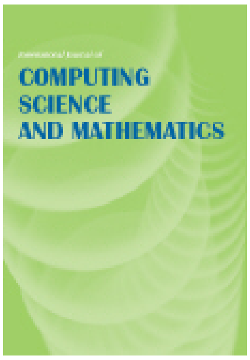
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Computing Science and Mathematics

ISSN online: 1752-5063 - ISSN print: 1752-5055
https://www.inderscience.com/ijcsm

Nonparametric path function estimation of Fourier series at low oscillations for modelling timely paying credit

Adji Achmad Rinaldo Fernandes, Fathiyatul Laili Nur Rasyidah, Atiek Iriany

DOI: 10.1504/IJCSM.2024.10062440

Article History:

Received:	30 September 2021
Last revised:	25 March 2023
Accepted:	30 March 2023
Published online:	22 February 2024

Nonparametric path function estimation of Fourier series at low oscillations for modelling timely paying credit

Adji Achmad Rinaldo Fernandes*,
Fathiyatul Laili Nur Rasyidah and Atiek Iriany

Faculty of Mathematics and Natural Sciences,
Department of Statistics,
Brawijaya University,
Malang City, 65145, Indonesia
Email: fernandes@ub.ac.id
Email: laili.nur.r1@gmail.com
Email: atiekiriany.ub@gmail.com
*Corresponding author

Abstract: The development of the nonparametric path model is carried out with the assumption that each function formed has the same data pattern shape. However, in actual cases, several cases are often encountered where the data patterns formed are different from each of the calculated functions. This research aims to estimate the nonparametric path function of the Fourier series and to describe the lemma and theorem for the analysis of the nonparametric path of the Fourier series at low oscillation levels ($K = 2,3,4,5$). Primary data is obtained from customers at a Bank (Bank X) in Indonesia. The function estimation in nonparametric path analysis using the Fourier series approach is $\hat{\boldsymbol{\alpha}}(\boldsymbol{\lambda}) = \left(\boldsymbol{n}^{-1}\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{\lambda}\boldsymbol{D}\right)^{-1}\boldsymbol{n}^{-1}\boldsymbol{X}'\boldsymbol{y}$. The best nonparametric path model that can describe the 5C variable on Time to Pay through Willingness to Pay is when the oscillation $K = 4$ with R2 is 78%. This study applies the Fourier series approach to path analysis in the banking sector.

Keywords: banking; data; Fourier series; nonparametric path; on-time pay; oscillation; path analysis; statistics; time to pay; willingness to pay.

Reference to this paper should be made as follows: Fernandes, A.A.R., Nur Rasyidah, F.L. and Iriany, A. (2024) ‘Nonparametric path function estimation of Fourier series at low oscillations for modelling timely paying credit’, *Int. J. Computing Science and Mathematics*, Vol. 19, No. 1, pp.67–82.

Biographical notes: Since 2005 until now Adji Achmad Rinaldo Fernandes has worked as a Lecturer at Brawijaya University, Malang. Bachelor degree in Statistics obtained from the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Brawijaya University (2004). The Master of Science degree in Statistics was obtained from the School of Information and Technology, Hasselt University, Diepenbeek, Belgium (2007). The title of Doctor of Statistics was obtained from the Department of Statistics, Faculty of Mathematics and Natural Sciences, Sepuluh Nopember Institute of Technology Surabaya (2016).

Fathiyatul Laili Nur Rasyidah received Bachelor's degree in Statistics from the Department of Statistics, Faculty of Mathematics and Science, Universitas Brawijaya (2021). The title of Master of Science in Statistics was obtained from the Department of Statistics, Brawijaya University (2022).

Atiek Iriany works as a Lecturer at Brawijaya University, Malang. The Bachelor of Agriculture degree was obtained from the field of agricultural science, Universitas Brawijaya (1987). The title of Master of Science in Statistics was obtained from the Department of Statistics, Bogor Agricultural University (1992). Doctoral degree obtained from the field of statistics FMIPA, Airlangga University, Surabaya (2013).

1 Introduction

Statisticians employ a variety of descriptive and inferential techniques to gather, process, present, and analyse data in order to reach broad, practical conclusions. Regression analysis, according to Gujarati and Porter (2012), is the investigation of one variable (the dependent variable) in relation to one or more other factors (the independent variables), with the goal of determining the dependent variable's mean (population) value, means to look at dependencies. A variable derived from the dependent variable's value. Independent variables' known or predetermined values (repeated sampling). When determining whether an independent variable directly affects a dependent variable, regression analysis is used. Regression analysis can be done in three different ways: parametric, semiparametric, and nonparametric (Fernandes, 2021). The development of the nonparametric path model is carried out with the assumption that each function formed has the same data pattern shape, so that some researchers only use one approach. However, in actual cases in the field, there are often cases where the data pattern formed is different from each of the calculated functions. Therefore, researchers want to conduct research using the development of regression analysis, namely path analysis with a nonparametric Fourier series approach in modelling on time to pay credit in the banking sector. This research takes a case study in one of the banks in Indonesia (Bank X). Because the data is confidential, the name of the bank cannot be mentioned.

