Platelet Modeling in DHF Patients Using Local Polynomial Semiparametric Regression on Longitudinal Data

Tiani Wahyu Utami[1], Nur Chamidah[2*], Toha Saifudin[3]
[1]Doctoral Study Program, Faculty of Sciences and Technology, Airlangga University, Indonesia
[2,3]Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, Indonesia
tianiutami@unimus.ac.id[1], nur-c@fst.unair.ac.id[2], tohasaifudin@fst.unair.ac.id[3]

ABSTRACT

Regression analysis is one of the statistical methods used to model the relationship between response variables and predictor variables. Semiparametric regression is a combination of parametric and nonparametric regression. The estimator used in estimating the semiparametric regression model in this research is the Local Polynomial. Longitudinal data can be found in the health sector, including dengue hemorrhagic fever (DHF) data. The laboratory criteria for indication of DHF is thrombocytopenia. This research aims to obtain platelets model for DHF patients that can be used for forecasting so that it is hoped that it can provide information to the medical team in treating DHF patients. The estimated model used is Local Polynomial semiparametric regression on longitudinal data. The response variables in this research were platelets of DHF patients, which were influenced by hemoglobin as parametric predictor variable and examination time while hospitalized as nonparametric predictor variable. In the local polynomial regression model, it is necessary to select the optimal bandwidth and polynomial order method, GCV. The optimum bandwidth selection based on the GCV method obtained is 1.5 and polynomial order of 2, then applied to DHF patient platelet data, producing an estimated local polynomial semiparametric regression model that follows the actual data pattern. Modeling the platelets of DHF patients obtained using a local polynomial estimator resulted in an R^2 value of 84.25% and MAPE of 4.5%, indicating highly accurate forecasting, so it can be concluded that the resulting model is better at predicting.

———————————— ◆ ————————————

A. INTRODUCTION

Regression analysis is a statistical method that aims to model the relationship between the response and predictor variables. The regression analysis approach was carried out in three ways, including parametric, nonparametric, and semiparametric regression. Parametric regression is performed when the shape of the regression curve is known (Laome et al., 2020). it is assumed that the shape of the regression curve is known based on previous information and theory or experience. Therefore, the regression curve estimator is obtained by estimating its parameters. Meanwhile, nonparametric regression is performed if the shape of the regression curve is unknown (Budiantara et al., 2015). The nonparametric regression method makes no assumptions about the form of a particular curve or information about the shape of the regression curve (Ramli et al., 2020). The nonparametric regression curve, contained in a certain function space, is assumed to be smooth. In the nonparametric regression procedure,

the data will be searched for the shape of the regression curve without being influenced by the researcher's subjectivity. Semiparametric regression is a combination of parametric and non-parametric regression (Utami et al., 2019). Semiparametric regression is used if one of the regression curves is unknown while the others are known. Several semiparametric regression methods that have been developed include semiparametric kernel, spline, local polynomial regression, and others. The following is research on semiparametric models, including Chamidah et al. (2016) used a local linear estimator to estimate the parameter model. Hidayah et al. (2019) used truncated splines applied in semiparametric regression. Lestari and Chamidah (2020) used a semiparametric model to smooth the spline.

In regression, there are two types of data: cross-sectional and longitudinal. Cross-sectional data has different characteristics from longitudinal data. In cross-sectional data, the observations of *n* subjects were independent, and each subject only has an observation. In longitudinal data, the observations of *n* subjects were also independent, and each subject was observed several times (Nidhomuddin et al., 2019). According to Wu and Zhang (2006), there are several advantages to researching longitudinal data, including identifying individual changes and requiring fewer subjects due to repeated observations. In addition, the estimation is more efficient because each observation is carried out. Longitudinal data is done because it can provide information about changes from time to time or the dynamics of changes to provide an overview of data fluctuations from time to time. Research on longitudinal data includes Setyawati et al. (2021) using a spline estimator for semiparametric regression, Islamiyati et al. (2019) with penalized spline regression and Nidhomuddin et al. (2022) using a local linear estimator for confidence interval of the parameter.

Longitudinal data are often found in the health sector because of the nature of observations made repeatedly over a certain period. One example of a case is Dengue Hemorrhagic Fever. Dengue Hemorrhagic Fever (DHF) is a viral infectious disease transmitted through the bite of Aedes aegypti and Aedes albopictus mosquitoes, varying from mild to severe (Hidayah et al., 2017). DHF is a threat to global health. The incidence of dengue fever tends to increase yearly; dengue fever is currently an endemic disease with a high mortality rate. 20,000 cases of death (Syumarta et al., 2014). Based on data from the Ministry of Health of the Republic of Indonesia, in 2022, the number of DHF cases was 45,387, with 432 deaths. DHF is an acute viral infectious disease caused by the dengue virus, which is characterized by fever for 2-7 days accompanied by bleeding manifestations and decreased platelets (thrombocytopenia). Laboratory criteria for indication of DHF are thrombocytopenia ($<100,000/mm3$) (Ojha et al., 2017).

