures in obtaining the multi-responses estimator form of curve *f* of nonparametric regression by using PWLS. The addition of weight is conducted since in the longitudinal data, the similar subjects are dependent, while the different subjects are independent so that the correlation between the observations needs a weight on the least quadratic optimization.

In this research, the nonparametric regression model was used for longitudinal data by using *smoothing spline* approach and adding the Σ weight in estimating the regression function implemented in the data of baby growth. The estimation on the parameter, in which considering the penalty controlling the roughness/smoothness of the curve, is called as PWLS method. In contrast, if it does not use the penalty, it is called as WLS method. Therefore, this research is aimed at testing the efficiency of the use of PWLS and WLS, which one is better.

Based on the background above, the objectives of this research are; (1) to obtain the estimation of smoothing spline nonparametric regression based on PWLS and WLS, (2) to obtain the error variance-covariance matrix based weighted estimation, (3) to examine the efficiency of Spline Estimator Curve in Smoothing Spline Nonparametric Regression Model Based on Weighted Least Square WLS and Penalized Weighted Least Square PWLS on the Longitudinal Data in the data of baby growth. One of the ways to measure the baby growth is by recording the age and weight of the baby monthly and having it written in a card called as Kartu Menuju Sehat (KMS) to detect the malnutrition in toddlers.

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 among subjects assumed to be independent of each other, but between observations in the same subjects is dependent [9-11]. Many researchers have developed a spline estimator in nonparametric regression model for longitudinal data ([2,3,6,12]). 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:

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

Information:

y_{it} : Response on the subject all the time observation i and all t,

x_{it} : Predictors on the subject all the time observation i and all t,

f_i : Regression curve predictor relationship with the response on the subject i,

N : number of subjects,

T : The number of observations of each subject,

P : the number of predictors,

ε_{it} : Random error on the subject all the time observation i and all t,

The regression model in equation (1) to include as a regression curve f_i that accommodates no liability observations on the same 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 ($\underset{\sim}{\varepsilon}$) = (vector measuring NT) and the variance-covariance matrix Var ($\underset{\sim}{\varepsilon}$) = Σ (matrix measuring NT × NT) as follows ([2-7]):

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

The matrix Σ can be simplified into sub-matrices Σ_i and 0.

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

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

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

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 in the same subjects. This covariance can be worth not 0, which accommodates the correlation between observations in the same subjects. On the other hand, the sub-matrix that 0 is the matrix of all elements of value 0 states that the covariance between observations in different subjects are mutually 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 (smooth), in the sense of space is contained in a particular function, especially Sobolev spaces or written ([13-15])

$f_i \in W_2^m[a_i, b_i]$ where:

$$W_2^m[a_i, b_i] = \left\{ f_i : f_i, f_i^{(1)}, ..., f_i^{(m-1)} \text{ kontinu absolut;} \int_{a_i}^{b_i} [f_i^{(m)}(x_{it})]^2 dx_{it} < \infty \right\}, \quad (3)$$

or a constant 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 Σ as has 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 [10,11]:

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

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. RESULT AND DISCUSSION

There are three research objectives that will be completed in this paper. First, to obtain the estimation of smoothing spline nonparametric regression with penalty, PWLS (written in Theorem 1), and without penalty, WLS (written in Theorem 2). Second, to obtain the error variance-covariance matrix based weighted estimation. Third, to examine the efficiency of the Spline Estimator curve in the PWLS and WLS based Smoothing Spline Nonparametric Regression Model in the baby growth.