The requirement for financial support from financial institutions is critical for the welfare and development of corporations. Giving clients credit makes it simpler for them to satisfy their everyday demands. According to Widayati and Putri (2019), one of the advantages of credit is that it promotes economic expansion and growth. One of the financial institutions that provides credit products is a bank. Customers have access to a range of credit options. Home Ownership Loan is one of the financial products that banks provide (KPR). The secret to our company's success is how well we service our loans. When the loan management process is overworked, banks experience issues, one of which lowers the bank's operational profit. The customer must make on-time loan payments in order to mitigate the danger of negative credit. His 5C Principles (Character, Capacity, Equity, Collateral, and Terms) can be used by banks to assess a borrower's ability to make timely loan payments.

According to Fernandes et al. (2017), Wright invented path analysis in 1934. Additionally, based on Solimun's path analysis from 2002, the cause-and-effect model of the relationship between variables is tested. To ascertain whether there is a causal connection between exogenous and endogenous variables, one method is to perform a path analysis. Path analysis is employed to ascertain not only the exogenous variables' direct effects on the endogenous variables, but also the indirect effects that the exogenous variables have on the endogenous variables through the mediating role of the endogenous variables.

Parameter-based path analysis and non-parameter-based path analysis are the two categories under which path analysis is broken down. When linearity assumptions are met, parameter-based route analysis can be utilised, but when they are not, nonlinear and/or nonparametric analysis are two viable studies. I can manage it. Solimun (2010) outlines his six underlying presumptions for path analysis. That is, the variables are related in a linear and additive manner, the residuals are normally distributed, the pattern of the relationship between the variables is recursive, and the link between the variables is minimum intrinsic variables in scales of measurement intervals, (5) Study variables were accurately measured, and (6) Analytical models were chosen in accordance with pertinent theories and notions. The linearity assumption is one that might cause changes to the model. The model's form is affected by linearity assumptions. If the linearity assumptions are satisfied, path analysis is parametric; otherwise, one of two things can happen. When the nonlinearity's shape is known but its shape is unknowable and lacking information, nonlinear path analysis is used. The nonparametric route analysis method is then used to that data pattern. Testing for linearity is a method for identifying relationships between variables. The regression specification error test (RESET) approach is one of them.

Nurjanah et al. (2015) conducted a study on the selection of oscillatory parameters in Fourier series nonparametric regression analysis using high values of the vibration parameter between 60 and 99. The findings demonstrate that, in comparison to 99 oscillations, 70 oscillations can yield a sufficiently high R2, and that numerous parameters, up to 72 parameters, must be evaluated. According to Wisisono et al. (2018), we must estimate various parameters up to 16 using oscillations with values ranging from 1 to 18, as oscillations with a value of 16 can offer noticeably greater R2 than oscillations with a value of 18 is displayed. Soliha et al. (2018) using vibration parameters with values ranging from 3 to 6, vibrations with a value of 3 have a suitably high R2 compared to vibrations with a value of 6, necessitating the estimation of up to 8 parameters has demonstrated the capacity to deliver We can infer from some of these research that high oscillation levels do not always translate into high R2. Therefore, 2, 3, 4, and 5 vibration levels are compared in this study. The amount of characteristics that need to be calculated determines which vibration levels to compare. According to Soliha et al. (2018), vibrational parameters below 10 already produced R2 values that were quite high if the vibrational parameters were too large.

Trigonometric polynomials with flexible properties, such as Fourier series, enable the model to respond to the local characteristics of the data in a useful way. Prahutama (2013) and Sholiha et al. conducted a survey on the modelling of Fourier series using nonparametric regression acted upon (2018). Prahutama (2013) studied the open unemployment rate in East Java using a Fourier series nonparametric regression analysis. Then, to estimate the sales strategy of common Madurese snacks, Sholiha et al. (2018) conducted a nonparametric regression analysis research of Fourier series based on sine and cosine. Sine and cosine data, which typically exhibit periodic functions, can be overcome by Fourier series (Prahutama, 2013), which has the advantage of being able to do so (Wisisono et al., 2018). Periodic refers to a circumstance that happens on a regular basis (Nurjanah et al., 2015). The repeating data pattern, which repeats the values of an endogenous variable for several exogenous variables, is one type of data pattern appropriate for the Fourier series technique (Prahutama, 2013). Based on the foregoing description, the objective of this study is to estimate the Fourier series' nonparametric path functions and to present lemmas and theorems for the analysis of these routes at low vibrational levels. The application of the Fourier series approach to route analysis to simulate the promptness of loan payments in the banking industry is a novel aspect of this study.

