International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 3 Issue 10, October – 2019, Pages: 1-4

Estimation of Multiresponse Nonparametric Regression Model Using Smoothing Spline Estimator: A Simulation Study

Budi Lestari
Department of Mathematics, Faculty of Mathematics and Natural Sciences, The University of Jember
Jalan Kalimantan 37 Kampus Tegal Boto, Jember 68121 Indonesia
Doctoral Study Program, Faculty of Science and Technology, Airlangga University
Jalan Mulyorejo Kampus C UNAIR Surabaya 60115 Indonesia
lestari.statistician@gmail.com and budi.lestari-2016@fst.unair.ac.id

Fatmawati
Department of Mathematics, Faculty of Science and Technology, Airlangga University
Jalan Mulyorejo Kampus C UNAIR Surabaya 60115 Indonesia
Corresponding Author's E-Mail: fatma47unair@gmail.com and fatmawati@fst.unair.ac.id

I Nyoman Budiantara
Department of Mathematics, Faculty of Mathematics, Computing and Data Science, Sepuluh Nopember Institute of Technology
Jalan Arif Rahman Hakim, Kampus Sukolilo, Keputih, Surabaya 60111 Indonesia
nyomanbudiantara65@gmail.com

Abstract: *Estimation of a regression model is equivalent to estimate a regression function of the regression model. The regression function draws association between response variable and predictor variable. Multiresponse nonparametric regression model provides powerful tools for modeling the function which represents association between response variable and predictor variable where there are correlations between responses. It means that estimating regression function is the main problem in the multiresponse nonparametric regression model. Smoothing spline estimator has powerful and flexible properties for estimating the regression function in this model. Therefore, in this paper we present the simulation study to determine regression function estimate of multiresponse nonparametric regression model based on smoothing spline estimator. In this simulation study we apply the smoothing spline estimator to three multiresponse nonparametric regression models represented by trigonometric, polynomial, and exponential models. The result shows that based on plots of estimation results, the smoothing spline estimator has good performance and flexibility for estimating the regression function of the multiresponse nonparametric regression models.*

Keywords — Exponential Model, Generalized Cross Validation, Multiresponse Nonparametric Regression, Polinomial Model, Smoothing Spline Estimator, Trigonometric Model.

1. INTRODUCTION

Research on smoothing spline models has attracted a great deal of attention in recent years, and the methodology has been widely used in many areas. Smoothing spline estimator with its powerful and flexible properties is one of the most popular estimators used for estimating regression function in the nonparametric regression model as well as in the multiresponse nonparametric regression model. Researchers who have considered spline estimator for estimating regression function of the nonparametric regression and multiresponse nonparametric regression models are [1-3] who have used original spline estimator to estimate regression function of smooth data, [4-5] who have introduced M-type spline to overcome outliers in nonparametric regression, [6] who has constructed confidence interval for original spline model by using Bayesian approach, [7] who compared generalized cross validation (GCV) and generalized maximum likelihood (GML) for choosing the smoothing parameter in the generalized spline smoothing problem, [8-9] who introduced relaxed spline and quantile spline, [10] who has studied smoothing spline models with correlated random errors, [11] who introduced some techniques for spline statistical model building by using reproducing kernel Hilbert spaces, [12] who proposed a method that combines smoothing spline estimates of different smoothness to form a final improved estimate, [13] who investigated asymptotic property of smoothing splines estimators in functional linear regression with errors-in-variables, [14] who have studied smoothing spline estimation of variance functions, [15] who showed goodness of spline estimator rather than kernel estimator in estimating nonparametric regression model for gross national product data, [16] who have studied the determination of an optimal smoothing parameter for nonparametric regression using smoothing spline, [17] who has studied spline smoothing for estimating nonparametric functions from bivariate data with the same correlation of errors, [18] who proposed methods for estimating nonparametric regression model with spatially correlated errors, [19] and [20] have studied spline estimators in multiresponse nonparametric regression model with equal correlation of errors and unequal correlation of errors, respectively, [21] who have used multiresponse nonparametric regression model approach to design child growth chart, [22] who proposed a mathematical statistics method for estimating regression function of the multiresponse nonparametric regression model in case of heteroscedasticity of variance, [23] who discussed estimating regression function of the homoscedastic multiresponse nonparametric regression in which the number of observations were unbalanced, [24-26] who proposed smoothing spline estimator for estimating of the multiresponse nonparametric regression model by using reproducing kernel Hilbert space (RKHS), [27] who have discussed construction of covariance matrix in case of homoscedasticity of variances of errors, [28] who discussed estimating of both covariance matrix and optimal smoothing parameter, and [29] who used spline to show ability of covariance matrix.

