



# **PROCEEDINGS**

# 2<sup>nd</sup>BASIC SCIENCE International Conference Malang

February 24-25<sup>th</sup>, 2012

# 2<sup>nd</sup> Basic Science International Conference 2012 February 24 – 25<sup>th</sup>, 2012

# **Proceedings**

RESEARCH INNOVATION ON MODELING, SIMULATION, AND ITS APPLICATIONS

### **Editors**

Dr. Wuryansari Muharini Kusumawinahyu **Danny Prasetyo Hartanto** Rizka Firdausi Mutya Fani Atsomya

Mathematics Department on Behalf of Faculty of Sciences Brawijaya University















#### Message from Dean of Faculty of Sciences

On behalf of the faculty of Sciences we are very happy to welcome you at the 2<sup>nd</sup> International Conference on Basic Sciences 2012. Our appreciation to Prof. Dr. Ir.Yogi Sugito, rector of the University of Brawijaya for his supports. My gratitude to all of the keynote and invited speakers who made this conference succeed.

We can share our knowledge and experiences during this conference. Next, we would like to thank all of the steering and organizing committee for their efforts in succeeding this conference, especially to Ratno Bagus Edy Wibowo, Ph.D. as the chair of the organizing committee.

Last but not least for all of the participants, we are thankful for your cooperation, contribution and very valuable support for this event.

Thank you,

Prof. Dr. Marjono, MPhil.

#### INTERNATIONAL ADVISORY COMMITTEE

- · Prof.Dr. Ir. Yogi Sugito, Brawijaya University, Indonesia
- Prof.Dr. Marjono, M.Phil, Brawijaya University, Indonesia
- · Prof. Kiyoshi Okuda, DVM.,Dr.med.,Vet.,Ph.D., Faculty of Agriculture, Okayama University, Japan
- Prof. Don Faust, Dept. Mathematics and Computer Science, North Michigan University,
   USA

#### SCIENTIFIC COMMITTEE

- Prof. M. E. O'Neill (Sydney University Australia)
- · Dr. Edgar Paski (BCIT-Canada)
- · Dr. Francois Malherbe (Swinburne University of Technology Australia)
- · Dr. Lilibeth Coo (University of Philippine Philippine)
- · Dr. Agus Suryanto, M.Sc (Brawijaya University Indonesia)
- · Dr. Wuryansari Muharini, M.Sc (Brawijaya University Indonesia)
- Dr. Rahma Fitriani, M.Sc (Brawijaya University Indonesia)
- · Dr. M. Bernadetha (Brawijaya University Indonesia)
- Dr. Rurini Retnowati (Brawijaya University Indonesia)
- · Dr. Endang Ariesoesilaningsih (Brawijaya University Indonesia)
- · Dr. Abdur Rouf (Brawijaya University Indonesia)
- Drs. Abdul Rouf Alghofari, Ph.D. (Brawijaya University-Indonesia)