2 Literature review

2.1 Parametric regression analysis

In Kutner et al. (2005)'s regression analysis, relationships between two or more quantitative variables are used to predict responses or outcomes using other variables. You can perform a straightforward linear regression analysis if you use predictors to forecast the response variable. An example of a straightforward linear regression model is:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{1}$$

with:

Y_i : the value for the response at the *i*th observation

β_0 : intercept parameter

β_1 : slope parameter

X_i : the value for the predictor variable on the *i*th observation

ε_i : error on observation *i*.

In situations where there are numerous predictor variables, multiple linear regression analysis is used. In cases where more than one predictor variable is used to predict the response variable, Kutner et al. (2005) state that multiple linear regression analysis is one of the most frequently employed statistical techniques. The subsequent are examples of multiple linear regression models' possible expressions:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i \tag{2}$$

The general linear regression model can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip-1} + \varepsilon_i \tag{3}$$

with:

Y_i : the value of the response at the *i*th observation

β_0 : intercept parameter

$\beta_1, \beta_2, \ldots, \beta_p$: slope parameters

$X_{i1}, \dots, X_{ip-1}$: nilai variabel prediktor dalam pengamatan ke-i.

i: 1, 2, ..., n.

n: number of observations

ε_i: error on observation i.

Matrix approaches can be used to resolve the parameter estimation issue for multiple regression analysis with two or more predictors. The generic equation for multiple linear population regression models with up to p–1 predictors is equation (3). The regression equation can be expressed as follows with respect to n observations and p predictors:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_p X_{1p-1} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_p X_{2p-1} + \varepsilon_2 \\
\vdots \quad & \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_p X_{np-1} + \varepsilon_n
\end{aligned} \tag{4}$$

From the above formula, we can write in matrix form:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p-1} \\ 1 & X_{31} & X_{32} & \cdots & X_{3p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \tag{5}$$

A linear regression model in matrix form can be written as:

$$\tilde{Y} = \boldsymbol{X}\tilde{\beta} + \tilde{\varepsilon} \tag{6}$$

Parameter estimation for parametric regression models using the matrix approach is done by minimising the number of squared errors.

$$\tilde{\varepsilon} = \tilde{Y} - \boldsymbol{X}\tilde{\beta} \tag{7}$$

$$\tilde{\varepsilon}'\tilde{\varepsilon} = \left(\tilde{Y} - \boldsymbol{X}\tilde{\beta}\right)'\left(\tilde{Y} - \boldsymbol{X}\tilde{\beta}\right)$$

Then it is derived with the parameter, namely β and equates to zero

$$\tilde{\varepsilon}'\tilde{\varepsilon} = \tilde{Y}'\tilde{Y} - \tilde{Y}'\tilde{\beta}\boldsymbol{X} - \tilde{\beta}'\boldsymbol{X}'Y + \tilde{\beta}'\boldsymbol{X}'\boldsymbol{X}\tilde{\beta}$$

$$\tilde{\varepsilon}'\tilde{\varepsilon} = \tilde{Y}'\tilde{Y} - 2\tilde{\beta}'\boldsymbol{X}'Y + \left(\boldsymbol{X}'\boldsymbol{X}\right)\tilde{\beta}^2$$

$$\frac{\partial \tilde{\varepsilon}'\tilde{\varepsilon}}{\partial \beta} = \frac{\partial \tilde{Y}'\tilde{Y} - 2\tilde{\beta}'\boldsymbol{X}'\tilde{Y} + \left(\boldsymbol{X}'\boldsymbol{X}\right)\tilde{\beta}^2}{\partial \beta}$$

$$\frac{\partial \underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon}}{\partial \beta} = 0 - 2\mathbf{X}'\underset{\sim}{Y} + 2\left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\beta}$$

$$0 = 0 - 2\mathbf{X}'\underset{\sim}{Y} + 2\left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\hat{\beta}}$$

$$2\left(\boldsymbol{X}'\boldsymbol{X}\right)\hat{\tilde{\beta}} = 2\boldsymbol{X}'\tilde{Y}$$

$$\left(\boldsymbol{X}'\boldsymbol{X}\right)\hat{\tilde{\beta}} = \boldsymbol{X}'\tilde{Y}$$

$$\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\left(\boldsymbol{X}'\boldsymbol{X}\right)\hat{\tilde{\beta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\tilde{Y}$$

$$\boldsymbol{I}\hat{\tilde{\beta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\tilde{Y}$$

Therefore, the parameter estimates for multiple linear regression using the matrix approach are:

$$\hat{\tilde{\beta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\tilde{Y} \tag{8}$$

2.2 *Nonparametric regression analysis*

When modelling data patterns, nonparametric regression is a very flexible regression that reduces researcher subjectivity (Soliha et al., 2018). When the typical requirements of parametric regression analysis (normality assumption, non-multicollinearity, and homoscedasticity) are not satisfied, nonparametric regression analysis is applied. In circumstances where there is little or no prior knowledge of the regression curve or data pattern, this strategy works best (Eubank, 1999).