Thrombocytopenia is one of the laboratory criteria that can support the diagnosis of DHF. At the beginning of the patient's admission to the hospital, the platelet count is usually normal, then decreases according to the severity of the disease and the occurrence of bleeding (Faridah et al., 2022). In DHF patients, thrombocytopenia occurs due to the appearance of antibodies against platelets due to the formation of antigen-antibody complexes (Azeredo et al., 2015). Platelets are produced in the bone marrow by fragmentation of the megakaryocyte cytoplasm. The average platelet count is 150,000-400,000/mm3. In patients with dengue infection, a platelet count of $\leq 100,000$ is usually found between the 3rd to 7th day when suffering from DHF (Jayashree et al., 2011). Hemoglobin (Hb) is essential in diagnosing DHF, especially when plasma leakage can lead to shock. Hemoglobin in the first days is generally normal or slightly

decreased because the dengue virus is still at the replication stage then its levels will rise following the increase in hemoconcentration and are the earliest hematological abnormalities found in DHF patients (Silitonga et al., 2020). In cases of DHF, decreased Obstacles to the formation of erythrocytes in the spinal cord cause hb values.

The measurement of medical data for DHF sufferers varies significantly, so that DHF analysis needs to be considered on every day of treatment. This is due to the possibility that, at any time, changes will occur and require medical action. Therefore, it is necessary to model daily DHF regarding the patient's platelets. This research will use a local polynomial semiparametric regression approach on longitudinal data to estimate the regression model. In the local polynomial regression model, it is necessary to select the optimal bandwidth and polynomial order method, (i.e. using GCV to select optimal bandwidth) (Fitriyani et al., 2021). The type of kernel function used is the Gaussian kernel. The software used to determine the bandwidth and estimate the model is software *R*. The regression model was applied to the DHF patient's platelet data while being treated at Roemani Hospital based on the examination time and the patient's hemoglobin. The goodness criterion used is MAPE. The purpose of this research is to obtain excellent and accurate DHF patient platelet modeling in forecasting so that it is hoped that it can provide information to the medical team in treating DHF sufferers.

B. METHODS

The type of research used in this research is quantitative, which is systematic about a situation by collecting data that can be measured using statistical techniques. The data used in this research are secondary data obtained from the medical records of the Roemani Hospital, Semarang City, Central Java Province, in 2019. The following is a flowchart of the research procedures carried out, as shown in Figure 1.
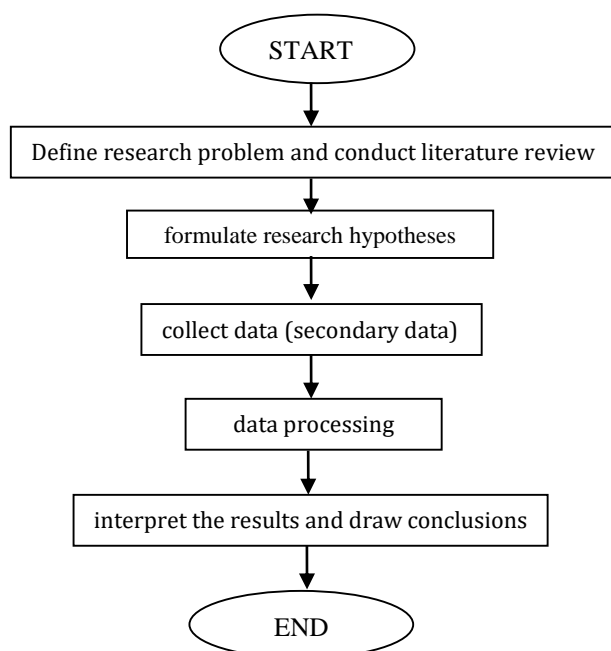

Figure 1. Flowchart of the research procedures

The data is regarding the platelet of DHF (Dengue Hemorrhagic Fever) sufferers affected by hemoglobin and examination time while hospitalized. The local polynomial semiparametric regression approach to longitudinal data in this research uses a response variable with two predictor variables, namely stating the response variable for the i th subject observation at the k th time, z_{ik} and t_{ik} declaring the predictor variables so that the relationship between z_{ik}, t_{ik} and y_{ik} following a semiparametric regression model on longitudinal data. The response variable in this research was the platelet level of dengue fever patients during hospitalization (thousands/μL), while the parametric component predictor variable was hemoglobin of DHF patients while hospitalized (gram/dL) and the nonparametric component predictor variable when examining patient platelet levels dengue fever while hospitalized (days). The number of subjects taken was 6 for Grade II DHF patients. Applying longitudinal data semiparametric polynomial local regression model to analyze the effect of hemoglobin and time of examination on platelet per day in DHF patients while being treated at Roemani Hospital, Semarang City, with the following steps:

    a. Identification of parametric and nonparametric variables

    b. Create a paired data plot that satisfies the semiparametric regression model:

$$y_{ik} = \beta_0 + z_{ik}\beta_1 + z_{ik}{}^2\beta_2 + ... + z_{ik}{}^p\beta_p + f(t_{ik}) + e_{ik} \; ; k = 1, 2, ..., n_i ; i = 1, 2, ..., n \tag{1}$$

    c. Model equation 1 can be stated as follows:

$$y_{ik}^* = f(t_{ik}) + e_{ik} \text{ atau } \boldsymbol{y^*} = \boldsymbol{f} + \boldsymbol{\varepsilon} \text{ dengan } \boldsymbol{y^*} = \boldsymbol{y} - \boldsymbol{T\beta}$$

    d. The state $f(t_{ik})$ can be approximated locally by a Taylor Series of degree p such that:

$$f(t_{ik}) = f(t) + (t_{ik} - t)f^{(1)}(t) + (t_{ik} - t)^2 \frac{f^{(2)}(t)}{2!} + ... + (t_{ik} - t)^p \frac{f^{(p)}(t)}{p!}$$

If $\alpha_r = \dfrac{f^{(r)}(t)}{r!}$ then $f(t_{ik}) = \alpha_0 + (t_{ik} - t)\alpha_1 + (t_{ik} - t)^2\alpha_2 + ... + (t_{ik} - t)^p\alpha_p$

    e. Estimating parameters by minimizing Weighted Least Square (WLS) criteria:

$$WLS = \sum_{i=1}^{n} \sum_{k=1}^{n_i} \left(y_{ik}^* - f(t_{ik})\right)^2 G_h(t_{ik} - t) \tag{2}$$

    f. Determine the type of weighting, namely, the kernel function used is Gaussian. Gaussian Kernel is $G(./h) = \frac{1}{\sqrt{2\pi}} exp\left(\frac{1}{2}(-./h^2)\right)$, where $G_h(.) = G(./h)/h$ is the Kernel function and h is bandwidth (Darnah et al., 2019).

    g. Choose the polynomial order (p) and the optimal bandwidth (h) value that minimizes Generalized cross-validation (GCV) (Aydın et al., 2016):

$$GCV(h) = \frac{N^{-1} \sum_{i=1}^{n} \sum_{k=1}^{n_i} [y_{ik} - \hat{y}_{ik}]^2}{(N^{-1} tr[I - \boldsymbol{A}(h)])^2} \tag{3}$$

h. Modeling the platelet of DHF patients using a local polynomial estimator.
i. Estimation result of local polynomial semiparametric regression models on DHF platelet.
j. Calculate the R^2 and MAPE value.

$$MAPE = \sum_{i=1}^{n} \sum_{k}^{m_i} \left| \frac{y_{ik} - \hat{y}_{ik}}{\hat{y}_{ik}} \right| \times 100\% \text{ and } R^2 = \frac{\sum_{i=1}^{n} \sum_{k=1}^{m_i} (\hat{y}_{ik} - \bar{y})^2}{\sum_{i=1}^{n} \sum_{k=1}^{m_i} (y_{ik} - \bar{y})^2} ; 0 \le R^2 \le 1$$

Based on Lewis (1982), MAPE values can be interpreted into 4 categories, namely, as shown in Table 1.

Table 1. MAPE Criteria for Model Evaluation

MAPE	Forecasting Power
< 10%	Highly accurate forecasting
10% ~ 20%	Good forecasting
20% ~50%	Reasonable forecasting
> 5%	Weak and inaccurate forecasting

The smaller the MAPE value, the smaller the estimation error. Conversely, the greater the MAPE value, the greater the estimation error. The results of a prediction method have excellent forecasting ability if the MAPE value is <10% (Setiabudi and Hermawan, 2021). R^2 is a value that shows how much the independent variable influences the dependent variable. R^2 is a number that ranges from 0 to 1, indicating the magnitude of the combination of independent variables that affect the value of the dependent variable. The R-squared value (R^2) assesses how much influence a particular independent latent variable has on the dependent latent variable. There are three grouping categories in the R^2 value: strong, moderate, and weak (Hair et al., 2011). They stated that an R^2 value of 0.75 is in the strong category, an R^2 value of 0.50 is in the moderate category, and an R^2 value of 0.25 is in the weak category (Hair et al., 2011).