These researchers mentioned above have not discussed about estimation of multiresponse nonparametric regression model via a simulation study for some type of multiresponse nonparametric regression models. Therefore, in this paper we give numerical example as an illustration of smoothing spline estimator in estimating the regression function of the three model types of multiresponse nonparametric regression models.

2. MATERIAL AND METHODS

In this simulation study we apply three model types of multiresponse nonparametric regression model where in every model consists of a predictor variable (t) and three response variables (y_1, y_2, y_3). Data ($n=100$) is generated from these three model types, i.e., trigonometric model, polinomial model, and exponential model. These models are given in Eq. (1), Eq. (2), and Eq. (3), respectively.

Trigonometric Model:

$$\left.\begin{aligned} y_{1i} &= 5+3\sin(2\pi t_{1i}^2)+\varepsilon_{1i}, \quad i=1,2,...n \\ y_{2i} &= 3+3\sin(2\pi t_{2i}^2)+\varepsilon_{2i}, \quad i=1,2,...n \\ y_{3i} &= 1+3\sin(2\pi t_{3i}^2)+\varepsilon_{3i}, \quad i=1,2,...n \end{aligned}\right\} \qquad (1)$$

Polinomial Model:

$$\left.\begin{aligned} y_{1i} &= 5+3t_{1i}^2+\varepsilon_{1i}, \quad i=1,2,...n \\ y_{2i} &= 3+3t_{2i}^2+\varepsilon_{2i}, \quad i=1,2,...n \\ y_{3i} &= 1+3t_{3i}^2+\varepsilon_{3i}, \quad i=1,2,...n \end{aligned}\right\} \qquad (2)$$

Exponential Model:

$$\left.\begin{aligned} y_{1i} &= \frac{1}{\sqrt{2\pi}}\exp(t_{1i}^2)+\varepsilon_{1i}, \quad i=1,2,...n \\ y_{2i} &= 0.5\frac{1}{\sqrt{2\pi}}\exp(t_{2i}^2)+\varepsilon_{2i}, \quad i=1,2,...n \\ y_{3i} &= 0.4\frac{1}{\sqrt{2\pi}}\exp(t_{3i}^2)+\varepsilon_{3i}, \quad i=1,2,...n \end{aligned}\right\} \qquad (3)$$

The estimated regression function can be obtained by taking solution of penalized weighted least square optimization by using reproducing kernel Hilbert space approach as has been discussed by [21-26, 30].

3. RESULTS AND DISCUSSION

In this section, we give results and discussion about estimating regression functions of trigonometric, polynomial, and exponential models.

3.1. Estimation of Trigonometric Model

In this simulation, for trigonometric model in (1) we take sample size of $n=100$, and correlation values $\rho_{12}=0.5$, $\rho_{13}=0.7$, $\rho_{23}=0.4$, and variances $\sigma_1^2=\sigma_2^2=\sigma_3^2=1$. The following figure (Fig. 1) gives plots of estimated response ($\hat{y}$) versus predictor (t) for simulation data of trigonometric model in (1), i.e., plot of $\hat{y}_1$ versus t_1 (above), plot of $\hat{y}_2$ versus t_2 (center), and plot of $\hat{y}_3$ versus t_3 (below).

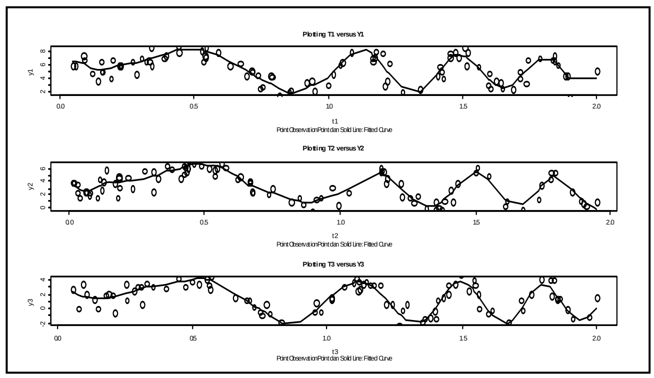

Figure 1. Plots of estimation curves for trigonometric model.