FIREMORD

WESSAGE FROM

THURSDAY TIFE

COMMENTS

KEMMOTE SPEA Advitia, D., M. V

Wathematics at Lab/Wath Indon

Aceg Purgon, University, Ass Adviser for pr Institute for Cl

Widecular Sc 5/4/C-P/ON/ZAT

Dr. Takeshi Igav penetics for end m endangered f

Dr. Agus Suryar

In ing Setiawa

TRIBUTED

MATICS

Darmaid, The 6

L.H. Wiryanto, S

Low Affinish, 1

- WER NOTCH

## **CONTENTS**

FOREWORD	ii
MESSAGE FROM DEAN OF FACULTY OF SCIENCES	iii
COMMITTEE	iv
CONTENTS	٧
*EYNOTE SPEAKERS  Adytia, D., M. Woran, Prof. E. van Groesen, Faculty of Electrical Engineering,  Mathematics and Computer Science (EEMCS), University of Twente, Netherlands -  LabMath Indonesia: Effect of a possible Anak Krakatau explosion in the Jakarta Bay	K – 1
Acep Purqon, Ph.D., Assistant Professor, Computational Bioscience, Kanazawa University, Assistant Professor, Earth Physics and Complex systems, ITB Special Actives for promotion of globalization, Kanazawa University Guest researcher, Institute for Chemical Research, Kyoto University Guest researcher, Institute for Wolecular Science, Japan: SELF-ASSEMBLY FLUCTUATION MODES AND SYNCHRONIZATION PATTERS IN BIONANOCLUSTERS	K – 6
Takeshi Igawa, Hiroshima University, Japan, A case study of conservation cenetics for endangered species in Japan: Landscape genetics and captive breeding mendangered frogs of genus Odorrana.	K – 7
INVITED SPEAKERS	
Dr. Agus Suryanto, M.Sc., Brawijaya University, Indonesia	l – 1
Dr. Ing. Setiawan P. Sakti., M.Eng., Brawijaya University, Indonesia	1-7
CONTRIBUTED PAPERS	
MATHEMATICS	
Darmajid, The Geometric Representation of Semisimple Module	M – 1
Wiryanto, Stream Near a Separation of Solid Boundary	M – 5
E Herawaty, THE BOUNDEDNESS OF SUPERPOSITION OPERATOR ON $[L_\Phi]$ $^((L))$	M – 9
Meiza, Fuzzy Parameters Estimation Method for The Cusp Catastrophe Model	M – 14
Alfiniyah, THE EFFORTS TO MINIMIZE FLOODING ALONG THE RIVER THROUGH	M – 18
* WEIR NOTCH DESIGN (CASE STUDY)	

ersity,

Lily Montarcih Limantara, Optimization of Linear Programming Model For Water	M – 21
Irrigation Area of Jatimlerek, East Java of Indonesia	
Mulia Astuti, A Special Class of Integral Hypergraphs	M – 24
Sigit Pancahayani, Backbone Chromatic Number on Harary Graph	M – 27
Endang Wahyu H, THE NET SINGLE PREMIUM IN COLLECTIVE LIFE INSURANCE BASED	M – 30
ON INSURED STATUS	
Marsudi, STABILITY ANALYSIS OF THE HIV/AIDS EPIDEMIC MODEL WITH TWO AGE	M – 33
GROUPS IN INDONESIA	
Muhammad Hamzah Syahruddin, Modeling Groundwater Flow With Boundary	M – 38
Element Method (BEM)	
STATISTICS	
Anwar Fitrianto, Classical and a Robust Method in Dealing with Outlier in Factorial	S – 1
Designs:An Empirical Example	
Maylita Hasyim, SPATIAL MIXTURE SURVIVAL MODEL BY NORMAL DISTRIBUTION OF	S <b>-</b> 5
FRAILTY FOR INCIDENCE RATE OF DENGUE HEMORRHAGIC FEVER (DHF) IN	
PAMEKASAN	
Mohammad Masjkur, MODELING PHOSPHATE SORPTION ON KAOLINITIC AND	S – 10
SMECTITIC SOILS	
anwar fitrianto, Effects of an Outlying Observation on Multiple Linear Regression	S <b>–</b> 19
Model: An Empirical Example in Forward Selection Procedures	
Idhia Sriliana, TRUNCATED SPLINE REGRESSION IN LINEAR PARTIAL MODEL FOR	S – 23
LONGITUDINAL DATA	
Lilik Hidayati, MULTIVARIABLE CUBIC SPLINE REGRESSION IN SCORE MODELING	S – 27
NATIONAL EXAM	
Andoyo Supriyantono, Estimation of (Co)Variance Components and Genetic	S - 31
Parameter of Withers Height, Chest Girth and Body Length of Bali Cattle Using	
Animal Model	

**Nur Chamidah** 

Case of Hetero

Doris Enwanto

Linear Discrimi

Ardilliansah Fa

Effect for Wood

Solimum, Inves

Validity and U

Dan Angered

divalid

W. Art Zamuli

Satelite mag

Saisio

CMP.TH S

Actorial E.T.