C. RESULT AND DISCUSSION
1. Identification of Parametric and Nonparametric Variables

The local polynomial semiparametric regression model for longitudinal data was applied to cases of DHF sufferers at Roemani Hospital Semarang City in 2019. This research wanted to see the relationship between platelet count as a response variable to the predictor variable, namely hemoglobin and time of observation. The characteristics of Grade II DHF patients were symptoms of fever, headaches, pain behind the eyeballs, aches, and joint pain accompanied by spontaneous bleeding such as red spots on the skin, nosebleeds, bleeding, gums, vomiting blood, or black stools. Presented in Table 2 below are the characteristics of the five patients used in this research based on the patient's age and gender, consisting of 4 male patients and 1 female patient:

Table 2. Characteristics of the DHF patients

No	Patient	Age	Gander
1	Patient 1	63	Male
2	Patient 2	16	Male
3	Patient 3	4	Female
4	Patient 4	24	Male
5	Patient 5	88	Male

This research used semiparametric regression with a local polynomial approach. This research aims to model the platelet levels of DHF patients to get a picture of the changes in each patient when hospitalized. The predictors used in this research are the amount of hemoglobin levels and the time of observation. In this research, the cases used were DHF patients of Grade II. The number of patients observed was five patients, and each patient was observed for six days while being treated at the hospital. Semiparametric regression has a curve with a known pattern approximated by parametric and an unknown pattern approximated by non-parametric. The parametric and non-parametric variables have to be identified through the scatterplot. The scatterplot is made between response and predictor variables. The relationship pattern formed between the platelet response variable and the predictor variable of observation time and hemoglobin for DHF patients of Grade II can be seen from the following scatterplot, as shown in Figure 2 and Figure 3.
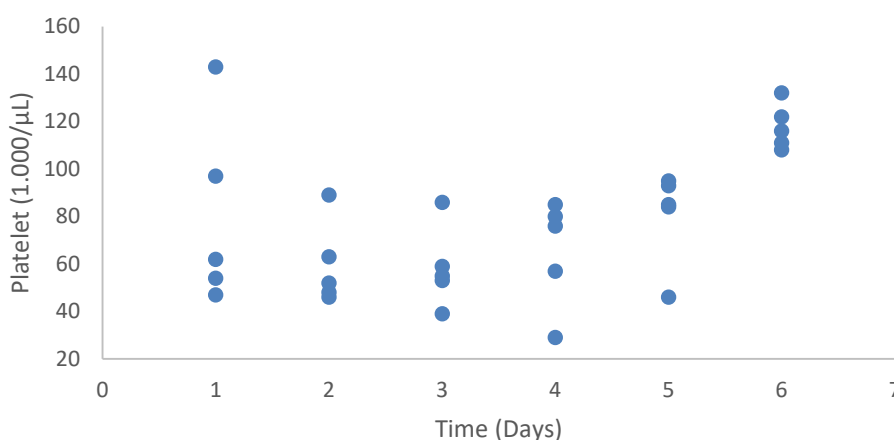

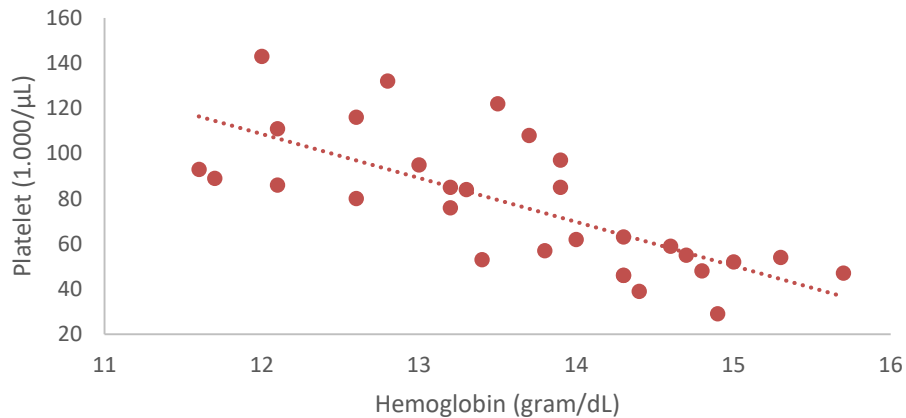**Figure 2.** Scatterplot between the platelet and observation time

Figure 3. Scatterplot between the platelet and hemoglobin

Based on the results of the scatterplot, Figure 2 shows that the response variable has an unknown pattern curve, so it can be concluded that a non-parametric approach approximates the predictor variable of observation time. Figure 3 shows the response variable has a known pattern curve, so it can be supposed that a parametric approach is used in the hemoglobin predictor variable.

2. Semiparametric Regression Model Polynomial Local Longitudinal Data

a. Optimum Bandwidth Selection and Polynomial Order

Given the observation data y_{ik}, z_{ik}, t_{ik}, the local semiparametric regression polynomial model for longitudinal data is as follows, which contains as many as n subjects with the ith subject as many as k time of observations, where y_{ik} is the response variable for the ith subject and the kth time of observation, while z_{ik} and t_{ik} are predictor variables. The relationship between the response variable (y_{ik}) and the predictor variable (z_{ik}) has a known functional form, namely linear, which is a parametric component, while the relationship between the response variable (y_{ik}) and the predictor variable (t_{ik}) has a non-functional form. Note that this is a nonparametric component. These relationships can be modeled with semiparametric regression, which can be seen in the following equation:

$$y_{ik} = \beta_0 + z_{ik}\beta_1 + f(t_{ik}) + e_{ik}, k = 1, 2, ..., n_i; i = 1, 2, ..., n \tag{4}$$

