BAREKENG: Journal of Mathematics and Its Applications
June 2024 Volume 18 Issue 2 Page 1249–1260
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss2pp1249-1260

MODELING CHRONIC FILARIASIS CASES IN WEST JAVA USING A MULTIVARIATE ADAPTIVE APPROACH REGRESSION SPLINES

Ardi Kurniawan[1*], Mochammad Firmansyah[2], Toha Saifudin[3]

[1,2,3] *Department of Statistics, Faculty of Science and Technology, Universitas Airlangga*
Jln. Dr. Ir H.Soekarno, Muyorejo, Surabaya, 60115, Indonesia

*Corresponding author's e-mail: * **ardi-k@fst.unair.ac.id***

Article History:

Received: 10th January 2024
Revised: 20th February 2024
Accepted: 22nd April 2024

Keywords:

Modeling;
Chronic Number of Filariasis;
Multivariate Adaptive Regression Spline.

ABSTRACT

One of the most crippling infectious diseases in the world is filariasis. Indonesia is a unitary country with 34 provinces, where West Java is one of the 5 provinces with the most filariasis sufferers in Indonesia as of 2021. Reinfection occurs in places that have implemented POMP. Therefore, monitoring operations must be carried out to track the emergence of new cases and risk factors for transmission. The aim of this research focuses on describing and modeling the number of chronic filariasis in West Java, as well as interpreting the best model results obtained. The method used is a method with a nonparametric regression approach, namely Multivariate Adaptive Regression Spline. The results of the research show that the best model obtained is a combination of 15 base functions, maximum interaction 2, and minimum observation between knots 1. From this model, the predictor variable that has the most influence on the response variable in order based on the level of variable importance is the Percentage of Population Access to Facilities Decent Sanitation, Percentage of Households with Clean and Healthy Behavior (PHBS), Sex Ratio, and Percentage of Poor Population. The interpretation of the best model is that the variable Percentage of Population Access to Adequate Sanitation Facilities above 6,650% will contribute to a reduction in the number of chronic filariasis; the Sex Ratio variable below 103,300 will contribute in the form of a reduction in the number of chronic filariasis. it can be seen that the predictor variable that has the most influence on the response variable is the variable Percentage of Population Access to Proper Sanitation Facilities (X_2) with an importance level of 100%.

How to cite this article:

A. Kurniawan, M. Firmansyah and T. Saifudin., "MODELING CHRONIC FILARIASIS CASES IN WEST JAVA USING A MULTIVARIATE ADAPTIVE APPROACH REGRESSION SPLINES," *BAREKENG: J. Math. & App.,* vol. 18, iss. 2, pp. 1249-1260, June, 2024.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

One of the most crippling infectious diseases in the world is filariasis. The worms that cause filariasis, also known as elephantiasis, invade the lymph nodes and ducts to produce infection. Many mosquito species spread this disease. This can impact the productivity of those who suffer at work, burden families and result in significant economic losses for the nation [1]. Apart from that, filariasis has long been of global concern because this disease is in third place on the list of diseases in the world that are most commonly transmitted by mosquitoes after malaria and dengue fever [2]. In 2019, the World Health Organization (WHO) stated that there were 859 million people in 50 endemic countries who were at risk of contracting filariasis. The majority of filariasis sufferers have chronic symptoms with an estimated 36 million people who have not recovered [3].

Indonesia is a unitary country with 34 provinces, where West Java is one of the 5 provinces with the most filariasis sufferers in Indonesia until 2021. Further details of the 5 provinces, namely Papua with 3629 cases, East Nusa Tenggara 1307 cases, West Papua 620 cases, West Java 588 cases, and Aceh 523 cases [4]. Apart from that, according to the West Java Provincial Health Service report (2019), of the 25 districts/cities, 11 of them are endemic areas and are spread across 266 villages in 147 sub-districts and have a total of 428 people with chronic cases and positive Microfilariae [5].

Various studies and analyzes have been carried out to analyze influential factors and equation models that can help the government as material for deeper studies regarding chronic cases of filariasis. The research on filariasis cases above indicates that there are many factors that can influence filariasis cases, however, most previous researchers only discussed the influence of predictor variables on response variables [6]. Therefore, this research is intended to see whether or not there is an influence of predictor variables, either simultaneously or individually, where the predictor variables use a combination of predictor variables from previous research which have a significant effect on filariasis cases and see whether there is an interaction between predictor variables using statistical methods, namely MARS [7]. The predictor variables used in this research are the percentage of households with Clean and Healthy Living Behavior the percentage of poor people, population access to proper sanitation facilities, and the sex ratio.

Research on modeling filariasis cases was conducted by Sri & Ernawati (2022) using the Geographically Weighted Negative Binomial Regression (GWNBR) method, stating that the percentage of households with Clean and Healthy Living Behavior (PHBS) has a significant influence on the factors that cause filariasis cases. In this way, the selection of the MARS method for modeling filariasis cases in West Java Province is expected to obtain the best model results by taking into account factors that are thought to influence filariasis cases. Apart from that, it is hoped that the research results obtained later will be useful as material for government evaluation and consideration for taking concrete steps to reduce the prevalence of chronic cases of filariasis and support the Sustainable Development Goals (SDGs) action plan issued by BAPPENAS in 2020 with a focus on the number third, namely ensuring healthy lives and promoting prosperity for all people at all ages [8].