When data patterns are pressed against unknowable data, parametric regression can be dangerous. In other words, it may result in an erroneous regression model, which would bias the results of the hypothesis test. Nonparametric regression models can be employed as follows to uncover patterns in the relationships between predictor variables (*X*) and response variables (*Y*) whose curve shape is unknown (Soliha et al., 2018).

$$Y_i = \hat{f}\left(x_i\right) + \varepsilon_i \tag{9}$$

If the linearity assumption is met, then the analysis is continued by using parametric path analysis by fulfilling the following assumptions $\varepsilon_i \sim N(0,\sigma^2)$. However, if the linearity assumptions are not met, the analysis uses nonlinear and/or nonparametric paths.

where:

Y_i : the value of the response variable on the *i*th observation

x_i : the value of the predictor variable in the *i*th observation.

$\hat{f}$: regression curve.

i: 1, 2, ..., n.

n: number of observations

ε_i: error on observation *i*.

2.3 Path analysis

Path analysis was first developed in 1934 by a geneticist, Sewall Wright. Path analysis is a technique for estimating the effect of a set of independent variables on the dependent variable from a series of observed correlations. The purpose of path analysis is to measure the direct effect on each separate path in the system thereby finding out the extent to which the variation of a given effect can be determined by each cause. Solimun (2010) describes six assumptions that underlie path analysis, namely:

1. The relationship between variables is linear and additive. The assumption of linearity can be checked with a scatter plot, but the results will be subjective. Another way of checking the assumption of linearity is with the Regression Specification Error Test (RESET) introduced by Ramsey in 1969.
2. The errors are evenly spaced out (residual normality). The Kolmogorov-Smirnov test is used to determine whether residuals are normal. Widarjono (2005) claims that if the residuals are normally distributed, testing the influence of predictor variables on the response variable is valid.
3. Recursive relationships exist between the variables (one-way causal flow system). The recursive model has the following characteristics:
   a. Between ε_i are mutually free.
   b. Between ε_i and X_i are free.
4. Minimal endogenous variables in interval measures.
5. Research variables were accurately measured
6. Analysed models are specified using relevant theories and concepts.

The linearity assumption is one that might cause changes to the model. The model's form is affected by linearity assumptions. If the linearity assumptions are satisfied, path analysis is parametric; otherwise, there are two possibilities. Recursive relationships exist between the variables (one-way causal flow system).

2.4 Nonparametric path analysis Fourier series

Fourier series path analysis is a technique for nonparametric path analysis. Trigonometric polynomials called Fourier series are adaptable and can handle the local character of the data well (Wisisono et al., 2018). Data with triangular distributions (sine and cosine) can be addressed by Fourier series, which is advantageous (Prahutama, 2013). Repetition of dependent variable values for various independent variables is a data pattern that is appropriate for the Fourier series technique. The nonparametric path function estimator for the Fourier series is (Tripena, 2009):

$$\hat{f}_\lambda(x_i) = \hat{b}(\lambda)x_i + \frac{1}{2}\hat{a}_0(\lambda) + \sum_{k=1}^{K}\hat{a}_k(\lambda)\cos kx_i \tag{10}$$

3 Methodology

Primary data was used in this study. The data scale used is a second-order Likert scale, primary data was obtained from one of the banks in Indonesia (Bank X). Five exogenous factors, one intervening intrinsic variable, and one intrinsic variable all constituted the confiscation of data, measurement that can be changed based on the average result for each item. Latent variable data was obtained by using the average scale of all indicators for each variable in this approach. After obtaining the latent variables based on the average score of the questionnaire results, the data were analysed by using Fourier series nonparametric path function estimates.

4 Result

4.1 The lemma and theorem of nonparametric path analysis Fourier series

Lemma 4.1: *Form of Fourier Series Nonparametric Path Model at Oscillation Rate = 2. If given paired data (*$\mathbf{X}_{1i}, \mathbf{X}_{2i}, \mathbf{Y}_{1i}, \mathbf{Y}_{2i}$*) with i = 1, 2, ..., n; following the nonparametric pathway analysis model, we obtain the shape of the nonparametric pathway analysis function as shown in equation (11).*

$$\tilde{\hat{f}} = \boldsymbol{X}\tilde{\alpha} \tag{11}$$

Writing the function for each response variable yields equation (12).

$$\begin{aligned} Y_{1i} &= f_1(X_{1i}, X_{2i}) + \varepsilon_{1i} \\ Y_{2i} &= f_2(X_{1i}, X_{2i}, Y_{1i}) + \varepsilon_{2i} \end{aligned} \tag{12}$$