Theorem 1: Estimation of Curve Regression with Penalty (PWLS)

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

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

then the spline estimator that minimizes PWLS

$$\min_{f_i \in W_2^m[a_i,b_i], i=1,2,\ldots,N} \left\{ M^{-1}\left(\underset{\sim}{y} - \underset{\sim}{f}\right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y} - \underset{\sim}{f}\right) + \sum_{i=1}^{N} \lambda_i \int_{a_i}^{b_i} \left[f_i^{(m)}(x_{it}) \right]^2 dx_{it} \right\} \tag{5}$$

is

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

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

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

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 [2-7]:

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

where **T** is $(NT)´(Nm)$ 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
$$\mathbf{T}_i = \begin{pmatrix} \langle \eta_{i1}, \phi_{i1} \rangle & \langle \eta_{i1}, \phi_{i2} \rangle & \cdots & \langle \eta_{i1}, \phi_{im} \rangle \\ \langle \eta_{i2}, \phi_{i1} \rangle & \langle \eta_{i2}, \phi_{i2} \rangle & \cdots & \langle \eta_{i2}, \phi_{im} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_{iT}, \phi_{i1} \rangle & \langle \eta_{iT}, \phi_{i2} \rangle & \cdots & \langle \eta_{iT}, \phi_{im} \rangle \end{pmatrix}, \tag{9}$$

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

$\underset{\sim}{d}$ is Nm-sized vector, from:

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

V is *(NT)´(NT)*-sized matrix as follow:

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

where
$$\mathbf{V}_i = \begin{pmatrix} \langle \xi_{i1}, \xi_{i1} \rangle & 0 & \cdots & 0 \\ \langle \xi_{i2}, \xi_{i1} \rangle & \langle \xi_{i2}, \xi_{i2} \rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \langle \xi_{iT}, \xi_{i1} \rangle & \langle \xi_{iT}, \xi_{i2} \rangle & \cdots & \langle \xi_{iT}, \xi_{iT} \rangle \end{pmatrix}, \tag{11}$$

$$\langle \xi_{it}, \xi_{is} \rangle = \int_a^b \frac{(x_{it}-u)_+^{m-1}(x_{is}-u)_+^{m-1}}{\left((m-1)!\right)^2} du, t = 1,2,\dots,T; s = 1,2,\dots,T \tag{4.8}$$

$\underset{\sim}{c}$ is NT-sized vector, from:

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

Nonparametric regressionanalysis is conducted to get estimator of regression curve $\underset{\sim}{f}$. To get the estimation, Reproducing Kernel Hilbert Space (RKHS) is used. The purpose is to obtain the estimation of $\underset{\sim}{f}$ that meets PWLS optimization [10,11]:

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

with restricted:

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

Then, space function $\mathcal{H} = W_2^m[a_i, b_i]$ used is order-2 Sobolev space defined as follow:

$$W_2^m[a_i, b_i] = \left\{ f_i : \int_{a_i}^{b_i} [f_i^{(m)}(x_{it})]^2 \, dx_{it} < \infty \right\},$$

Where $a_i \leq x_{it} \leq b_i$ and $i = 1,2,...,N$. Based on the space, nornm of every $f_i \in W_2^m[a_i, b_i]$ 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:

$$\min_{\substack{f_i \in W_2^m[a_i, b_i] \\ i=1,2,..,N}} \left\{ \left\| \mathbf{\Sigma}^{-\frac{1}{2}} \underset{\sim}{\varepsilon} \right\|^2 \right\} = \min_{\substack{f_i \in W_2^m[a_i, b_i] \\ i=1,2,..,N}} \left\{ \left\| \mathbf{\Sigma}^{-\frac{1}{2}} (\underset{\sim}{y} - \underset{\sim}{f}) \right\|^2 \right\}, \tag{14}$$

With restricted in equation (15) into:

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

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

$$\min_{\substack{f_i \in W_2^m[a_i, b_i], \\ i=1,2,...,N}} \left\{ M^{-1} \left(\underset{\sim}{y} - \underset{\sim}{f} \right)^{\mathrm{T}} \mathbf{\Sigma}^{-1} \left(\underset{\sim}{y} - \underset{\sim}{f} \right) + \sum_{i=1}^{N} \lambda_i \int_{a_i}^{b_i} \left[f_i^{(m)}(x_{it}) \right]^2 dx_{it} \right\}, \tag{16}$$

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 :

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

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

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

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:

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

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

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

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

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

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

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

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

Solving optimization (21) is obtained by conducting derivative $\mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})$ partially towards $\underset{\sim}{c}$ 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:

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

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

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

When matrix **U** is presented as:

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

equation (23) can be stated as:

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

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

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

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

Furthermore, partial derivative:

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

results in:

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

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

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

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

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 Σ^{-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:

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

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

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

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

Error variance-covariance matrix $\hat{\mathbf{\Sigma}}$ will be presented in the next section (Theorem 3), 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

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

Theorem 2: Estimation of Curve Regression Without Penalty (WLS)

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 WLS

$$\min_{f_i \in \mathrm{W}_2^m[a_i,b_i],\, i=1,2,\ldots,N} \left\{ M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)\right\} \tag{5}$$

is

$$\hat{\underset{\sim}{f}} = \mathbf{A}\underset{\sim}{y},\ \text{with:}$$

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}. \tag{6}$$

Proof: Considering the equation that is function $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d}$,then the nonparametric regression model (1) can be stated as [2-7]:

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

The purpose is to obtain the estimation of $\underset{\sim}{f}$ that meets WLS optimization [11]:

$$\min_{\substack{f_i \in \mathcal{H} \\ i=1,2,\ldots,N}} \left\{ \left\| \mathbf{\Sigma}^{-\frac{1}{2}} \underset{\sim}{\varepsilon} \right\|^2 \right\} = \min_{\substack{f_k \in \mathcal{H} \\ i=1,2,\ldots,N}} \left\{ \left\| \mathbf{\Sigma}^{-\frac{1}{2}} (\underset{\sim}{y} - \underset{\sim}{f}) \right\|^2 \right\}, \quad (22)$$

Or

$$M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}} \mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right) = M^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)^{\mathrm{T}} \mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right). \quad (23)$$

with $M=NT$. Solve WLS using goodness of fit (23) without penalty, as follow result:

$$\begin{aligned}
&\min_{\underset{\sim}{d} \in \Re^{Nm}} \left\{ M^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)^{\mathrm{T}} \mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right) \right\} \\
&= \min_{\underset{\sim}{d} \in \Re^{Nm}} \left\{ \left(\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)^{\mathrm{T}} \mathbf{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)\right) M^{-1} \right\} \\
&= \min_{\underset{\sim}{d} \in \Re^{Nm}} \left\{ \left[(\underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} - \underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d} - \underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d} \right] M^{-1} \right\} \\
&= \min_{\underset{\sim}{d} \in \Re^{Nm}} \left\{ \left[(\underset{\sim}{y}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} - 2\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d} \right] M^{-1} \right\} \\
&= \min_{\underset{\sim}{d} \in \Re^{Nm}} \left\{ \mathrm{Q}(\underset{\sim}{d}) \right\}. \quad (24)
\end{aligned}$$

Solve the optimation of (24), coming from differential of $\mathrm{Q}(\underset{\sim}{d})$ by $\underset{\sim}{d}$, and equally to zero, as follow:

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

The result:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} = \underset{\sim}{0} \quad (25)$$

The the solving of curve estimation from (25) as follow:

$$\hat{\underset{\sim}{f}} = \mathbf{T}\hat{\underset{\sim}{d}}$$

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

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

$$\hat{\underset{\sim}{f}} = \mathbf{A}\underset{\sim}{y}, \quad (26)$$

with

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}..$$

The second purpose of the study us to estimate error variance matrix as weighting in PWLS or WLS. In single-response case, there is weighting that accommodates correlation between responses (Fernandes, [2-3]). Estimation for error variance-covarian matrix shown in Theorem 3.