2. RESEARCH METHODS

2.1 Method and Types of Research

This research uses a regression method with a non-parametric approach, namely the Multivariate Adaptive Regression Spline (MARS) method. The type of research used in this study is quantitative research. Quantitative research methods are one type of research whose specifications are systematic, planned and clearly structured from the beginning to the making of the research design. Quantitative research is a type of research that produces findings obtained using statistical procedures or other means of quantification (measurement) [9]. Quantitative research methods are used to examine certain populations or samples, sampling techniques are generally carried out randomly, data collection using research instruments, quantitative or statistical data analysis with the aim of testing predetermined hypotheses [10].

2.2 Research Variables

The data used in this research is secondary data taken from the official website of the Central Bureau of Statistics of the Republic of Indonesia and the Profile of the West Java Provincial Health Service in 2019.

The data used is the latest including data on the percentage of chronic filariasis cases, the percentage of districts/cities that have clean and healthy living behavior policies, the percentage of poor people, the percentage of population access to proper sanitation facilities, and the Sex Ratio in 27 districts/cities in West Java in in 2019.Research variables are everything determined by the researcher in the form of attributes or objects that have certain variations to be studied and then conclusions drawn [11]. Research variables are divided into predictor variables and response variables. Predictor variables are variables that are thought to cause influence on the response variable. Meanwhile, the response variable is a variable that appears as a result or impact of the influence caused by the predictor variable [12]. The research variables used in this study are presented in Table 1.

Table 1. Research Variable

Variabel		Information	Scale
Response	Y	Number of Chronic Filariasis Cases	Ratio
Predictor	X_1	Percentage of Households with Clean and Healthy Living Behavior (PHBS)	Ratio
	X_2	Percentage of Poor Population	Ratio
	X_3	Percentage of Population Access to Adequate Sanitation Facilities	Ratio
	X_4	Sex Ratio	Ratio

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

Descriptive statistics is the initial stage in data exploration which. Presentation of research data is in the form of bar charts assisted by Microsoft Excel software and scatterplots with the help of applications. The bar diagram between the response variables, namely the Number of Chronic Filariasis Cases with each predictor variable, namely the Percentage of Households with Clean and Healthy Behavior, the Percentage of Poor Population, the Percentage of Population Access to Proper Sanitation Facilities, and the Type Ratio The genders are presented as follows [13].

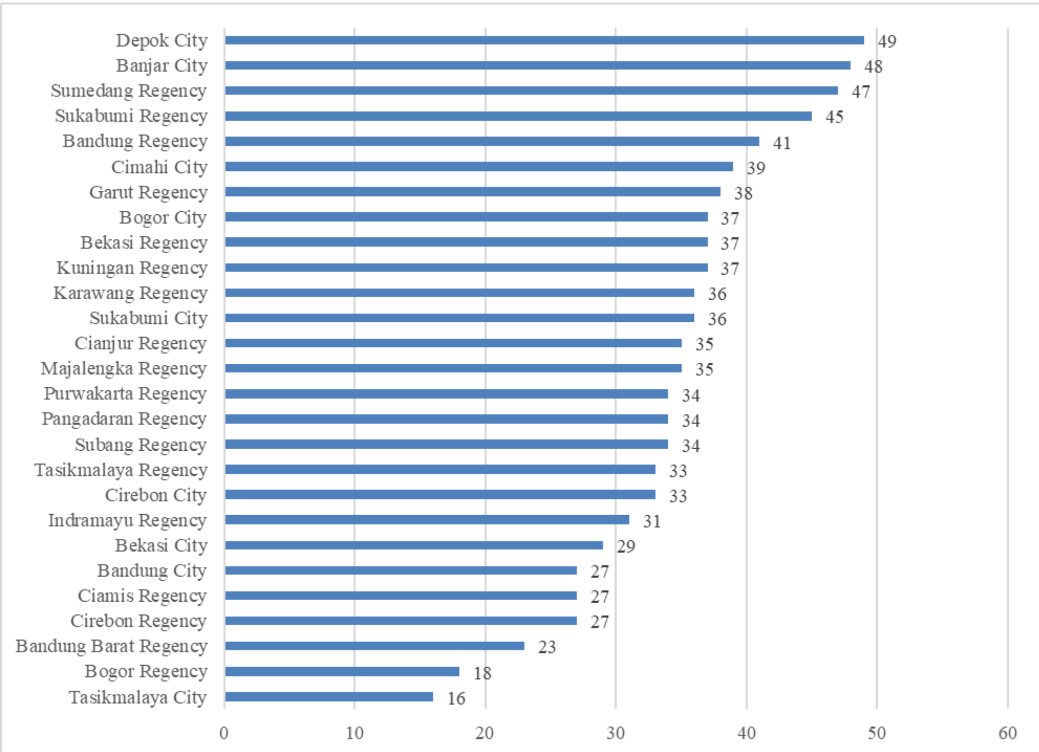

Figure 1. Variable of Bar Chart Y

Based on Figure 1, it can be seen that the highest number of chronic filariasis in West Java is in Cianjur Regency with 49 cases. Meanwhile, the lowest number of chronic filariasis in West Java was in Tasikmalaya

Regency with 16 cases. This is in accordance with the facts on the ground that the city of Depok, with its busy conditions and large population, is in line with the number of chronic filariasis cases.