The equation model is as follows

$$\begin{aligned} \hat{f}_{1i} &= \frac{1}{2}a_{01} + b_{11}X_{1i} + \gamma_{11}\cos X_{1i} + \gamma_{21}\cos 2X_{1i} + b_{21}X_{2i} + \gamma_{31}\cos X_{2i} + \gamma_{41}\cos 2X_{2i} \\ \hat{f}_{2i} &= \frac{1}{2}a_{02} + b_{12}X_{1i} + \gamma_{12}\cos X_{1i} + \gamma_{22}\cos 2X_{1i} + b_{22}X_{2i} + \gamma_{32}\cos X_{2i} + \gamma_{42}\cos 2X_{2i} \\ &\quad + b_{32}Y_{1i} + \gamma_{52} cos Y_{1i} + \gamma_{62}\cos 2Y_{1i} \end{aligned} \tag{13}$$

With the form of the matrix X is as follows

$$
\boldsymbol{X} = \begin{pmatrix}
\frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2x_{21} & 0 & 0 & 0 & \cdots & 0 \\
\frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2x_{22} & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2x_{2n} & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2y_{11} \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2y_{12} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2y_{1n}
\end{pmatrix} \tag{14}
$$

And $\underset{\sim}{a}$ are as follows

$$
\tilde{a}^T = \begin{pmatrix} \frac{1}{2}a_{01} & b_{11} & \cdots & \gamma_{41} & \frac{1}{2}a_{02} & b_{12} & \cdots & \gamma_{62} \end{pmatrix} \tag{15}
$$

So the dimension of the matrix for nonparametric path analysis of the series at the oscillation level = 2 is

$$
\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx17}\alpha_{17x1} \tag{16}
$$

where:

$\tilde{f}\left(X_{ij}\right)$: Vector nonparametric regression function of the *i*th observation nonparametric of the *j*th exogenous variable

X_{ij}: The *j*th exogenous variable matrix on the *i*th observation

$\tilde{\alpha}_{ij}$: Parameter vector of the *j*th observation of the *i*th exogenous variable.

4.2 Assumption of linearity

The relationship between the variables Character (X1), Capacity (X2), Capital (X3), Collateral (X4), and Condition on Time (X5) to Obedient Paying Behaviour (Y2) through the Willingness to Pay (Y1) will be examined using path analytical modelling. A linearity test is the first step in the path analysis modelling process because it determines the type of relationship between the variables. To determine whether a relationship between variables is linear or non-linear, use the Ramsey Reset test. The results of the Ramsey Reset test between the variables in the study are shown in Table 1.

Table 1 Linearity assumption test results

Research variable			*p-value*	*Conclusion*	*Research variable*			*p-value*	*Conclusion*
X1	→	Y1	0.909	Linear	X2	→	Y2	0.092	Linear
X2	→	Y1	0.001	Not Linear	X3	→	Y2	0.159	Linear
X3	→	Y1	0.234	Linear	X4	→	Y2	0.517	Linear
X4	→	Y1	0.129	Linear	X5	→	Y2	0.178	Linear
X5	→	Y1	0.029	Not Linear	Y1	→	Y2	0.033	Not Linear
X1	→	Y2	0.130	Linear					

We can see from the results of the aforementioned linearity test that there are three non-linear interactions between the variables: X2 and Y1, X5 and Y1, and Y1 and Y2. Therefore, nonparametric pathway analysis can be used for modelling.

4.3 Comparison of GCV of each lambda on each oscillation

4.3.1 Selection of the best model for each oscillation

With values ranging from 0.2 to 0.9, the nonparametric path function at $K = 2, 3$, and 4 is estimated. First, λ the GCV value will be used to choose the best value. The GCV value will be used to choose the best value. The lowest/minimum GCV number will be the outcome λ in an optimal value. Table 2 shows the GCV value for each value of λ.

Table 2 Selection of the best smoothing parameters

Oscillation	*Lambda*	*GCV*	*Oscillation*	*Lambda*	*GCV*
2	**0.2**	**89937.560**	4	0.2	56588.770
	0.3	1560106.000		0.3	53643.340
	0.4	4598515.000		0.4	52256.130
	0.5	664767.600		0.5	51449.470
	0.6	332864.500		0.6	50922.030
	0.7	227785.800		0.7	50550.240
	0.8	178524.700		0.8	50274.070
	0.9	150476.400		**0.9**	**50060.820**
3	0.2	94060.860	5	0.2	56588.770
	0.3	73789.970		0.3	53643.340
	0.4	66028.350		0.4	52256.130
	0.5	61945.170		0.5	51449.470
	0.6	59429.910		0.6	50922.030
	0.7	57725.930		0.7	50550.240
	0.8	56495.590		0.8	50274.070
	0.9	**55565.640**		**0.9**	**50060.820**

This provides the lowest lambda value at any vibration level in Table 2. The best path analysis model for vibration levels 2, 3, 4, and 5 is also determined using a coefficient of determination calculation. Table 3 displays the coefficient of determination for each oscillation.