Equation (4) can also be expressed in matrix form so that it becomes:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\beta} + \mathbf{f} + \boldsymbol{\varepsilon}$$

$$y_{ik}^* = f(t_{ik}) + e_{ik} \text{ then } \mathbf{y}^* = \mathbf{f} + \boldsymbol{\varepsilon}; \ \mathbf{y}^* = \mathbf{y} - \mathbf{Z}\boldsymbol{\beta}$$

Next, denoting $f(t_{ik})$ is approximated by a local polynomial estimator of degree p so that it becomes.

$$f(t_{ik}) = \alpha_0 + (t_{ik} - t)\alpha_1 + (t_{ik} - t)^2\alpha_2 + \ldots + (t_{ik} - t)^p\alpha_p, \text{ where } \alpha_r = \frac{f^{(r)}(t)}{r!}.$$

The parameter estimation used is a method that minimizes error, namely the WLS method. The Local Polynomial Estimator for $\hat{f}$ is obtained using the WLS method so that it is obtained:

$$\hat{f} = T_{t_0}(T_{t_0}{}^T K_h T_{t_0})^{-1} T_{t_0}{}^T G_h \, y*$$

with $z_{ij} = [1, (t_{ik} - t_0), (t_{ik} - t_0)^2, \ldots, (t_{ik} - t_0)^p]^T$, $T_i = [t_{i1}, t_{i2}, \ldots, t_{in_i}]^T$,

$$T_{t_0} = [T_1^T, \ldots, T_n^T], G_h = diag(G_{1h}, \ldots, G_{nh}), G(./h) = \frac{1}{\sqrt{2\pi}} exp\left(\frac{1}{2}(-./h^2)\right)$$

The estimator *f* can be expressed as follows:

$$\hat{f} = A \, y* = A(\, y - Z\beta) \text{ dengan } A = T_{t_0}\left(T_{t_0}{}^T G_h T_{t_0}\right)^{-1} T_{t_0}{}^T G_h$$

$$\hat{\beta} = [Z^T(I - A)^T(I - A)Z]^{-1}Z^T(I - A)^T(I - A)y^{\cdot}$$

with, $A_{par} = Z[Z^T(I - A)^T(I - A)Z]^{-1}Z^T(I - A)^T(I - A).$

$$A_{nonpar} = A(I - A_{par}).$$

$$A^*(h) = A_{par} + A(I - A_{par}) \text{ (Wu dan Zhang, 2006)}.$$

One of the most important things in making a regression model using a local polynomial estimator is choosing the optimum bandwidth and polynomial order to obtain a local polynomial semiparametric regression model on longitudinal data as follows:

$$\hat{y}_{ik} = \hat{\beta}_0 + z_{ik}\hat{\beta}_1 + \hat{\alpha}_0 + (t_{ik} - t)\hat{\alpha}_1 + (t_{ik} - t)^2\hat{\alpha}_2 + \ldots + (t_{ik} - t)^p\hat{\alpha}_p$$

This research uses the GCV method to select the optimum bandwidth and polynomial order. Before estimating the model parameters, determine the optimal bandwidth and polynomial order. The degree of the local polynomial used to estimate the model is $p = 1$, $p = 2$, $p = 3$ and $p = 4$. In the following, Table 3 is presented to determine the optimum bandwidth and polynomial order based on the GCV method as shown in Table 3.

Table 3. Bandwidth and polynomial order using GCV

Bandwidth	Polynomial Order	GCV
3.2	1	581.643
1.5	2	275.065
2.1	3	292.358
2.5	4	298.001

The selection of optimal bandwidth and polynomial order are based on the smallest GCV value. Table 3 obtained a minimum GCV value of 275.065, so the optimal bandwidth value of 1.5 and polynomial order of 2 can be determined. The bandwidth results are

used as a guideline in deciding model parameter estimates so that the modeling of DHF patient platelet levels is obtained.

b. Modeling the platelet of DHF patients using the local polynomial estimator

After obtaining the optimal bandwidth value and polynomial order, the next step is to estimate the parameters using a local polynomial estimator. Based on the estimation of the model parameters, the platelet model equation for DHF patients can be formed. The following is the platelet modeling equation for DHF patients of Grade II in the 1st patient at the 3rd observation time, presented as follows:

$$\widehat{y}_{13} = \widehat{\beta}_0 + z_{13}\widehat{\beta}_1 + \widehat{\alpha}_0 + (t_{13} - 3)\widehat{\alpha}_1 + (t_{13} - 3)^2 \widehat{\alpha}_2$$
$$\widehat{y}_{13} = 6.11 \times 10^{-16} - z_{13}1.45 + 258.29 - (t_{13} - 3)1.95 + (t_{13} - 3)^2 5.79 \qquad (5)$$