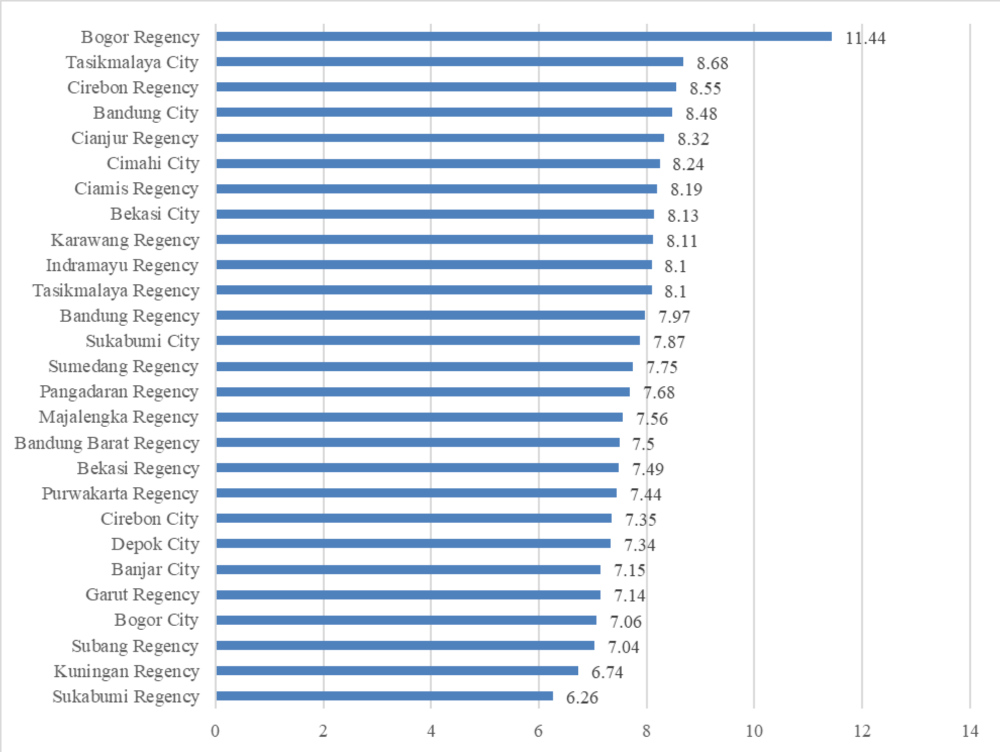

Figure 2. Bar chart of variable X_1

Based on **Figure 2**, It can be seen that the highest percentage of households with clean and healthy behavior in West Java is in Banjar City, namely 11.44%. Meanwhile, the lowest percentage of households with clean and healthy behavior in West Java is in Tasikmalaya City, namely 6.26%. This is in accordance with the fact that the city of Bogor is a city that can implement PBHS and shows the good performance of the city government.

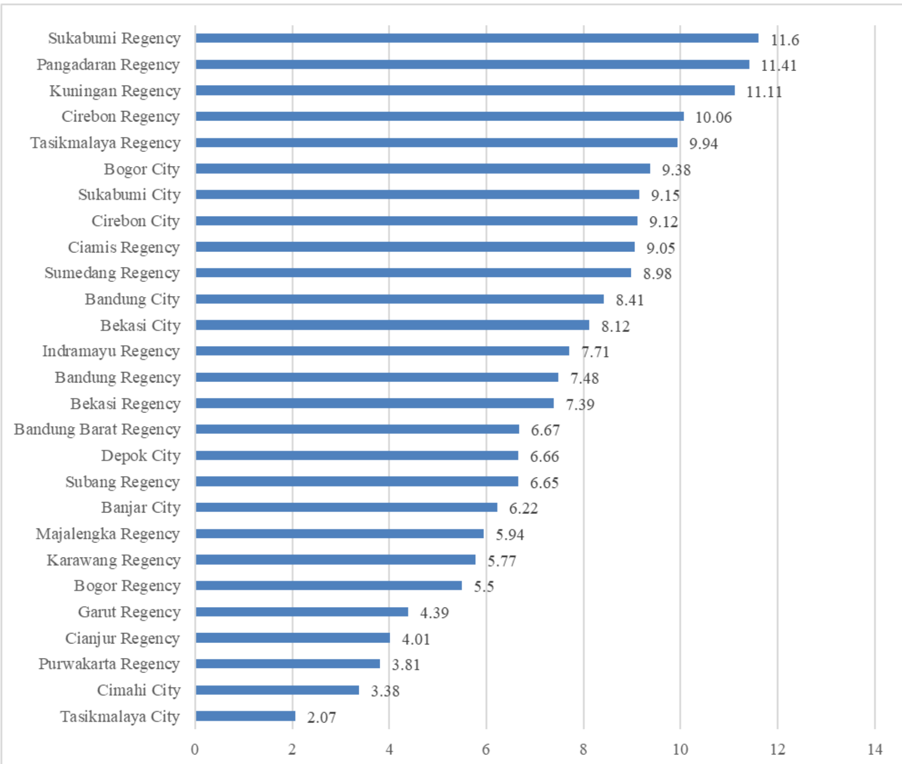

Figure 3. Bar Chart of Variable X_2

Based on Figure 3, It can be seen that the highest percentage of poor people in West Java is in Tasikmalaya City, namely 11.6%. Meanwhile, the lowest percentage of poor people in West Java is in Depok City, namely 2.07%. This is supported by central statistical data showing that Sukabumi has the highest percentage of poor people, thus showing the potential for a greater number of chronic cases of filariasis.