Table 3 Each oscillation's coefficient of determination value

Oscillation	*Coefficient of determination*
2	0.762
3	0.775
4	0.780
5	0.770

Using Table 3, we find that $K = 4$ has the highest coefficient of determination (78%). So the best model when oscillating is 4 with a lambda of 0.9 and many parameters in the range 31 to 37 R2. 78% indicates that the variance is correct. Time to payment is 78% explained by personality, customer, capital, collateral, and economic conditions, with the rest explained by other variables. In line with research conducted by Apriana et al. (2017) entitled Analysis of the Effect of Credit Prerequisites (5C) on the Smooth Payment of Customer Installments at Bank Kalsel Antasari Sentra Unit Banjarmasin, the results show that character, and collateral have a significant effect on the smoothness of credit payments. Another study conducted by Wahyuni (2019) entitled Analysis of the Effect of 5C on the Smooth Payment of the Future Package of BTPNS, it was found that economic conditions, character, and capital had a positive effect on the smoothness or accuracy of payments. The following is a nonparametric path analysis function with the best Fourier series approach:

$$\begin{aligned}\hat{f}_1 &= 2.737 + 0.348x_{1i} + 0.001 cos x_{1i} + 0.001 cos 2x_{1i} - 0.001 cos 3x_{1i} + 0.001 cos 4x_{1i} + 0.001 cos 5x_{1i} \\ &+ 0.328x_{2i} + 0.373 cos x_{2i} - 0.001 cos 2x_{2i} + 0.001 cos 3x_{2i} + 0.001 cos 4x_{2i} + 0.001 cos 5x_{2i} + 0.001x_{3i} \\ &- 0.269 cos x_{3i} - 0.194 cos 2x_{3i} + 0.001 cos 3x_{3i} + 0.001 cos 4x_{3i} + 0.001 cos 5x_{3i} + 0.001x_{4i} \\ &+ 0.001 cos x_{4i} - 0.460 cos 2x_{4i} - 0.252 cos 3x_{4i} + 0.001 cos 4x_{4i} + 0.001 cos 5x_{4i} + 0.001x_{5i} \\ &+ 0.001 cos x_{5i} + 0.001 cos 2x_{5i} + 0.166 cos 3x_{5i} + 0.015 cos 4x_{5i} - 0.008 cos 5x_{5i}\end{aligned}$$

$$\begin{aligned}\hat{f}_2 &= 0.001 + 0.001x_{1i} + 0.001 cos x_{1i} + 0.001 cos 2x_{1i} + 0.159 cos 3x_{1i} + 0.125 cos 4x_{1i} + 0.011 cos 5x_{1i} \\ &+ 0.002x_{2i} + 0.001 cos x_{2i} + 0.001 cos 2x_{2i} + 0.001 cos 3x_{2i} - 0.158 cos 4x_{2i} - 0.157 cos 5x_{2i} + 0.030x_{3i} \\ &+ 0.001 \cos x_{3i} + 0.001 cos 2x_{3i} + 0.001 cos 3x_{3i} + 0.001 cos 4x_{3i} - 0.004 cos 5x_{3i} + 0.080x_{4i} + 0.001 \cos x_{4i} \\ &+ 0.001 cos 2x_{4i} + 0.001 cos 3x_{4i} + 0.001 cos 4x_{4i} + 0.001 cos 5x_{4i} + 0.424x_{5i} - 0.403 \cos x_{5i} - 0.014 os 2x_{5i} \\ &+ 0.003 cos 3x_{5i} - 0.001 cos 4x_{5i} + 0.001 cos 5x_{5i} + 0.001y_{1i} - 2.141 \cos y_{1i} - 1.259 cos 2y_{1i} - 0.010 cos 3y_{1i} \\ &+ 0.002 cos 4y_{1i} - 0.001 cos 5y_{1i}\end{aligned}$$

The actual and predicted data plots on the nonparametric path function of the Fourier Series when the oscillation is equal to 4 can be seen in Figures 1–11.