Theorem 3

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

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Sigma}}_{NN} \end{bmatrix} \tag{27}$$

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:

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

$$L(\underset{\sim}{f}, \mathbf{\Sigma}|\underset{\sim}{y}) = \frac{1}{(2\pi)^{TM/2}|\mathbf{\Sigma}|^{T/2}}\exp\left(-\frac{1}{2}\sum_{t=1}^{T}(\underset{\sim}{y} - \underset{\sim}{f})'\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \underset{\sim}{f})\right) \tag{15}$$

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

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

$$\sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})'\boldsymbol{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f}) = \sum_{t=1}^{T}\text{tr}\left[\boldsymbol{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right]$$
$$= \text{tr}\left[\boldsymbol{\Sigma}^{-1}\sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right] \quad (16)$$

Thus, equation (15) is substituted using equation (16) and the result is as follow:

$$L(\underset{\sim}{f},\boldsymbol{\Sigma}|\underset{\sim}{y}) = \frac{1}{(2\pi)^{TM/2}|\boldsymbol{\Sigma}|^{T/2}}\exp\left(-\frac{1}{2}\text{tr}\left[\boldsymbol{\Sigma}^{-1}\sum_{t=1}^{T}(\underset{\sim}{y}-\underset{\sim}{f})(\underset{\sim}{y}-\underset{\sim}{f})'\right]\right) \quad (17)$$

Estimator for variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ is obtained by maximizing function of $L(\underset{\sim}{f},\boldsymbol{\Sigma}|\underset{\sim}{y})$, through $\frac{\partial L(\underset{\sim}{f},\boldsymbol{\Sigma}|\underset{\sim}{y})}{\partial\boldsymbol{\Sigma}} = \mathbf{0}$. As mentioned in Fernandes [4-5], likelihood function in equation (16) will meet the maximum condition if $\hat{\boldsymbol{\Sigma}} = \frac{1}{2b}\mathbf{B}$, 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:

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{2b}\mathbf{B}$$
$$\hat{\boldsymbol{\Sigma}} = \frac{\sum_{t=1}^{T}(\underset{\sim}{y}-\hat{\underset{\sim}{f}})(\underset{\sim}{y}-\hat{\underset{\sim}{f}})'}{T} \quad (18)$$

Random error variance-covariance matrix in the study is similar to equation (18) or can be reformulated as:

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

Hence, equation (17) can be reformulated as:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Based on the elaboration of equation (18), estimation of $\hat{\boldsymbol{\Sigma}}_{11}$ is as follow:

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

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

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

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

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

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

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

Or it can be formulated that:

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} \hat{\mathbf{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{\Sigma}}_{NN} \end{bmatrix} \tag{19}$$

with

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

Estimation of variance-covariance matrix in equation (19) can be used to predict regression curve equation (13).

The theoretical findings above resulted in the estimation of the smoothing spline nonparametric regression curve for PWLS based longitudinal data (with penalty) in equation (21) and for the WLS based longitudinal data (without penalty) in equation (26). The application on the data used the weighted estimation in equation (27). The data used in this research is the data of the babies visiting Dinoyo Community Health Center of Malang City using *Kartu Menuju Sehat* KMS. The data only involved $N = 4$ babies. The babies describe the baby growth aged 0-24 months. The observed response is the weight of the baby (y) in several monthly-observation periods for 24 months. The predictor used in this research is the age of the baby (x). Table 1 is the output of coefficient estimation for PWLS (including coefficients c and d), and WLS (including only coefficient c) methods.

Table 1

The values of $\hat{d}$ and $\hat{c}$

	x1	*x2*	*x3*	*x4*
d1	0,14364	0,14992	0,18288	0,23721
d2	0,94124	2,07205	0,88867	0,82028
c1	-0,23115	-51,1723	-0,19994	-0,20281
c2	-0,32721	-25,9204	-0,21121	-0,26378
c3	-0,20049	33,6623	-0,06852	-0,13940
c4	-0,08740	54,6730	-0,01119	-0,02170
c25	-0,02962	0,03788	-0,04292	-0,02244

The following is the comparison between the PWLS and WLS based smoothing spline nonparametric regression model. Figure 1 is the estimation of smoothing spline nonparametric regression curve with WLS (red line) and PWLS (green line).