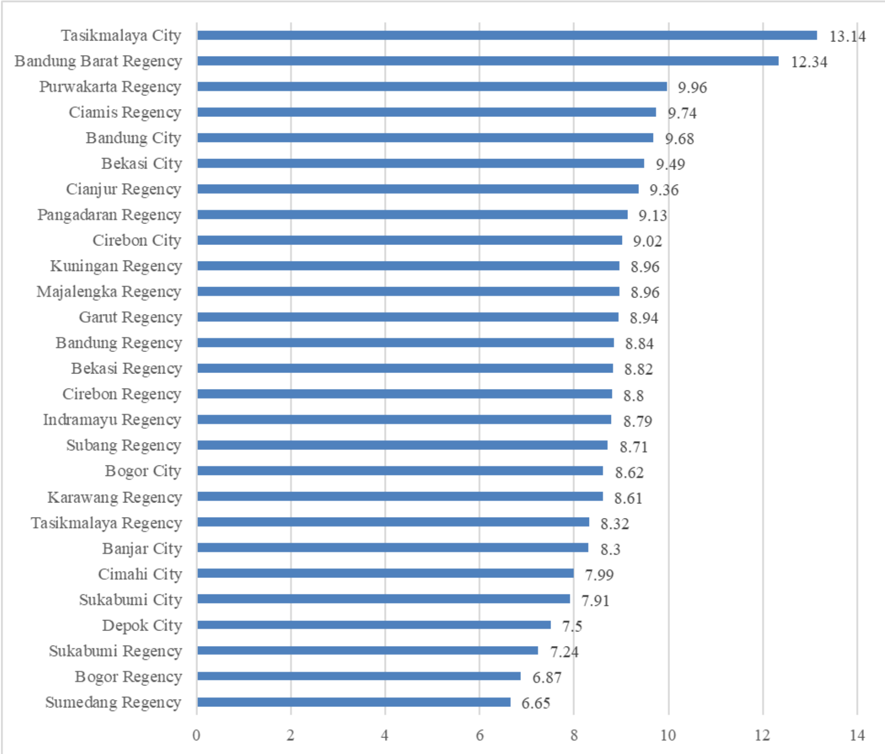

Figure 4. **Bar Chart of Variable X_3**

Based on Figure 4, It can be seen that the highest percentage of population access to proper sanitation facilities in West Java is in Depok City, namely 13.14%. Meanwhile, the lowest percentage of population access to proper sanitation facilities in West Java is in Garut Regency, namely 6.65%.

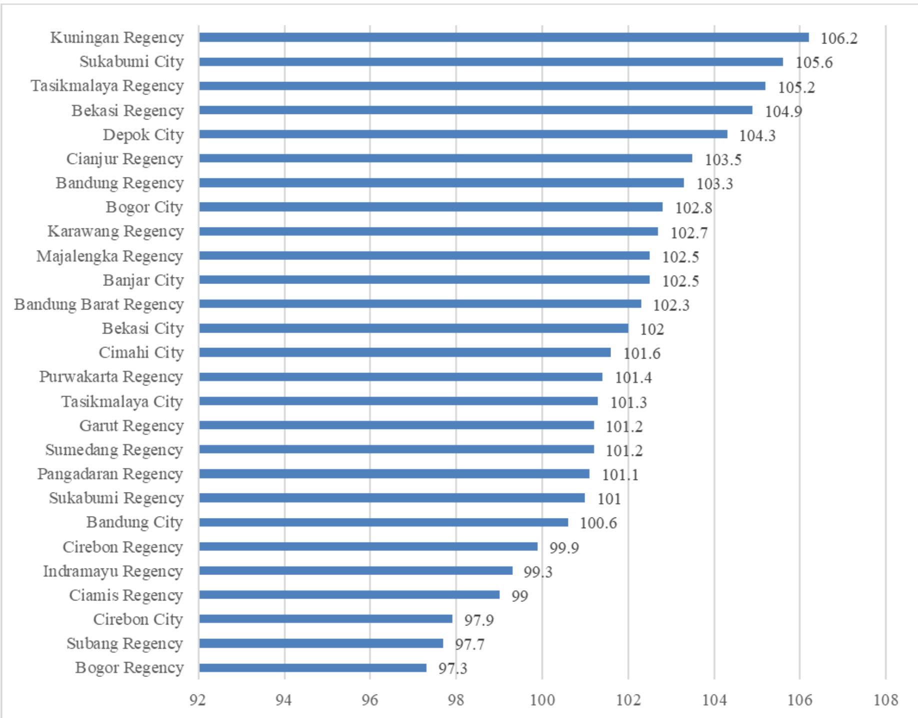

Figure 5. **Bar Chart of Variable X_4**

Based on **Figure 5**, It can be seen that the highest sex ratio in West Java is in Indramayu Regency with a value of 106.2. Meanwhile, the lowest sex ratio in West Java is in Banjar City, namely 97.3. This is in accordance with the facts on the ground that the city's male to female sex ratio is indeed high

3.2 Best Model Selection Analysis

In this study, modeling the number of chronic cases of filariasis in 27 regencies/cities in West Java Province, Indonesia using the MARS method based on 4 predictor variables is as follows.

3.2.1 Combination of Basis Functions

Calculations are carried out using MARS software by combining Basis Function (BF), Maximum Interaction (MI), and Minimum Observation between knots (MO). The Basis Function (BF) value is 8 to 16, the Maximum Interaction (MI) is 1 and 2 and the Minimum Observation between knots (MO) is 0, 1, 2, and 3. The criterion for the best model is the Generalized Cross Validation (GCV) value the lowest or minimum. If there are two or more models that have the same minimum GCV value, it can be seen from the maximum R^2 value and the minimum Mean Square Error (MSE) value.