Figure 1 The plot of actual and predicted data between variables X_1 and Y_1 (see online version for colours)

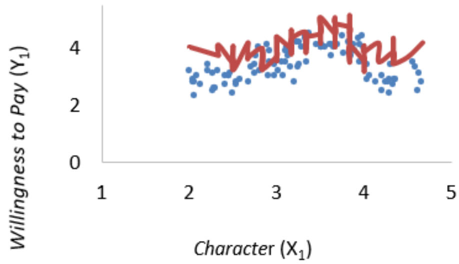


Figure 2 The plot of actual and predicted data between variables X_2 and Y_1 (see online version for colours)

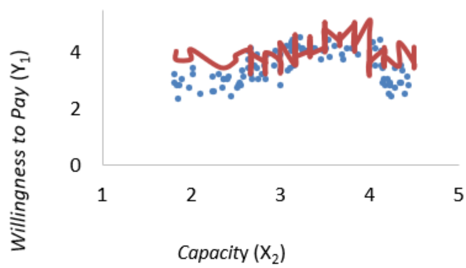


Figure 3 The plot of actual and predicted data between variables X_3 and Y_1 (see online version for colours)

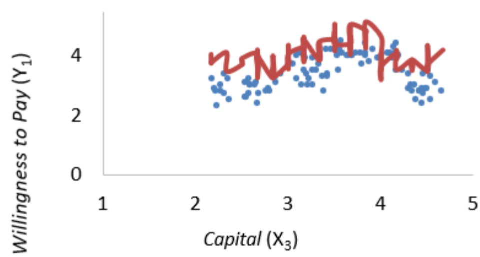


Figure 4 The plot of actual and predicted data between variables X_4 and Y_1 (see online version for colours)

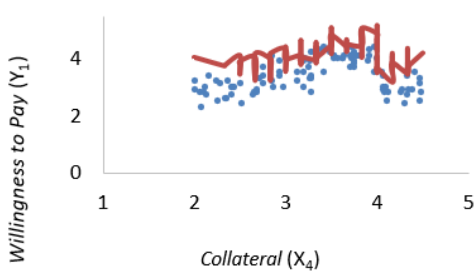


Figure 5 The plot of actual and predicted data between variables X_5 and Y_1 (see online version for colours)

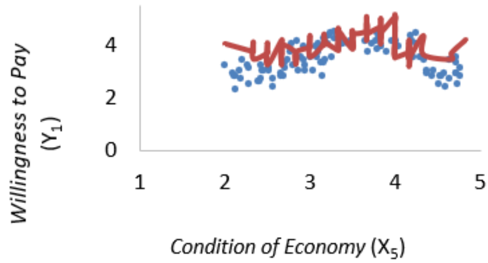


Figure 6 The plot of actual and predicted data between variables X_1 and Y_2 (see online version for colours)

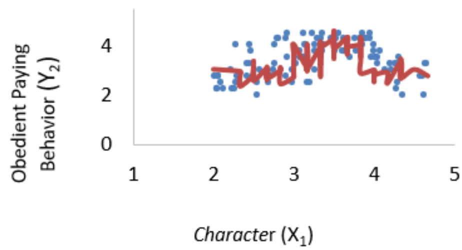


Figure 7 The plot of actual and predicted data between variables X_2 and Y_2 (see online version for colours)

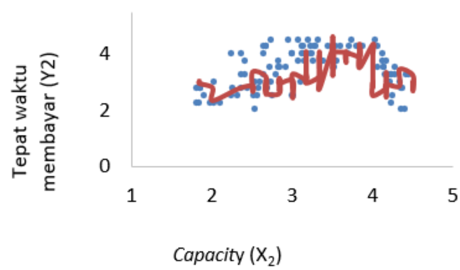


Figure 8 The plot of actual and predicted data between variables X_3 and Y_2 (see online version for colours)

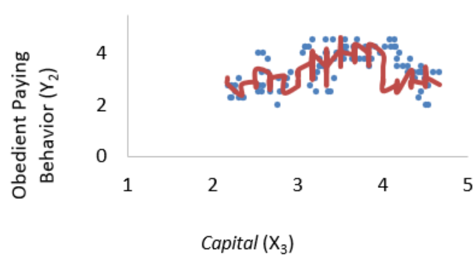


Figure 9 The plot of actual and predicted data between variables X_4 and Y_2 (see online version for colours)

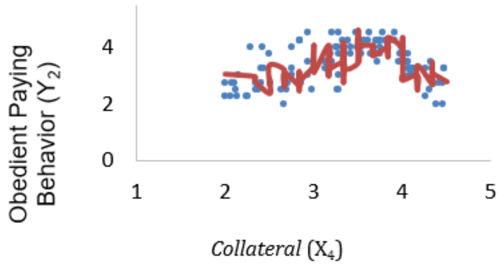


Figure 10 The plot of actual and predicted data between variables X_5 and Y_2 (see online version for colours)

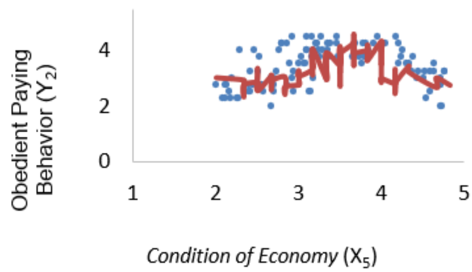


Figure 11 The plot of actual and predicted data between variables Y_1 and Y_2 (see online version for colours)

Based on the actual data plots and the predictions made for each variable in Figures 1–11, we find that using the Fourier series approach to estimate the path function does a good job of explaining the model. This is indicated by the blue dots extending above the red line or around the line of the fitted Fourier series function where the actual data were predicted, resulting in small residuals.