Based on the simulation results, for trigonometric model we obtain estimated optimal smoothing parameters of $\hat{\lambda}_1=6.886826\times10^{-8}$ (for the first response, $\hat{y}_1$), $\hat{\lambda}_2=7.748048\times10^{-8}$ (for the second response, $\hat{y}_2$), and $\hat{\lambda}_3=8.303897\times10^{-8}$ (for the third response, $\hat{y}_3$). Also, we get minimum generalized cross validation (GCV) value of 1.473032.

3.2. Estimation of Polinomial Model

Similar to section 3.1, in this simulation, for polynomial model in (2) we also take sample size of $n=100$, and correlation values $\rho_{12}=0.5$, $\rho_{13}=0.7$, $\rho_{23}=0.4$, and variances $\sigma_1^2=\sigma_2^2=\sigma_3^2=1$. Figure 2 gives plots of estimated response ($\hat{y}$) versus predictor (t) for simulation data of polinomial model in (2), i.e., plot of $\hat{y}_1$ versus

t_1 (above), plot of $\hat{y}_2$ versus t_2 (center), and plot of $\hat{y}_3$ versus t_3 (below).

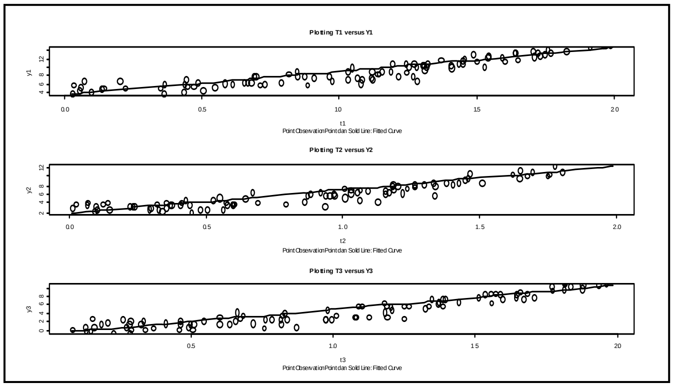

Figure 2. Plots of estimation curves for polinomial model.

Based on the simulation results, for trigonometric model we obtain estimated optimal smoothing parameters of $\hat{\lambda}_1 = 1$ (for the first response, $\hat{y}_1$), $\hat{\lambda}_2 = 1$ (for the second response, $\hat{y}_2$), and $\hat{\lambda}_3 = 0.2539217$ (for the third response, $\hat{y}_3$). Also, we get minimum generalized cross validation (GCV) value of 1.790529.

3.3. Estimation of Exponential Model

Similar to sections 3.1 and 3.2, in this simulation, for exponential model in (3) we also take sample size of $n = 100$, and correlation values $\rho_{12} = 0.5$, $\rho_{13} = 0.7$, $\rho_{23} = 0.4$, and variances $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$. Figure 3 gives plots of estimated response ($\hat{y}$) versus predictor (t) for simulation data of exponential model in (3), i.e., plot of $\hat{y}_1$ versus t_1 (above), plot of $\hat{y}_2$ versus t_2 (center), and plot of $\hat{y}_3$ versus t_3 (below).

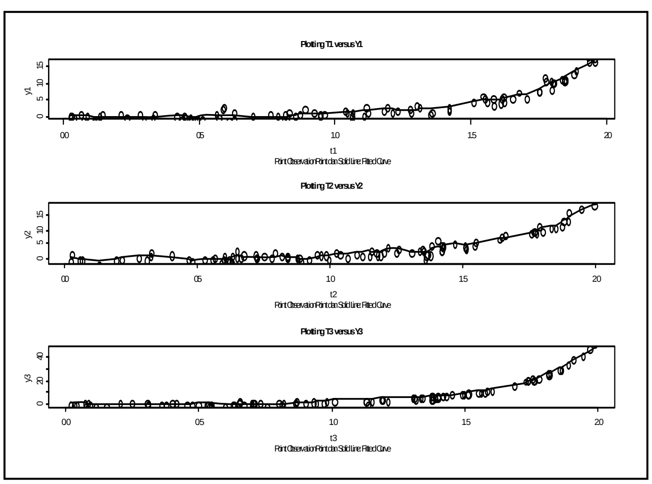

Figure 3. Plots of estimation curves for exponential model.