Table 3. 15 Base Function Combinations

BF	MI	MO	GCV	R^2(Naive)	MSE
15	1	0	18.873	0.909	7.149
		1	18.121	0.912	6.864
		2	18.121	0.912	6.864
		3	21.280	0.835	11.855
	2	0	25.094	0.922	6.083
		1	**17.235**	**0.916**	**6.528**
		2	22.862	0.911	6.940
		3	21.280	0.835	11.855

Based on **Table 3**, it can be seen that the best model is obtained from a combination of fifteen basis functions, a maximum interaction of one, and a minimum of observations between knots of three. The Generalized Cross Validation (GCV) value is 17,235, R^2 is 0.916, and the Mean Square Error (MSE) is 6,528 from the results of this combination. The R^2 value of 0.916 means that the variation in the value of the response variable Number of Chronic Filariasis (Y) which can be explained by the predictor variable (X) is 91.6%. This is due to many factors that cause chronic cases of filariasis.

3.2.2 Basis Function Estimation

The best estimated model for modeling the number of chronic filariasis in West Java is as follows.

Table 4. Best Significant Model Basis Function Estimation

Basis Functions	Parameter Estimation
Intersept/Constant	53.504
$BF_2 = \max(0, X_3 - 6.650)$	-4.479
$BF_4 = \max(0, 103.300 - X_4)$	-0.944
$BF_5 = \max(0, X_2 - 2.070) \times BF_2$	-0.658
$BF_6 = max(0, X_1 - 7.350) \times BF_2$	-6.404

Based on **Table 4**, the following is the MARS model for estimating chronic cases of filariasis in West Java Province as follows.

$$\hat{Y} = 53.504 - 4.479BF_2 - 0.944BF_4 - 0.658BF_5 - 6.404BF_6$$

Then substitute the values $X_3 > 6.650$; $X_4 < 103.300$; $X_2 > 2.070$; $and\ X_1 > 7.350$; This equation can be rearranged as follows.

$$\hat{Y} = 53.504 - 4.479(X_3 - 6.650) - 0.944(103.300 - X_4) - 0.658(X_2 - 2.070)(X_3 - 6.650) - 6.404(X_1 - 7.350)(X_3 - 6.650)$$

3.2.3 Residual Assumption Test

Testing the residual assumption is very necessary for further inference. The residuals are assumed to be normally distributed and the residuals from each observation are assumed to have a constant variance and have a value of σ^2. The hypothesis used in testing the residual assumption is as follows.

H_0 : Residuals are normally distributed
H_1 : The residuals are not normally distributed

Testing is carried out using software assistance. The results of the test statistical calculations using the Kolmogorov-Smirnov normality test are as follows.

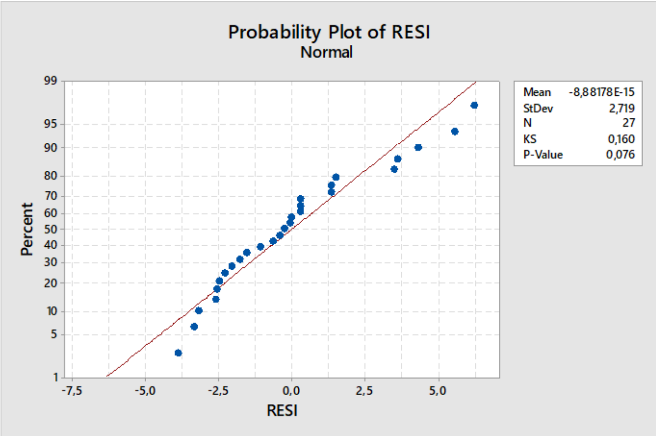

Figure 1. Residual Normality Assumption Test Plot

Source : Software Minitab 18

The critical area in this test is repulsion H_0 if the p-value is less than the significance level (α). Based on **Figure 1**, it can be seen that the p-value is 0.076, which is more than the significance level ($\alpha = 0.05$). Therefore, a decision of failure to reject was obtained H_0, so the conclusion is that the residuals are normally distributed. This test is carried out to detect symptoms of heteroscedasticity in the residuals. Testing was carried out using the Glejser test. The hypothesis used in testing heteroscedasticity using the Glejser test is as follows.

$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_{34}^2 = \sigma^2$
H_1 : There is at least one $\sigma_i^2 \neq \sigma^2; i = 1, 2, \ldots, 34$

The results of the heteroscedasticity detection test using the Glejser test with the help of Minitab software are as follows.

Table 5. Glejser Test

P-Value	Decision
1.000	Failed to Reject H_0

The critical area in this test is repulsion H_0 if the p-value is less than the significance level (α). Based on **Table 5**, it can be seen that the p-value is 1.000, which is more than the significance level ($\alpha = 0.05$). Therefore, a decision of failure to reject was obtained H_0, so the conclusion is that there was no case of heteroscedasticity.

3.3 MARS Model Significance Test

3.3.1 Simultaneous Test of MARS Model Base Function Coefficients

The hypothesis used in simultaneous testing of basis function coefficients is as follows.