5 Conclusions

Function estimation in nonparametric path analysis by Fourier series approach is:

$$\hat{\alpha}(\lambda) = (n^{-1}X'X + \lambda D)^{-1} n^{-1} X' y$$

The nonparametric path function formed for each oscillation is obtained from the lambda with the optimal value. That is, when the GCV is minimal. The best model for each vibration is:

a $K = 2$ while $\lambda = 0.2$

b $K = 3$ while $\lambda = 0.9$

c $K = 4$ while $\lambda = 0.9$

d $K = 5$ while $\lambda = 0.9$

The best nonparametric path model that can explain the 5C variable of time to pay by willingness to pay is the case with an oscillation $K = 4$ in R2 of 78%. Effect of simulation on the coefficient of determination: The larger the sample size used, the larger the coefficient of determination obtained; the smaller the sample size used, the smaller the coefficient of determination. Optimal vibration and lambda levels are reached when vibration and lambda are neither too high nor too low. The best point is 1 and the best degree is when squared. A model with optimal vibration, lambda, order, and node provides a high coefficient of determination.

References

Apriana, S., Artiningsih, D.W. and Irwansyah, I. (2017) 'Analisis pengaruh prasyarat kredit (5C) terhadap kelancaran pembayaran angsuran nasabah di bank kalsel unit sentra antasari banjarmasin', *Jurnal Bisnis Dan Pembangunan*, Vol. 6, No. 1, pp.10–18.

Eubank, R.L. (1999) *Nonparametric Regression and Spline Smoothing*, CRC Press.

Fernandes, A.A.R. (2021) *Analisis Regresi Dalam Pendekatan Fleksibel: Ilustrasi Dengan Paket Program R*, Universitas Brawijaya Press, Indonesia.

Fernandes, A.A.R., Solimun, and Arisoesilaningsih, E. (2017) 'Estimation of spline function in nonparametric path analysis based on penalized weighted least square (PWLS)', *AIP Conference Proceedings*, Vol. 1913, No. 1, p.020030.

Gujarati, D.N. and Porter, D.C. (2012) *Dasar-Dasar Ekonometrika, Edisi 5*, Salemba Empat, Jakarta.

Kutner, M.H., Nachtsheim, C.J., Neter, J. and Li, W. (2005) *Applied Linear Statistical Models* (*Vol. 5*), McGraw-Hill Irwin, New York.

Nurjanah, F., Utami, T.W. and Nur, I.M. (2015) 'Model regresi nonparametrik dengan pendekatan deret fourier pada pola data curah hujan di kota semarang', *Jurnal Statistika Universitas Muhammadiyah Semarang*, Vol. 3, No. 2, pp.8–14.

Prahutama, A. (2013) *Model Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur*, Prosiding Seminar Nasional Statistika UNDIP.

Sholiha, A., Kuzairi, K. and Madianto, M.F.F. (2018) 'Estimator deret fourier dalam regresi nonparametrik dengan pembobot untuk perencanaan penjualan camilan khas madura', *Zeta-Math Journal*, Vol. 4, No. 1, pp.18–23.

Solimun, M.S. (2002) *Multivariate Analysis Structural Equation Modelling (SEM) Lisrel dan Amos*, Fakultas MIPA Universitas Brawijaya, Malang.

Solimun (2010) *Metode Partial Least Square-PLS*, CV Citra Malang, Malang.

Tripena, A. (2009) 'Pemilihan parameter penghalus pada estimator deret fourier dalam regresi nonparametrik', *Jurnal Ilmiah Matematika dan Pendidikan Matematika Universitas Jenderal Soedirman Purwokerto*, Vol. 2, pp.2–16.

Wahyuni, A.L. (2019) *Analisis Pengaruh 5C Terhadap Kelancaran Pembayaran Pembiayaan Paket Masa Depan BTPNS (Studi Kasus Pada MMS Tambun Selatan)*, Doctoral dissertation, STIE Indonesia Banking School.

Widayati, R. and Putri, D.E. (2019) *Pelaksanaan Kredit Pada Bank Perkreditan Rakyat LPN Pasar Baru Durian Sawahlunto*, https://doi.org/10.31219/osf.io/jw2za

Wisisono, I.R.N., Nurwahidah, A.I. and Andriyana, Y. (2018) 'Regresi nonparametrik dengan pendekatan deret fourier pada data debit air sungai citarum', *Jurnal Matematika* '*MANTIK'*, Vol. 4, No. 2, pp.75–82.