Based on the simulation results, for trigonometric model we obtain estimated optimal smoothing parameters of $\hat{\lambda}_1 = 1.295002 \times 10^{-6}$ (for the first response, $\hat{y}_1$), $\hat{\lambda}_2 = 6.890408 \times 10^{-8}$ (for the second response, $\hat{y}_2$), and $\hat{\lambda}_3 = 7.628884 \times 10^{-7}$ (for the third response, $\hat{y}_3$). Also, we get minimum generalized cross validation (GCV) value of 1.88195.

4. CONCLUSION

Based on plots of estimation results given in Fig. 1-3 that present plots of estimation curves of multiresponse nonparametric regression models from types of trigonometric, polynomial, and exponential models show that the smoothing spline estimator has good performance and flexibility for estimating the regression function of the multiresponse nonparametric regression models. It means that the smoothing spline estimator is suitable to use in estimating multiresponse nonparametric regression models from all types of mathematical models.

5. ACKNOWLEDGMENT

Authors thank to Director of the Directorate of Research and Public Service, the Directorate General of Reinforcing of Research and Development, the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for financial support of this research via the Doctoral Dissertation Research (*Penelitian Disertasi Doktor - PDD*) Grant in the fiscal year 2019.

6. REFERENCE

[1] Kimeldorf, G., and Wahba, G. (1971). Some result on tchebycheffian spline functions. J. of Math. Anal. and Appl., Vol. 33, pp. 82-95.

[2] Craven, P., and Wahba, G. (1979). Smoothing noisy data with spline function: Estimating the correct degree of smoothing by the method of generalized cross validation. Numer. Math., Vol. 31, pp. 377-403.

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

[4] Cox, D. D. (1983). Asymptotic for M-type smoothing spline. Annals. Statistics. Vol. 11, pp. 530-551.

[5] Cox, D. D., and O'Sullivan, F. (1996). Penalized likelihood type estimators for generalized nonparametric regression. J. Mult. Anal., Vol. 56, pp. 185-206.

[6] Wahba, G. (1983). Bayesian confidence intervals for the cross-validated smoothing spline. J. Royal Stat. Society, Series B., Vol. 45, pp. 133-150.

[7] Wahba, G. (1985). A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. Annals Stat., Vol. 13, pp. 1378-1402.

[8] Oehlert, G. W. (1992). Relaxed boundary smoothing spline. Annals. Statistics, Vol. 20, pp. 146-160.

[9] Koenker, R., Pin, N. G., and Portnoy, S. (1994). Quantile smoothing splines. Biometrics, Vol. 81, pp. 673-680.

[10] Wang, Y. (1998). Smoothing spline models with correlated random errors. J. Amer. Stat. Assoc., Vol. 93, pp. 341-348.

[11] Wahba, G. (2000). An introduction to model building with reproducing kernel hilbert spaces. Available at: http://www.stat.wisc.edu/~wahba/ftp1/interf/rootpar1r.pdf

[12] Lee, T., C., M. (2004). Improved smoothing spline regression by combines estimates of different smoothness. Stat. & Prob. Letters, Vol. 67, pp. 133-140.

[13] Cardot, Crambes, H.C., Kneip, A., and Sarda, P. (2007). Smoothing splines estimators in functional linear regression with errors-in-variables. Computational Stat.Data Anal., Vol. 51, pp. 4832-4848.

[14] Liu, A., Tong, T., and Wang, Y. (2007). Smoothing Spline Estimation of Variance Functions. J. Comput. Graphical Stat., Vol. 16, pp. 312-329.

[15] Aydin, D. (2007). A comparison of the nonparametric regression models using smoothing spline and kernel regression. World Acad. Sci. Eng. Technol., Vol. 26, pp. 730-734.

[16] Aydin, D., Memmedhi, M., and Omay, R.E. (2013). Smoothing parameter selection for nonparametric regression using smoothing spline. European J. of Pure & Appl. Math., Vol. 6, pp. 222-238.