The following is the platelet modeling equation for DHF patients of Grade II in the 4th patient at the 1st day observation time, presented as follows:

$$\widehat{y}_{41} = \widehat{\beta}_0 + z_{41}\widehat{\beta}_1 + \widehat{\alpha}_0 + (t_{41} - 4)\widehat{\alpha}_1 + (t_{41} - 4)^2 \widehat{\alpha}_2$$
$$\widehat{y}_{41} = 9.30 \times 10^{-16} - z_{41}1.48 + 289.58 - (t_{41} - 4)27.42 + (t_{41} - 4)^2 6.49 \qquad (6)$$

Based on equation 5, in the 1st DHF patient during the 3rd day of observation, while being hospitalized, an increase in hemoglobin of 1 gram/dL will reduce platelets by 1.450/μL, assuming the variable of observation time is in constant or fixed model. Based on equation 6, in the 4th DHF patient during the 1st day of observation while hospitalized, if there is an increase in hemoglobin of 1 gram/dL, the platelets of the DHF patient will decrease by 1.480/μL assuming the variable of observation time is in constant or fixed model.

3. Estimation Result of Local Polynomial Semiparametric Regression Models on DHF Platelet

Based on the GCV method, the optimal bandwidth obtained is $h = 1.5$ and polynomial order of 2, so the estimated value of the platelets of DHF patients can be generated. In the following, a graph will be presented of the estimation results of DHF patient thrombocytes using local polynomial semiparametric regression model, then compared with the actual platelet data of DHF patients. Estimation results are presented in green lines, and actual data are presented in blue dots as shown in Figure 4.
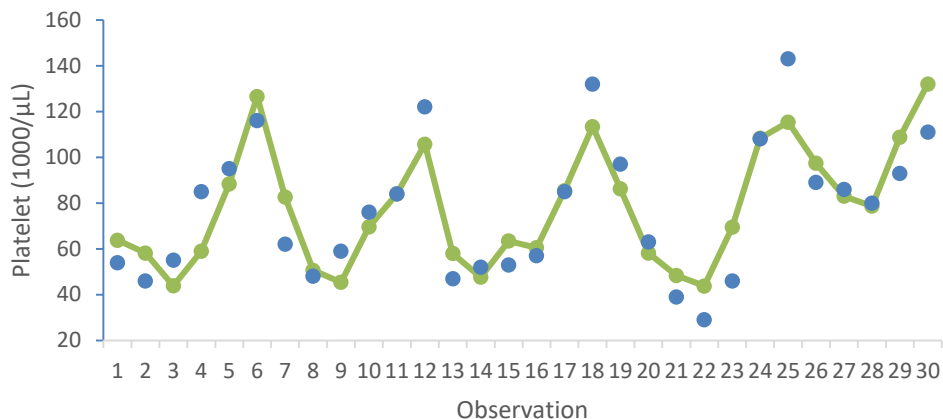
Figure 4. Compare actual data and estimation result

The estimation results of the local polynomial semiparametric regression model showed that the lowest platelets were experienced by the 1st and 4th patients with DHF, namely 44.000/μL. The decrease in platelets was experienced by 1st patient on the 3rd day while being treated at the hospital and also shared by the 4th patient on the 4th day while being treated at the hospital. Information on the occurrence of a decrease in platelets is important for the medical team to treat DHF patients. The 1st patient has the characteristics of a 63-year-old male, while the 4th patient has the characteristics of a 24-year-old male. Based on the graph presented in Figure 4, it can be seen that the platelets from the estimation results of the local polynomial semiparametric regression model follow the pattern of the actual data. The local polynomial semiparametric regression model's estimation results can be declared the best model if the resulting R^2 is close to 100%. The coefficient of determination obtained from the estimation of the local polynomial semiparametric regression model on the platelet data of DHF patients is 84.25% in the strong category. This indicates that all predictor variables (hemoglobin and time of observation) simultaneously affect 84.25% of the platelets of DHF patients. At the same time, the remaining 15.75% is influenced by other variables not tested in the research. The smaller the MAPE, the more accurate a model is in forecasting. The accuracy of the estimation results of the local polynomial semiparametric regression model, as seen from the MAPE value of 4.5%, indicates the ability to forecast highly accurately. Models with a MAPE value of less than 10% are better forecasting models. Therefore, it can be concluded that modeling platelet levels using local polynomial semiparametric regression is a better forecasting model. Previous research carried out by Side et al. (2020) used a different estimator, namely Spline, which produced a coefficient of determination value of 75.3%, which means that the estimator approach used in this research was local polynomial which made the coefficient of determination value greater than previous research, namely equal to 84.25%.