$H_0 : a_2 = a_4 = a_5 = a_6 = 0$
H_1 : There is at least one $a_m \neq 0; m = 2,4,5,6$

Table 6. Simultaneous Test of MARS Model Basis Function Coefficients

Test Statistics	Value
F Test Statistics	60.352
P-value	0.152881×10^{-10}

The critical area of simultaneous testing is rejecting H_0 if the F_count is more than $F_{(0.05,4,22)}$ or the p-value is less than the significance level ($\alpha = 0.05$). Based on Table 6, the calculated p-value is 60,352, which is more than $F_{(0,05,4,22)} = 2{,}82$. Apart from that, the resulting p-value is $0{,}152881 \times 10^{-10}$, which is less than the significance level (α=0.05). Therefore, the decision obtained is to reject H_0, so the conclusion is that there is at least one a_m that is not equal to zero, with $m = 2{,}4{,}5{,}6$. This can be interpreted that the model obtained is appropriate and shows a relationship between the basis function coefficients and the response variable.

3.3.2 Partial Test of MARS Model Base Function Coefficients

The hypothesis used in partial testing of the basis function coefficients is as follows.

$H_0 : a_m = 0, m = 2{,}4{,}5{,}6$
$H_1 : a_m \neq 0, m = 2{,}4{,}5{,}6$

Table 7. Partial Test of MARS Model Basis Function Coefficients

Parameter	*Estimate*	S.E.	*T-Ratio*	*P-value*
Constant	53.504	1.941	27.563	0.999201×10^{-15}
BF_2	-4.479	0.394	-11.355	0.114089×10^{-9}
BF_4	-0.944	0.300	-3.150	0.005
BF_5	-0.658	0.221	-2.970	0.007
BF_6	-6.404	0.704	-9.099	0.653134×10^{-8}

The critical area of the partial test is to reject H_0 if $|t_{count}| > t_{(0.0025,23)}$ or p-value less than the significance level (α=0.05). Based on Table 7, the $|t_{count}|$ value is obtained of each basis function in the model is more than $t_{(0.025,23)} = 2{,}069$. In addition, the p-value of each basis function in the model is less than the significance level (α=0.05). Therefore, the decision obtained is to reject H_0, so the conclusion is that there is at least one a_m that is not equal to zero, with m=2,4,5,6. This can be interpreted that the model obtained shows a relationship between the basis function coefficients and the response variable.

3.3.3 Variable Importance Level

The level of variable importance is used to rank the predictor variables that influence the response variable. The level of importance of variables in modeling data on the number of chronic filariasis in West Java is as follows.

Table 8. Variable Importance Level

Variable	Variable Name	Level of Importance	GCV reduction
X_3	Percentage of Population Access to Adequate Sanitation Facilities	100%	84.438
X_1	Percentage of Households with Clean and Healthy Behavior	78.485%	58.632
X_4	Sex Ratio	9.633%	17.859
X_2	Percentage of Poor Population	1.129%	17.244

Based on Table 8, it can be seen that the predictor variable that has the most influence on the response variable is the variable Percentage of Population Access to Proper Sanitation Facilities (X_2) with an importance level of 100%. Apart from that, the Percentage of Population Access to Adequate Sanitation Facilities variable can reduce the Generalized Cross Validation (GCV) value by 84,438, if this variable is included in the model. Furthermore, the predictor variables that influence the response variable based on the order of importance are the Percentage of Households with Clean and Healthy Behavior (X_1) with an importance level of 78.485% and the Sex Ratio (X_4) with an importance level of 9.633%. Meanwhile, the predictor variable which has an importance level of 1.129% is the variable Percentage of Poor Population (X_2). If viewed based on the contribution of predictor variables in reducing the Generalized Cross Validation (GCV) value, it is found that the Percentage of Households with Clean and Healthy Behavior variable (X_1) can reduce the GCV value by 58,632, the Sex Ratio (X_4) can reducing the GCV value by 17,859, and the Percentage of Poor Population (X_2) can reduce the GCV value by 17,244. The reduction in GCV value applies if this variable is included in the model.

3.3.4 Mean Absolute Percentage Error Value

The Mean Absolute Percentage Error (MAPE) value is used to measure the accuracy of the best model estimation. Calculating the MAPE value can use the formula :

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t} \times 100\%$$

The MAPE value obtained is 7.01983%. This shows that the MAPE value is in the range of less than 10%, meaning that the estimation results show a very accurate ability to determine the best model.

3.4 Best MARS Model Interpretation

The best MARS model has been obtained with a combination of fifteen basis functions, maximum interactions of two, and minimum observations between knots of one. After getting the best model and testing significant variables, as well as assumptions on the residuals, the response variable and the estimated results can be plotted to compare the two values. The plot is as follows.

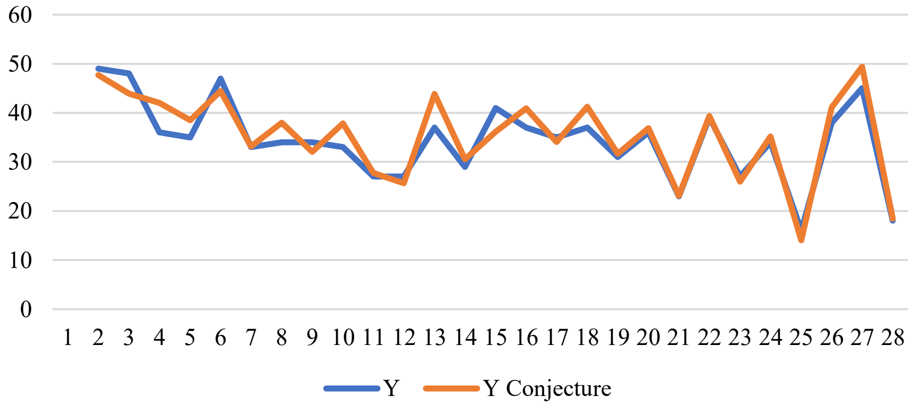

Figure 2. Plot $\boldsymbol{Y}$ with $\hat{\boldsymbol{Y}}$

Based on **Figure 2**, it can be seen that the estimated value $(\hat{Y})$ is close to the value from factual data, namely the Number of Chronic Filariasis (Y) where the x-axis describes the district/city and the y-axis describes the number of cases. Apart from that, based on the best model that has been obtained, the interpretation results are as follows.