[17] Wang, Y., Guo, W., and Brown, W. B. (2000). Spline smoothing for bivariate data with applications to association between hormones. Statistica Sinica, Vol. 10, pp. 377-397.

[18] Fernandez, F. M., and Opsomer, J. D. (2005). Smoothing parameter selection methods for nonparametric regression with spatially correlated errors. Canadian J. Statistics, Vol. 33, pp, 279-295.

[19] Lestari, B., Budiantara, I. N., Sunaryo, S., and Mashuri, M. (2009). Spline estimator in homoscedastic multi-response nonparametric regression model. Proc. Indonesian Math. Soc. Int. Conf. on Math. & Its Appl., Oct. 12-13, Yogyakarta-Indonesia. (Available at: http://digilib.its.ac.id/ITS-Article-40312000750/18579).

[20] Lestari, B., Budiantara, I. N., Sunaryo, S., and Mashuri, M. (2010). Spline estimator in multi-response nonparametric regression model with unequal correlation of errors. J. Math. Stat., Vol. 6, pp. 327-332.

[21] Chamidah, N., Budiantara, I. N., Sunaryo, S., and Zain, I. (2012). Designing of child growth chart based on multi-response local polynomial modeling. J. Math. Stat., Vol. 8, pp. 342-347.

[22] Lestari, B., Budiantara, I. N., Sunaryo, S., and Mashuri, M. (2012). Spline smoothing for multiresponse nonparametric regression model in case of heteroscedasticity of variance. J. Math. and Stat., Vol. 8(3), pp. 377-384.

[23] Chamidah, N., and Lestari, B. (2016). Spline Estimator in Homoscedastic Multi-Response Nonparametric Regression Model in Case of Unbalanced Number of Observations. Far East Journal of Mathematical Sciences (FJMS), Vol. 100, pp. 1433-1453.

[24] Lestari, B., Fatmawati, and Budiantara, I. N. (2017). Estimasi Fungsi Regresi Nonparametrik Multirespon Menggunakan Reproducing Kernel Hilbert Space Berdasarkan Estimator Smoothing Spline. Proceed.ing of National Seminar on Mathematics and Its Applications (SNMA) 2017, Airlangga University, Surabaya, pp. 243-250.

[25] Lestari, B., Fatmawati, Budiantara, I. N., & Chamidah, N. (2018). Estimation of regression function in multiresponse nonparametric regression model using smoothing spline and kernel estimators. Journal of Physics: Conference Series, 1097, 012091. DOI: 10.1088/1742-6596/1097/1/012091.

[26] Lestari, B., Chamidah, N., & Saifudin, T. (2019). Estimasi fungsi regresi dalam model regresi nonparametrik birespon menggunakan estimator smoothing spline dan estimator kernel. Jurnal Matematika, Statistika & Komputasi (JMSK), Vol. 15(2), pp. 20-24. DOI: 10.20956/jmsk.v15i2.5565.

[27] Lestari, B., Anggraeni, D., and Saifudin, T. (2018). Estimation of Covariance Matrix based on Spline estimator in Homoscedastic Multiresponses Nonparametric Regression Model in Case of Unbalance Number of Observations. Far East Journal of Mathematical Sciences, Vol. 108(2), pp. 341-355.

[28] Lestari, B., Fatmawati, and Budiantara, I. N. (2019). Spline Estimator and Its Asymptotic Properties in Multiresponse Nonparametric Regression Model. Songklanakarin Journal of Science and Technology, In Press. (http://rdo.psu.ac.th/sjstweb/ArticleInPress.php).

[29] Islamiyati, A., Fatmawati, Chamidah, N. (2019). Ability of Covariance Matrix in Bi-Response Multi-Predictor Penalized Spline Model Through Longitudinal Data Simulation. International Journal of Academic and Applied Research (IJAAR), Vol. 3(3), pp. 8-11.

[30] Lestari, B., Fatmawati, and Budiantara, I. N. (2019). Smoothing Spline Estimator in Multiresponse Nonparametric Regression for Predicting Blood Pressures and Heart Rate. International Journal of Academic and Applied Research (IJAAR), Vol. 3(9), pp. 1-8.