Source Service

FISHE + 4 A

I E I I

North File

342 4 5 1

-----

# TRUNCATED SPLINE REGRESSION IN LINEAR PARTIAL MODEL FOR LONGITUDINAL DATA

Idhia Sriliana<sup>1)</sup>, I Nyoman Budiantara<sup>2)</sup>

Department, Faculty of Mathematics and Natural Sciences, Institut of Technology Sepuluh
Nopember, Surabaya, Indonesia (atha\_muflih@yahoo.co.id)

Lecture of Statistics Department, Faculty of Mathematics and Natural Sciences, Institut of Technology Sepuluh
Nopember, Surabaya, Indonesia (i. nyoman.b@statistika.its.ac.id)

#### **ABSTRACT**

**Theoretically** and practically, Spline regression estimator in semiparametric or partially linear model is not only suitable for handling cross section data, but it can be generalized for longitudinal data. Linear partial models for longitudinal data with  $y_{ij} = \mathbf{X}_{ij}^T \mathbf{\beta}_i + f\left(t_{ij}\right) + \varepsilon_{ij}$ ,  $j = 1, 2, ..., n_i$ ,  $i = 1, 2, ..., n_i$ , there are n subjects which have  $\mathbf{i} = \mathbf{i} = \mathbf{i} = \mathbf{i}$  observation. By using the truncated spline regression approach and the weighted least squares splinization, the estimation in partially linear model for longitudinal data was  $\hat{\mathbf{Y}} = \mathbf{A}(k_1, k_2, ..., k_n)\mathbf{Y}$ , where

Lationship among the number of leukocytes in leukemia patients (y) toward the number of platelets (x) and beckup time (t). The result of model estimation showed the number of leukocytes leukemia (y) had linear pattern with the number of platelets (x) and it also had pattern which followed a linear spline two knots toward checkup (t). This model had a GCV value of 0.4118387 with MSE value of 0.0001967 and the coefficient of determination ( $R^2$ ) of 0.9879.

Sovords: Leukemia, Linear Partial Model, Longitudinal Data, Truncated Spline

#### LINTRODUCTION

Regression analysis is one of the data analysis that is used to determine the model patterns of mationships between predictor variables and response variables. If the pattern of relationships between predictor variables and response variables are unknown, then the nonparametric regression is a recession approach which appropriate for the data of an unknown form of the regression curve Spline is one of the nonparametric regressions have specific statistical interpretation and an established visual interpretation [2]. Moreover, spline to handle the smooth character data and it has m excellent ability to handle data that changes its specific sub-interval [3]. Spline and spline estimator theoretically and practically somet only suitable for handling cross section data, be generalized to longitudinal data [4]. In a of longitudinal data, in general, the mercations are made of n mutually independent where each subject is observed repeatedly = If event time periods [5].

One of the methods that can be used for straing parameters in linear partial models for strainal data is truncated spline regression approach has advantages including easier mathematically interpretations same as the parametric before. Basically, the main issue in the strain of parameters using truncated spline

regression is about the selection of optimal knot points with ordinary least squares optimization. However, estimation of parameters in linear partial models for longitudinal data is conducted using weighted least squares optimization, which includes **W** value for the completion of the smallest squares optimization. This is caused by the observations of longitudinal data are dependent on the same subjects while for different subjects are independent, so that the correlation between observations cause ordinary least squares optimization need to be refined by a value.

There are several studies which analyze longitudinal data modeling using nonparametric regression approach. They are Wu and Chiang [6] who using kernel estimators, Zhang [7] using generalized spline regression, Rice and Wu [8] using the spline approach which based on a mixed effects model to estimate the regression curve for longitudinal data. On the other hand, there are several studies on longitudinal data based on the semiparametric model in particular linear partial model, including the Fan and Zhang [9], Guo [10], Liang and Xiao [11], and Laome [12].