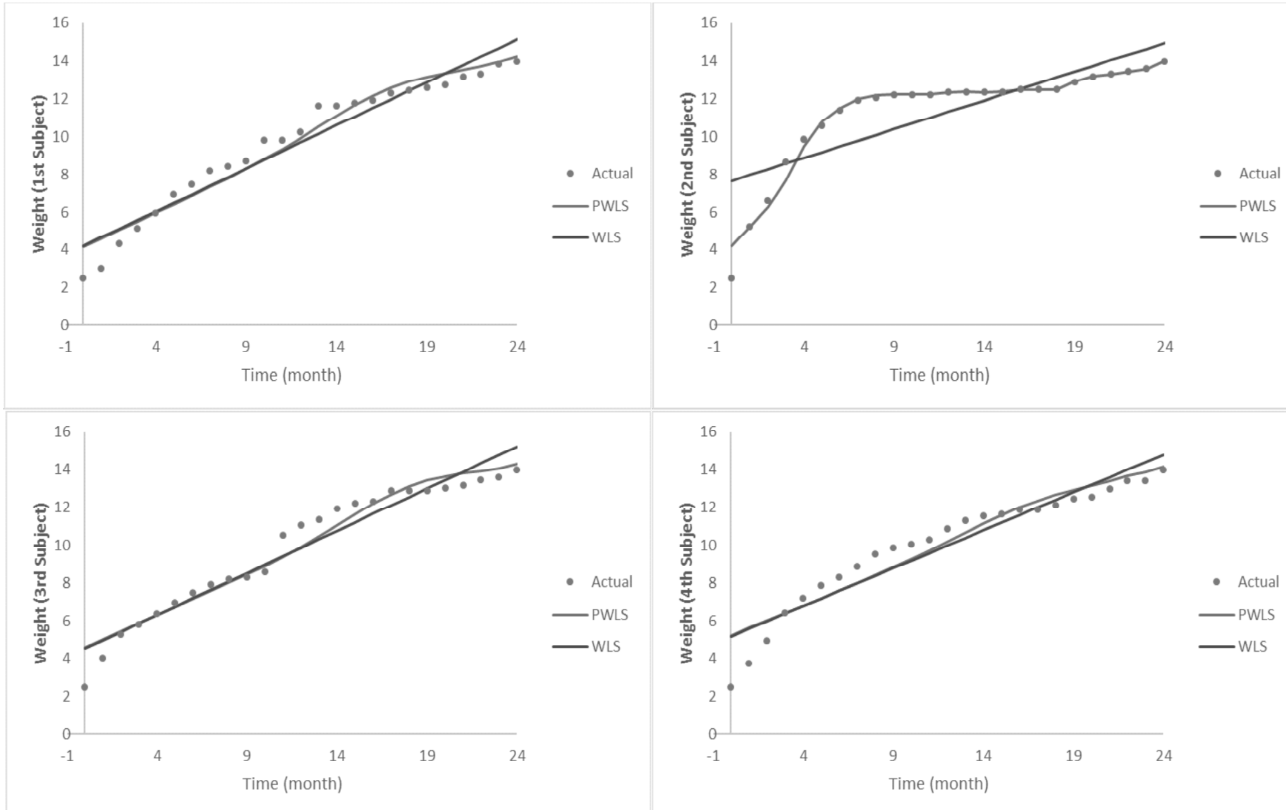

Figure 1: Estimation of the Curve of Smoothing Spline Nonparametric Regression Using WLS and PWLS

The comparison between the estimation results of the WLS and PWLS nonparametric regression shows that the prediction values obtained in the PWLS nonparametric regression model are near to the actual data, compared to the prediction values in WLS approach. It can be proved from the result of coefficient value of the determination on the spline nonparametric regression model for 91,1%. The estimation of WLS smoothing spline nonparametric regression model, which equals to parametric regression, that is global, has not yet estimated every observation point in depth. It is different from the estimation of PWLS based smoothing spline nonparametric regression model that is more local, so that it can estimate every observation point each other in detail very well.