3.4.1 Base Two Functions (BF_2)

$$BF_2 = \begin{cases} (X_3 - 6.650) & ;\ \text{for } X_3 > 6.650 \\ 0 & ;\ \text{for } X_3 \text{ another} \end{cases}$$

Interpretation of the value of the base two function (BF_2) with a coefficient of -4.479 in the best model in this equation means that every one unit increase in (BF_2) will reduce the number of chronic filtration by 4.479 with the base function (BF_4), (BF_5), and (BF_6) are considered constant. Apart from that, the value of the Percentage of Population Access to Proper Sanitation Facilities will make a significant contribution to districts/cities which has a value of more than 6.650% in the form of a reduction in the number of Chronic Filariasis by 4.479. It can be seen that all districts/cities in West Java have a value of (BF_2). This is in line with conditions in the field, because the higher the population's access to proper sanitation facilities, the greater the chance that the population will not be exposed to chronic filariasis. Apart from that, this is also due to the fact that environmental conditions where there are standing water around the house will become a potential breeding ground, especially puddles of water that are not maintained and contain water plants, so the risk of filariasis transmission can be reduced by collecting air around the house, circulating it so that the air does not pool or splash on larvae-eating fish. This activity can inhibit the breeding of mosquito vectors so that mosquito density can be reduced so that the transmission of filariasis can also be suppressed [14].

3.4.2 Base Four Functions ($\boldsymbol{BF_4}$)

$$BF_4 = \begin{cases} (103.300 - X_4) & ;\ \text{for } X_4 < 103.300 \\ 0 & ;\ \text{for } X_4 \text{ another} \end{cases}$$

Interpretation of the value of the four basis function (BF_4) with a coefficient of -0.944 in the best model means that every one unit increase in (BF_4) will reduce the number of chronic filariasis by 0.944 with the basis function (BF_2), (BF_5), and (BF_6) are considered constant. Apart from that, another meaning is that districts/cities that have a sex ratio value of less than 103.300 will make a significant contribution, namely in the form of reducing the number of chronic filariasis by 0.944. It can be seen that almost all districts/cities in West Java have a value of (BF_4). Meanwhile, the district/city that does not have (BF_4) is Bandung Regency. This is because men often work at night so they have a higher potential for contact with vectors. Research proves that men have a 4.7 times greater risk than women of facing filariasis [15].

3.4.3 Base Five Functions ($\boldsymbol{BF_5}$)

$$BF_5 = \begin{cases} (X_2 - 2.070)(X_3 - 6.650) & ;\ \text{for } X_2 > 2.070 \text{ and } X_3 > 6.650 \\ 0 & ;\ \text{for } X_2 \textit{ and } X_3 \text{ another} \end{cases}$$

Interpretation of the value of the five basis function (BF_5) with a coefficient of -0.658 in the best model means that every one unit increase in (BF_5) will reduce the number of chronic filariasis by 0.658 with the basis function (BF_2), (BF_4) and (BF_6) are considered constant. Apart from that, another meaning is that districts/cities that have a value for the Percentage of Poor Population variable of more than 2.070% and a value of Percentage of Population Access to Proper Sanitation Facilities of more than 6.650% will make a significant contribution, namely in the form of reducing the number of chronic filariasis by 0.658%. It can be seen that almost all districts/cities in West Java have a value of (BF_5). Meanwhile, the district/city that does not have (BF_5) is Tasikmalaya City.

3.4.4 Base Six Functions ($\boldsymbol{BF_6}$)

$$BF_6 = \begin{cases} (X_1 - 7.350)(X_3 - 6.650) & ;\ \text{for } X_1 > 7.350 \textit{ and } X_3 > 6.650 \\ 0 & ;\ \text{for } X_1 \textit{ and } X_3 \text{ another} \end{cases}$$

Interpretation of the value of the nine basis function (BF_6) with a coefficient of -6.404 in the best model means that every one unit increase in (BF_6) will reduce the number of chronic filariasis by 6.404 with the basis function (BF_2), (BF_4), and (BF_5) are considered constant. Apart from that, another meaning is that districts/cities that have a value on the Percentage of Households with Clean and Healthy Behavior variable of more than 7,350% and a value of Percentage of Population Access to Proper Sanitation Facilities of more than 6,650% will make a significant contribution, namely a reduction in the number of chronic filariasis by 6,404%. It can be seen that almost all districts/cities in West Java have a value of (BF_6). Meanwhile, the district/city that does not have (BF_6) is Cirebon City.