In this study, it used spline regression approach with weighted least square optimization to estimate parameters in partially linear model for longitudinal data. Then, it was conducted the selection of optimal knots point to find the best spline model estimator. In the last step, the proposed method was applied to predict the relationships pattern of the number of leukocytes in leukemia patients based on the number

of platelets and checkup time. The results showed truncated spline estimators in linear partial models for longitudinal data have the characteristics of linearity. The selection of optimal knot points was conducted by GCV method. Application of the model showed the pattern of the relationship among the number of leukocyte in leukemia patients, the number of platelets, and checkup time may be formed in a linear partial model, with the number of platelets as parametric component and checkup time as nonparametric component. Based on the result of model estimation, it showed that the number of leukocytes in leukemia patients had linear pattern on the number of platelets and it also had pattern which followed a linear spline two knots in checkup time.

#### II. METHODS

The method used in this study was truncated spline regression with weighted least square optimization. Spline is polynomial pieces that have continuous segmented characteristics thus it can explain the local characteristics of the data function. In general, the spline function of degree m with K knots is any function that written as:

$$s(t) = \sum_{p=0}^{m} \alpha_p t^p + \sum_{r=1}^{K} \delta_r (t - k_r)_+^m$$
 (1)

with,

$$(t-k_r)_{\cdot}^{*} = \begin{cases} (t-k_r)^{n} & , t \geq k_r \\ 0 & , t < k_r \end{cases}$$

where  $\alpha$  is polynomial parameter,  $\delta$  is truncated parameter, and  $k_1, k_2, ..., k_K$  are knot points.

In a study of longitudinal data, the observations were conducted on n subjects which were mutually independent, where each subject was observed continuously in a specified period. If it is given the longitudinal data  $(t_{ij}, x_{ij}, y_{ij})$ ,  $j = 1, 2, ..., n_i$ , i = 1, 2, ..., n, then the linear partial model for longitudinal data is obtained from [14]:

$$y_{ij} = \mathbf{X}_{ij}^{T} \mathbf{\beta}_{ij} + f(t_{ij}) + \varepsilon_{ij}, j = 1, 2, ..., n_{ij}, i = 1, 2, ..., n$$
 (2)

where  $\mathbf{\beta}_{i} = \left(\beta_{ij}, \beta_{i2}, ..., \beta_{il}\right)^{T}$  is  $l \times 1$ -dimensional vector of parametric regression coefficient  $\mathbf{X}_{ij}$ ,  $\mathbf{X}_{ij}^{T} = \left(\mathbf{X}_{ij}, \mathbf{X}_{2j}, ..., \mathbf{X}_{ij}\right)$ ,  $\mathbf{X}_{ij} = \left(x_{ij}, x_{ij2}, ..., x_{ijl}\right)$ ,  $\mathbf{X}_{2j} = \left(x_{2j1}, x_{2j2}, ..., x_{2jl}\right)$ , ...,  $\mathbf{X}_{nj} = \left(x_{nj}, x_{nj2}, ..., x_{njl}\right)$ ,  $f\left(t_{ij}\right)$  is an differentiable function and  $\mathcal{E}_{ij}$  is random error,  $j = 1, 2, ..., n_{ij}$ ,  $i = 1, 2, ..., n_{ij}$ .

The optimal curve of spline regression estimator depends on the location of the knots points  $k_1, k_2, ..., k_K$ . Therefore, it needs to choose optimal knots point to determine the best spline model. Knot points are the interface point to show the change of behavior of spline functions on different intervals.

The method used in the selection of optimal knot points is generalized cross validation (GCV), which is defined as follows [13]:

$$GCV(\mathbf{k}_{i}) = \frac{N^{4}Y'(\mathbf{I} - \mathbf{A}(\mathbf{k}_{i}))'W(\mathbf{I} - \mathbf{A}(\mathbf{k}_{i}))Y}{(N^{4}trace(\mathbf{I} - \mathbf{A}(\mathbf{k}_{i})))^{2}}$$
(3)

Estimation of linear partial models for longitudinal data with truncated spline regression approach was performed with these following steps:

- a. Approaching  $f(t_{ij})$  with truncated spline degree m and knots  $k_1, k_2, ..., k_K$
- b. Indicating the linear partial model for longitudinal data as follows:

$$y_{ij} = \mathbf{X}_{ij}^{T} \boldsymbol{\beta}_{i} + \sum_{p=0}^{m} \alpha_{ip} t_{ij}^{p} + \sum_{r=1}^{K} \delta_{ir} \left( t_{ij,r} - k_{ir} \right)_{+}^{m} + \varepsilon_{ij}$$

 $i = 1, 2, ..., n, \quad j = 1, 2, ..., n$ 

c. Writing model in step (b) as follows:

$$\mathbf{Y} = \mathbf{X}(\mathbf{k}_{i})\mathbf{B} + \boldsymbol{\varepsilon}, \ \mathbf{k}_{i} = \begin{pmatrix} k_{i1} & k_{i2} & \cdots & k_{iK} \end{pmatrix}^{T}$$

d. Determining  $\hat{\mathbf{B}}$  by using weighted least squares optimization:

$$\min_{\mathbf{B} \in \mathbb{D}^{n \times 1 \times 1 \times n \times 1}} \left\{ \left. \left( \mathbf{Y} - \mathbf{X} \left( \mathbf{k}_{i} \right) \mathbf{B} \right)^{T} \mathbf{W} \left( \mathbf{Y} - \mathbf{X} \left( \mathbf{k}_{i} \right) \mathbf{B} \right) \right. \right\}$$

e. Obtaining the spline estimation in linear partial model as follows:

model as follows:  

$$\hat{y}_{ij} = \mathbf{X}_{ij}^{T} \hat{\boldsymbol{\beta}}_{i} + \sum_{p=0}^{\infty} \hat{\alpha}_{ip} t_{ij}^{p} + \sum_{r=1}^{K} \hat{\delta}_{ir} \left( t_{ij} - k_{ir} \right)_{i}^{m}$$

Furthermore, the results of estimates were applied to observe the pattern of relationship among the number of leukemia leukocyte toward the number of platelets and the checkup time with these following steps:

- a. Creating data plot  $(t_{ij}, x_{ij}, y_{ij})$ ,
- b. Modeling the data with spline approach,
- Choosing the optimal knot points by the GCV method,
- d. Finding the smallest value of GCV,
- e. Estimating model patterns of relationships with leukemia leukocyte count,
- f. Calculating the value of determination coefficient and mean square error.

n Think there

lim title: and

#### III. RESULTS

A. Estimation Spline Truncated with Weighted Least Square Optimization in Linear Partial Models for Longitudinal Data

If the curve regression in a linear partial model for longitudinal data was approximated by truncated spline regression,

$$s(t_{ij}) = \sum_{p=0}^{m} \alpha_{ij} t_{ij}^{p} + \sum_{r=1}^{K} \delta_{ir} (t_{ij} - k_{ir})_{+}^{m}$$
 (4)

where i = 1, 2, ..., n, j = 1, 2, ..., n

the linear partial models longitudinal data can be expressed in matrix notation as follows:

$$\mathbf{Y} = \mathbf{X}(\mathbf{k}_{\perp})\mathbf{B} + \mathbf{\epsilon} \tag{5}$$

where 
$$\mathbf{k}_i = \begin{pmatrix} k_{ii} & k_{ij} & \cdots & k_{ir} \end{pmatrix}^T$$
,  $i = 1, 2, \dots, n$ 

where  $\mathbf{k}_{i} = \begin{pmatrix} k_{ii} & k_{i2} & \cdots & k_{iK} \end{pmatrix}^{T}$ ,  $i = 1, 2, \dots, n$ By using weighted **W**, estimates of **B** in (5) can be obtained by solving the optimization:

$$\min_{\mathbf{a}_{i,j} \in A(\mathbf{a}_{i,j})} \left\{ \left( \mathbf{Y} - \mathbf{X} \left( \mathbf{k}_{i,j} \right) \mathbf{B} \right)^{T} \mathbf{W} \left( \mathbf{Y} - \mathbf{X} \left( \mathbf{k}_{i,j} \right) \mathbf{B} \right) \right\}$$
 (6)

with W is a matrix:

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

so that, we obtained:

$$\hat{\mathbf{B}} = \left(\mathbf{X} \left(\bar{k}\right)^T \mathbf{W} \mathbf{X} \left(\bar{k}\right)\right)^{-1} \mathbf{X} \left(\bar{k}\right)^T \mathbf{W} \mathbf{Y} \tag{7}$$

Based on the estimation of B in (7), then obtained:

$$\hat{\mathbf{Y}} = \mathbf{X}(\mathbf{k}_{i})\hat{\mathbf{B}}$$

$$= \mathbf{A}(\mathbf{k}_{i}, \mathbf{k}_{i2}, \dots, \mathbf{k}_{iK})\mathbf{Y}, i = 1, 2, \dots, n$$
(8)

es

ial

ng he

ith

ion

ast

tels

del

ited

(4)

$$\mathbf{A}(\mathbf{k}_{a}, \mathbf{k}_{a}, ..., \mathbf{k}_{a}) = \mathbf{X}(\mathbf{k}_{a}) \left(\mathbf{X}(\mathbf{k}_{a})^{T} \mathbf{W} \mathbf{X}(\mathbf{k}_{a})\right)^{T} \mathbf{X}(\mathbf{k}_{a})^{T} \mathbf{W}.$$

#### 3. Applications to Data Number of Leukocytes in Leukemia Patients

The data used in this study was secondary data derived from studies conducted Oktiriani [14], that was the data from leukemia patients were treated at Surabaya Hajj General Hospital in 2009. It was about the development of leukocytes numbers in enkemia patients who were influenced the checkup ime and the number of platelets. Based on data from medical records of Surabaya Hajj General Hospital = 2009, there were four patients who had mutually independent leukemia and assumed only having leukemia. In each of patients was measured the number of leukocyte cell and platelet cell in a certain period to be analyzed by using a truncated spline model. The study was focused to predict the pattern of the relationship among the number of leukocytes in leukemia patients (y) toward the number of platelets (x) and time (t), where the number of platelets medically as explanatory wariables that was assumed as parametric component in partially linear model.

In the analysis process, it would be conducted mear partial model estimation simultaneously to obtained the optimal linear partial model. In addition, the criteria for the best model with a muncated spline approach were measured by the

smallest value of GCV at the point of optimum knots [15]. However, other criteria of the goodness of the model were considered through the MSE and the coefficient of determination R2. In the model estimation process, the calculation of the GCV was limited to 2 knots with polynomial degree m = 1. m=2, dan m=3, each of which is called a linear spline, quadratic spline, and cubic spline. Thus, the smallest GCV value was obtained through a combination of the number of knots and the degree of the polynomial used. Repetition which were conducted for each patient caused the correlation between observations in the same subjects. The indication of a correlation could be seen in the response variable that affected to the time variable. So the estimation of the optimal spline model was obtained from a weighting. In this case, the analysis was conducted by using weighting of variance covariance matrix.

Furthermore, the combinations of GCV with the number of knots and the degree of the polynomial spline for estimating the linear partial model are shown in Table 1.

Table 1. GCV values for The Model with Linear

Knots	The Degree of Polynomial Spline	GCV
1	1	0,6445039
1	2	3,1611685
1	3	4,0323595
2	1	0,4118387
2	2	3,0689753
2	3	4,1987952

Based on Table 1, it was obtained the smallest GCV values on a combination of two knots and one degree of the polynomial (linear spline) of 0.4118387. Knot points for each patient are shown in Table 2 below:

Table 2. Knot points with GCV Optimum Value

Subject	Knots	
	1	2
Patient 1	138	154
Patient 2	106	162
Patient 3	102	153
Patient 4	120	174

Thus, the estimation of linear partial model about the pattern of the relationship among the number of leukocytes in leukemia patients toward the number of platelets and checkup time is given:

Katha kanya

LINTRO

dewided int

[2] and the

uffecting w

mine regre

meression h

mudel which

While

EDMINISTRY IS I

Starty mising

MOUTE, IE IS CO

mady and so

$$\begin{split} \hat{y}_{ij} &= 0.24789 + 0.04287x_{ij} + 0.00065t_{ij} + 0.00237\left(t_{ij} - 138\right)_{-1}^{1} + \\ &- 0.03945\left(t_{ij} - 154\right)_{-1}^{1}, \quad j = 1, 2, ..., 7 \\ \hat{y}_{ij} &= 0.02549 + 0.00026x_{ij} - 0.00004t_{ij} + 0.00097\left(t_{ij} - 106\right)_{-1}^{1} + \\ &- 0.00079\left(t_{ij} - 162\right)_{-1}^{1}, \quad j = 1, 2, ..., 8 \\ \hat{y}_{ij} &= 0.01704 - 0.11162x_{ij} + 0.0001t_{ij} - 0.00008\left(t_{ij} - 102\right)_{-1}^{1} + \\ &- 0.0001\left(t_{ij} - 153\right)_{-1}^{1}, \quad j = 1, 2, ..., 9 \\ \hat{y}_{ij} &= 0.01503 + 0.00038x_{ij} + 0.00004t_{ij} + 0.00031\left(t_{ij} - 120\right)_{-1}^{1} + \\ &- 0.00137\left(t_{ij} - 174\right)_{-1}^{1}, \quad j = 1, 2, ..., 5 \end{split}$$

This model has a MSE value of 0.0001967 and coefficient of determination R<sup>2</sup> of 0.9879.

Based on (9), (10), (11), and (12) show that the estimation of the number of leukocytes for each leukemia patient  $(y_{ij})$  is as bellow:

- (i) It has linear pattern with the number of platelets  $(x_{ij})$ , and
- (ii) It patterns which follows a linear spline with two knots with checkup time  $(t_{\perp})$ .

Based on the model, it appears that in patient 2 and patient 4 have a tendency that the number of leukocytes keeps increasing with the length of time of checkup. It means that the longer time the patient undergoes the checkup; therefore the patient's condition keep declining instead, although in the beginning of each patient has different patterns. In contrast, patient 1 and patient 3 have a tendency that the number of leukocytes keeps decreasing along the length of time of checkup, which means that the longer the patient undergoes the checkup of the patient's condition improved. This difference may be influenced by the circumstances and background of different patients, as well as the influence of other factors that are not included in the model. While the pattern of the relationship between the number of platelets and the numbers of leukocytes, leukocyte numbers have a tendency to increase as the numbers of platelets also increase.

#### VII. CONCLUSION

If given a linear partial model for longitudinal data  $y_y = \mathbf{X}_y \mathbf{\beta}_i + f\left(t_y\right) + \varepsilon_y$ ,  $j=1,2,...,n_i$ , i=1,2,...,n then by using a truncated spline regression approach with weighted least squares (WLS) optimization, so it is obtained  $\hat{\mathbf{Y}} = \mathbf{A}\left(k_{i1},k_{i2},...,k_{ig}\right)\mathbf{Y}$ , where

$$\mathbf{A}(k_n, k_2, \dots, k_k) = \mathbf{X}(\mathbf{k}_1) (\mathbf{X}(\mathbf{k}_1)^T \mathbf{W} \mathbf{X}(\mathbf{k}_1))^T \mathbf{X}(\mathbf{k}_1)^T \mathbf{W}.$$

The results of model application show a pattern of relationship among the number of leukocytes in leukemia patients, the number of platelets, and checkup time in the same time. They may be formed in a linear partial model, with the number of platelets as parametric component and checkup time as nonparametric component. Model estimation results show that the number of leukocytes leukemia patients (y) have linear pattern toward the number of platelets (x) and pattern which follow a linear spline

two knots for checkup time (t). This model has a GCV value of 0.4118387 with MSE value of 0.0001967 and the coefficient of determination ( $\mathbb{R}^2$ ) of 0.9879.

#### ACKNOWLEDGEMENT

This paper mostly based on guidance from ICCBS II.