The estimator is said to be efficient if it has a minimum error variance. The efficiency of the estimator is the ration of minimum error variance of the estimator. Meanwhile, the efficiency relative is the ratio of the error variance of both compared estimators. For example, $\hat{g}_{PWLS}(x)$ and $\hat{g}_{WLS}(x)$ are the two estimators of the smoothing spline nonparametric regression of $g(\theta)$. If both estimators follow the general condition of Cramer-Rao, the Efficiency Relative (ER) of $\hat{g}_{PWLS}(x)$ and $\hat{g}_{WLS}(x)$ is defined as the ratio of the error variance as follow;

$$ER_{\theta}(\hat{g}_{PWLS}, \hat{g}_{WLS}) = \frac{MSE_{PWLS}(\theta)}{MSE_{WLS}(\theta)}$$

If $ER_{\theta}(\hat{g}_{PWLS}, \hat{g}_{WLS}) < 1$, then $\hat{g}_{PWLS}(x)$ is more efficient than $\hat{g}_{WLS}(x)$.

Table 2
Efficiency Relative of Two Methods

Subject	*MSE(PWLS)*	*MSE(WLS)*	*ER*
1	0.148	0.508	0.291
2	0.176	2.731	0.065
3	0.506	0.713	0.709
4	0.828	0.949	0.872
Overall			0.484

Table 2 above shows that those four PWLS based regression curves are more efficient than the WLS based regression curves, of which have only 6,5% efficiency compared to the PWLS curves. The highest similarity is on the subjects. In contrast, in the fourth subject, the PWLS and WLS estimations tend to be similar, with the Efficiency Relative near 100%. Overall, the estimation of the curve of PWLS based nonparametric regression is more efficient than the estimation of the curve of WLS based nonparametric regression. It can be seen from the efficiency of WLS based curve that is only 48.4% or less than 50%, compared to the efficiency of the PWLS based curve. In other word, the use of estimation of the curve of penalty (PWLS) based smoothing spline nonparametric regression has a better efficiency level than the WLS (without penalty) based.

4. CONCLUSIONS AND RECCOMENDATION

Based on the analysis result and discussion above, the conclusion of this research are follow:

(1) The estimation of smoothing spline nonparametric regression PWLS (with penalty) as follow:

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

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

Without penalty using WLS as follow:

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

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}.$$

(2) The estimation of error variance-covariance matrix is as follow:

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\Sigma}}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\boldsymbol{\Sigma}}_{NN} \end{bmatrix} \text{ with } \hat{\boldsymbol{\Sigma}}_{ij} = \frac{(\underset{\sim}{y}_i - \hat{\underset{\sim}{f}}_i)(\underset{\sim}{y}_j - \hat{\underset{\sim}{f}}_j)'}{T},$$

(3) The estimation of the curve of PWLS based nonparametric regression in the data of baby growth is more efficient than the estimation of the curve of WLS based nonparametric regression. It can be seen from the efficiency of WLS based curve that is only 48.4% or less than 50%, compared to the efficiency of the

PWLS based curve. In other word, the use of estimation of the curve of penalty (PWLS) based smoothing spline nonparametric regression has a better efficiency level than the WLS (without penalty) based.

The problem in this research is that it has not yet estimated the coefficient of the autocorrelation correctly, so that there is a need to conduct further research that is able to accommodate the estimation of the coefficient of autocorrelation, as well as to improve the efficiency in the simulation with different autocorrelation levels.

REFERENCES

[1] Verbekke, G., and Molenberghs, G. (2000), *Linear Mixed Model for Longitudinal Data*. Springer Series in statistics. New York: Springer Verlag.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] Budiantara, I.N, Ratnasari, V., Ratna, M., & Zain, I. (2015), The Combination of Spline and Kernel Estimator for Nonparametric Regression and its Properties, *Applied Mathematical Science*, Vol 9, No 122, 6083-6094.

[7] Lestari, B., Budiantara, I.N., Sunaryo, S., and Mashuri M. (2010), Spline Estimator in Multiresponse Nonparametric Regression Model with Unequal Correlation of Errors. *Journal of Mathematics and Statistics*, 6(3): 327-332.

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

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

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

[12] 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.

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

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

[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.