D. CONCLUSION AND SUGGESTIONS

Application of local polynomial semiparametric regression model to longitudinal data in cases of dengue hemorrhagic fever (DHF) with the response variable used is platelets and the predictor variables are hemoglobin and time of observation. Based on the results of the scatterplot, it was obtained that the predictor variable of observation time was approached

parametrically, while the hemoglobin predictor variable was approached non-parametrically. Optimum bandwidth selection based on the GCV method obtained $h = 1.5$ and polynomial order $p = 2$, then applied to DHF patient platelet data, which produces an estimation of local polynomial semiparametric regression model that follows the actual data pattern. The estimation results of the local polynomial semiparametric regression model showed that the lowest platelets were experienced by the 1st and 4th patients with DHF, namely 44.000/μL. The decrease in platelets was experienced by 1st patient on the 3rd day while being treated at the hospital and also shared by the 4th patient on the 4th day while being treated at the hospital. The 1st patient has the characteristics of a 63-year-old male, while the 4th patient has the characteristics of 24-year-old male. Modeling the platelets of DHF patients obtained using a local polynomial estimator resulted in R^2 value of 84.25% in the and MAPE of 4.5% indicating highly accurate forecasting, so it can be concluded that the resulting model is better at forecasting.

This research is only limited to cases of DHF patients with Grade II, so it needs to be studied further by adding the number of patients other than Grade II. DHF disease does not only take into account platelets, hemoglobin, and observation time but also considers hematocrit, leukocytes and body temperature of DHF patients. In addition, further research can be developed by including the variables of hematocrit, leukocytes and body temperature of DHF patients.

REFERENCES

Alzeredo, E. L., Monteiro, R. Q., Pinto, L. M. O. (2015). Thrombocytopenia in Dengue: Interrelationship between Virus and the Imbalance between Coagulation and Fibrinolysis and Inflammatory Mediators, *Mediators of Inflammation*, vol. 2015, Article ID 313842, 16 pages. doi: https://doi.org/10.1155/2015/313842.

Aydın, D., Günerib, Ö.İ., Fitc, A. (2016). Choice of Bandwidth for Nonparametric Regression Models using Kernel Smoothing: A Simulation Study, *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, Volume 26, No 1, pp 47-61. https://gssrr.org/index.php/JournalOfBasicAndApplied/article/view/5386/2773.

Budiantara, I.N., Ratnasari, V., Ratna, M. and Zain, I. (2015). The combination of spline and kernel estimator for nonparametric regression and its properties. *Applied Mathematical Sciences*. 9(122) pp. 6083-6094. DOI: https://doi.org/10.12988/ams.2015.58517.

Chamidah, N. and Rifada, M. (2016). Local linear estimator in bi-response semiparametric regression model for estimating median growth charts of children Far East. *Journal of Mathematical Sciences*. Vol. 99 (8), pp. 1233-1244. http://dx.doi.org/10.17654/MS100091433.

Darnah, M I Utoyo, Chamidah, N. (2019). Modeling of Maternal Mortality and Infant Mortality Cases in East Kalimantan using Poisson Regression Approach Based on Local Linear Estimator, *IOP Conference Series: Earth and Environmental Science*, 243 012023. https://iopscience.iop.org/article/10.1088/1755-1315/243/1/012023/pdf.

Faridah, I.N., Dania, H., Chen, Y.-H., Supadmi, W., Purwanto, B.D., Heriyanto, M.J., Aufa, M.A., Chang, W.-C. (2022). Perwitasari, D.A. Dynamic Changes of Platelet and Factors Related Dengue Haemorrhagic Fever: A Retrospective Study in Indonesian. *Diagnostics*, 12(4), 950. DOI: https://doi.org/10.3390/diagnostics12040950.

Fitriyani, N., Hadijati, M., Harsyiah, L., Sauri, M.S. (2021). Mixed Truncated Spline and Kernel Nonparametric Regression Model on Population Growth Rate in West Nusa Tenggara Province,

IOP Conference Series: Materials Science and Engineering, 1115 012054. DOI: https://doi.org/10.1088/1757-899X/1115/1/012054.

Hair, Jr., Joseph F., et. al. (2011). *Multivariate Data Analysis. Fifth Edition*. New Jersey: PrenticeHall, Inc.

Hidayah, N., Iskandar, Abidin, Z. (2017). Prevention of Dengue Hemorrhagic Fever (DHF) Associated with the Aedes aegypti Larvae Presence based on the Type of Water Source, *THE JOURNAL OF TROPICAL LIFE SCIENCE*, Vol. 7, No. 2, pp. 115 -120. DOI: https://doi.org/10.11594/jtls.07.02.05.

Hidayati L, Chamidah N. and Budiantara I. N. (2019). Spline Truncated Estimator in Multiresponse Semiparametric Regression Model for Computer based National Exam in West Nusa Tenggara. *IOP Conf. Series: Materials Science and Engineering.* 546 052029. doi: https://doi.org/10.1088/1757-899X/546/5/052029.

Islamiayati, 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)*, 3(3), 8-11 ISSN: 2000-005X. http://ijeais.org/wp-content/uploads/2019/03/IJAAR190302.pdf.

Jayashree, K., Manasa, G.C., Pallavi, P. et al. (2011). Evaluation of Platelets as Predictive Parameters in Dengue Fever. *Indian Journal of Hematology and Blood Transfusion*, vol. 27, pp 127–130. https://doi.org/10.1007/s12288-011-0075-1.