4. CONCLUSIONS

Based on the results of the analysis and discussion that has been carried out, the conclusions obtained from the research as follows:

1. Descriptive statistics for each variable in this study are the response variable. The highest number of chronic filariasis in West Java is in Cianjur Regency with 49 cases. Meanwhile, the lowest number of chronic filariasis in West Java was in Tasikmalaya Regency with 16 cases. In addition, the mean is 34.30 and the variance is 66.14. On the other hand, the scatterplot of each predictor variable against

the response variable shows that there is no particular data distribution pattern (trend) so it is suitable for use with a non-parametric regression approach.

2. The best model obtained using the Multivariate Adaptive Regression Spline (MARS) method is a combination of fifteen basis functions, a maximum of two interactions, and a minimum of one observation between knots. The Generalized Cross Validation (GCV) value is 17,235, R^2 is 0.916, and the Mean Square Error (MSE) is 6,528 from the results of this combination. The R^2 value of 0.916 means that the variation in the value of the response variable Number of Chronic Filariasis (Y) which can be explained by the predictor variable (X) is 91.6%. The following is the best model obtained using the MARS method.

$$\hat{Y} = 53.504 - 4.479BF_2 - 0.944BF_4 - 0.658BF_5 - 6.404BF_6$$

3. The best model interpretation in this research is as follows.
   a. Based on the best model that has been obtained, it is found that the interpretation of the value of the base two function (BF_2) means that the value of the Percentage of Population Access to Appropriate Sanitation Facilities will make a significant contribution to districts/cities which have a value of more than 6,650% in the form of a reduction in the number of Chronic Filariasis amounting to 4,479.
   b. Based on the best model that has been obtained, it is found that the interpretation of the base four function value (BF_4) means that districts/cities that have a Sex Ratio value of less than 103,300 will make a significant contribution, namely in the form of reducing the number of Chronic Filariasis by 0.944.
   c. Based on the best model that has been obtained, it is found that the interpretation of the value of the base five function (BF_5) means that districts/cities have a value for the Percentage of Poor Population variable of more than 2.070% and a value of Percentage of Population Access to Adequate Sanitation Facilities of more than 6.650%. will provide a significant contribution, namely in the form of reducing the number of chronic filariasis by 0.658.
   d. Based on the best model that has been obtained, it is found that the interpretation of the value of the base six function (BF_6) means that districts/cities have a value on the Percentage of Households with Clean and Healthy Behavior variable of more than 7,350% and a value The percentage of population access to proper sanitation facilities of more than 6,650% will provide a significant contribution, namely in the form of reducing the number of chronic filariasis by 6,404.

ACKNOWLEDGMENT

This research could not have been carried out without help and support from various parties. For this reason, the author would like to thank the Statistics Study Program, Faculty of Science and Technology, Airlangga University for its support and all parties who helped run this research activity smoothly.

REFERENCES

[1] R. Amelia, "Analysis of Risk Factors for Filariasis", *Unnes Journal of Public Health*, vol. 3, no. 1, pp. 2-3, March 2019.
[2] R. Novita, "The Impact of Climate Change on the Emergence of Mosquito Transmitted Diseases, Especially Lymphatic Filariasis", *Journal of Health Epidemiology and Communicable Diseases*, vol. 5, no. 1, pp. 30-39, May 2019.
[3] World Heath Organization. "Lymphatic Filariasis: Reporting Continued Progress Towards elimination as a Public Health Problem" Address: https://www.who.int/news/item/29-10-2020-lymphatic-filariasis-reporting-continued-progress-towards-elimination-as-a-publichealth-problem [accessed Juli. 03, 2021].
[4] Indonesian Ministry of Health, *Indonesian Health Profile.* Jakarta : Directorate General of Disease Prevention and Control (P2P) Ministry of Health of the Republic of Indonesia, 2021.
[5] West Java Provincial Health Service. Bandung : Filariasis Situation in West Java Province 2007-2011. West Java Provincial Health Service, 2017.
[6] Wicaksono, "Multivariate Adaptive Regression Spline (MARS) Modeling on Risk Factors for Diarrhea Disease Rates", *Gaussian Journal*, vol. 3, no. 2, pp. 253- 262, June 2014.
[7] F. Lembang, "Analysis of Poverty in West Southeast Maluku Regency Using the Multivariate Adaptive Regression Spline (MARS) Approach, Journal of Media Statistics", *Journal of Media Statistics*, vol. 12 no. 2, pp. 188-199, December 2019.
[8] Bappenas. *Jakarta:UNDP*. Indonesia: Progress Report on Achieving the Millennium Development Goals, 2016.

[9] V. Wiratna, *Quantitative Qualitative Research Methods and R&D*. Bandung: Alphabeta Bandung, 2014.
[10] A. Sholikhah. "Descriptive Statistics in Qualitative Research", *Journal of Communication*, vol. 10, no. 2, pp. 342-362, December 2016.
[11] Sugiyono, *Quantitative, Qualitative and R&D Research Methods*. Bandung: CV Alphabet, 2017.
[12] Winarno, Research Methodology in Physical Education. Malang: UM Press, 2013.
[13] L. Martias, "Descriptive Statistics as a Collection of Information", *FIHRIS: Journal of Library and Information Science*, vol. 16, no. 1, pp. 40-59, January 2021.
[14] Santoso, "Relationship between physical environmental factors and the incidence of filariasis in Indonesia," *Journal of Health Ecology*, vol. 13, no. 3, pp. 210-218, September 2018.
[15] P. Juriastuti, M. Kartika, IM. Djaja, D. Susanna, "Risk Factors for Filariasis in Jati Sampurna Village," *Makara Health*, vol. 14 no. 1, pp. 31-36, June 2017.