#### REFERENCES

- Lestari, B. & Budiantara, I.N., Spline Estimator of Triple Response Nonparametric Regression Model, Proposed Paper for Jurnal Ilmu Dasar, vol. 11, pp. 17-22, 2010.
- [2] Eubank, R.L., Spline Smoothing and Nonparametric Regression, New York, Marcel Deker, 1988.
- [3] Cox, D.D. & O'Sullivan, F., "Penalized Type Estimator for Generalized Nonparametric Regression, 1983", Journal of Multivariate Analysis, vol. 56, pp. 185-206, 1996.
- [4] Wu, H. & Zhang, J.T., Nonparametric Regression Method for Longitudinal Data Analisys: Mixed Effects Modeling Approaches, John Wiley and Sons, New York, 2006.
- [5] Diggle, P.J., Liang, K.Y., & Zeger, S.L., Analysis of Longitudinal Data, Oxford University Press, New York, 1994.
- [6] Wu, C.O. & Chiang, C.T., "Kernel Smoothing on Varying Coefficient Model with Longitudinal Dependent Variable", Statistica Sinica, vol. 10, pp. 433-456, 2000.
- [7] Zhang, H.P., "Multivariate Addaptive Spline for the Analysis of Longitudinal Data", Journal of Computational and Graphical Statistics, vol. 6, pp. 74-91, 1997.
- [8] Rice, J.A. & Wu, C.O., "Nonparametric Mixed Effects Models for Unequally Sampled Noisy Curve", *Biometrics*, vol. 57, pp. 253-259, 2001.
- [9] Fan, J. & Zhang, J.T., "Two Step Estimation on Functional Linear Models with Applications to Longitudinal Data", Journal of the Royal Statistical Sociaty, Seies B, vol. 62, pp. 303-322, 2000.
- [10] Guo, W., "Functional Mixed-Effect Models", Biometrics, vol. 58, pp. 121-128, 2002.
- [11] Liang, H. & Xiao, Y., "Penalized Spline For Longitudinal Data With An Application In AIDS Studies", Journal of Modern Applied Statistical Methods, vol. 73, pp. 13-22, 2006.
- [12] Laome, L., Regresi Semiparametrik untuk Data Longitudinal, Tesis Jurusan Statistika, ITS, Surabaya, 2008.
- [13] Wahba, G., Spline Models For Observasion Data, SIAM Pensylvania, 1990.
- [14] Oktiriani, M., Estimator Penalized Spline Dalam Model Regresi Semiparametrik Pada Data Longitudinal, Tugas Akhir Jurusan Statistika, UNAIR, Surabaya, 2010.
- [15] Islamiyati, A. & Budiantara, I.N., "Model Spline dengan Titik-Titik Knots dalam Regresi Non Parametrik", Jurnal INFERENSI, 3, 11-21, 2007

## Objective

The purpose of the conference is to contribute to the development of Modeling, Simulations and its applications and to bring together the researchers and practitioners from all over the world working in the related areas to present their research, to exchange new ideas, to discuss challenging issues, to foster future collaborations and to expose young researchers.

Organized by:







Sponsored by:



Official Site: www.basic.ub.ac.id

Mathematics Department Faculty of Sciences Brawijaya University Jl. Veteran, Malang 65145 Phone: +62-341-581033 e-mail: basicscience@ub.ac.id

ISBN 978-979-25-6033-6



Basic Science International Conference 2012 Mathematics Department on Behalf Faculty of Sciences **Brawijaya University** 

# **Certificate**

This certificate is awarded to:

Idhia Sriliana, S.Si.

in recognition of his/her contribution as a

Oral Participant

 $2^{nd}$  Basic Science International Conference (BaSIC) 2012 on February 24-25<sup>th</sup>, 2012 at Malang, East Java, Indonesia

BNI

nStat<sup>®</sup>

Ratno Bagus Edv Wibowo, Ph.D. President of BaSIC 2012, Brawijaya University Prof. Pr. Marjono, M.Phil.
Dean of Faculty of Sciences, Brawijaya University