Laome, L., Budiantara, I.N., Ratnasari, V. (2023). Estimation Curve of Mixed Spline Truncated and Fourier Series Estimator for Geographically Weighted Nonparametric Regression, *Mathematics*, 11, 152. https://doi.org/10.3390/math11010152.

Lestari, B. and Chamidah, N. (2020), Estimating Regression Function of Multi-Response Semiparametric Regression Model Using Smoothing Spline, *Journal of Southwest Jiaotong University*, 55(5), pp 1-9. http://www.jsju.org/index.php/journal/article/view/737/731.

Lewis, C.D. (1982). *International and Business Forecasting Methods*. Butterworths, London.

Nidhomuddin, Chamidah, N., Kurniawan, A. (2019). Admission Test Modelling of State Islamic College in Indonesia Using Local Linear for Bivariate Longitudinal Data, *IOP Conference Series: Materials Science and Engineering*, 546 052047, doi: https://doi.org/10.1088/1757-899X/546/5/052047.

Nidhomuddin, Chamidah, N., Kurniawan, A. (2022). Confidence Interval of The Parameter on Multipredictor Biresponse Longitudinal Data Analysis Using Local Linear Estimator for Modeling of Case Increase and Case Fatality Rates Covid-19 In Indonesia, *A Theoretical Discussion. Commun. Math. Biol. Neurosci.* (23), pp.1-12. ISSN: 2052-2541. doi: https://doi.org/10.1088/1742-6596/1764/1/012076.

Ojha, A., Nandi, D., Batra, H., Singhal, R., Annarapu, G.K., Bhattacharyya, S., Seth, T., Dar, L., Medigeshi, G.R., Vrati, S., et al. (2017). Platelet activation determines the severity of thrombocytopenia in dengue infection. *Scientific Reports*, volume 7, Article number: 41697. https://doi.org/10.1038/srep41697.

Prawanti, D. D., Budiantara, I. N., and Purnomo, J. D. T. (2019). Parameter Interval Estimation of Semiparametric Spline Truncated Regression Model for Longitudinal Data, *Journal IOP Conference Series: Materials Science and Engineering*, Volume 546, No 052053, 9th Annual Basic Science International Confrence IOP Publishing. DOI: https://doi.org/10.1088/1757-899X/546/5/052053.

Ramli, M., Ratnasari, V., Budiantara, I. N. (2020). Estimation of Matrix VarianceCovariance on Nonparametric Regression Spline Truncated for Longitudinal Data, *Journal of Physics: Conference Series*, 1562 012014. DOI: https://doi.org/10.1088/1742-6596/1562/1/012014

Setiabudi, I., Hermawan, I M A., (2021). Planning for alternative road development connecting East Balikpapan - North Balikpapan, *IOP Conference Series: Earth and Environmental Science*, 879 012017. DOI: https://doi.org/10.1088/1755-1315/879/1/012017

Setyawati, M., Chamidah, N. and Kurniawan, A. (2021). Modelling Scholastic Aptitude Test of State Islamic Colleges in Indonesia using Least Square Spline Estimator in Longitudinal Semiparametric Regression, *Journal of Physics Conference Series*, 1764 012077. doi: https://doi.org/10.1088/1742-6596/1764/1/012077

Side, S., Sanusi, W., Maksum, M W. (2020). Model Regresi Semiparametrik Spline untuk D ata Longitudinal pada Kasus Demam Berdarah Dengue di Kota Makassar. *Journal of Mathematics, Computations, and Statistics*, Vol 3 No 1. https://doi.org/10.35580/jmathcos.v3i1.19181

Silitonga, P., Dewi, B. E., Bustamam, A. et al. (2020). Correlation between laboratory characteristics and clinical degree of dengue as an initial stage in a development of machine learning predictor program, *AIP Conference Proceedings*, 2264, 030008. https://doi.org/10.1063/5.0023932

Syumarta, Y., Hanif, A.M., Rustam, E. (2014). Hubungan Jumlah Trombosit, Hematokrit dan Hemoglobin dengan Derajat Klinik Demam Berdarah Dengue pada Pasien Dewasa di RSUP. M. Djamil Padang. *Jurnal Kesehatan Andalas*. Vol 3 No 3. DOI: https://doi.org/10.25077/jka.v3i3.187

Utami, T.W., Prahutama, A., Karim, A., Achmad A R. F. (2019). Modelling rice production in Central Java using Semiparametric Regression of Local Polynomial Kernel. *Journal of Physics: Conference Series,* Volume 1217, The 8th International Seminar on New Paradigm and Innovation on Natural Science and Its Application 26 September 2018, Central Java, Indonesia. DOI: https://doi.org/10.1088/1742-6596/1217/1/012108

Wu, H., and Zhang, J.T. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis*. John Wiley & Sons, Inc., Canada. ISBN:9780470009673, DOI: https://doi.org/10.1002/0470009675