Modification of Multivariate Adaptive Regression Spline (MARS)

Septia Devi Prihastuti Yasmirullah¹, Bambang Widjanarko Otok¹*, Jerry Dwi Trijoyo Purnomo¹, and Dedy Dwi Prastyo¹

¹Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia

*E-mail: bambang_wo@statistika.its.ac.id

Abstract. Multivariate Adaptive Regression Spline (MARS) is a nonparametric regression method that can accommodate additive effects and interaction effects between predictor variables. Generally, MARS has been used for modeling pairs of data with continuous or categorical responses. One type of categorical data that needs special attention in modeling is count data. The count data is often encountered, especially in the health sector. The existence of count data motivates the development of the theory and application of the MARS method, which is the Multivariate Adaptive Poisson Regression Spline (MAPRS). The MAPRS is a combination of MARS and Poisson regression. It can accommodate and analyze the data according to its type and distribution. The application of MAPRS to model the count of Tuberculosis (TB) shows that it outperforms the Poisson regression

1. Introduction

Multivariate Adaptive Regression Spline (MARS) is a method that was introduced by Friedman [1] in 1991. This method is a nonparametric regression method that can accommodate additive and interaction effects between predictor variables [1]. MARS can obtain good predictive results for the shape of the regression curve from the unknown pattern relationships of response and predictors [2]. MARS also does not assume the functional relationship form of the response and predictor variables. It has a flexible and functional form. MARS can handle data that has changed behavior at certain sub-intervals because there is a knot in MARS, indicating a change in data behavior patterns.

The MARS method is a combination of the truncated spline method and recursive partitioning regression (RPR). The truncated spline method has limitations in determining the position and the number of knots used when involving multiple predictors. There will be so many combinations concerning the number of predictors, the position of the knots, and the number of knots [3]. In this case, the MARS method can overcome the weakness of the truncated spline because the knot determination in MARS is not sought individually from the combination but through an adaptive process. The adaptive process in MARS has been carried out with a stepwise algorithm, which includes forward and backward. Forward stepwise, build the model by adding truncated spline (knots and interaction) basis functions to obtain a model with the maximum number of basis functions. Meanwhile, backward stepwise is used to get a parsimony model by selecting the forward stepwise basis function. It has the most significant contribution to the estimated response based on the minimum Generalized Cross-Validation (GCV) value.
Generally, MARS modeling has been used on data with continuous or categorical responses. The response is continuous if the measurement scale is interval and ratio. Meanwhile, the response is categorical if the measurement scale is nominal or ordinal. Many previous studies have examined the MARS method for continuous and categorical responses. Previous research on MARS with a continuous response was applied to modeling eye fatigue with MARS [4], modeling of welfare indicators in Java with bi-responses MARS [5], and the simulation study and application of bi-responses nonparametric regression model using MARS [6], [7]. Meanwhile, previous research on MARS with a categorical response was applied to modeling diabetes using Bootstrap Aggregating MARS [8], classifying HIV/AIDS patients with random forest and MARS [9], classifying poor households in Jombang Regency with Bootstrap Aggregating MARS [10], and parameter estimation of Multivariate Adaptive Regression Spline (MARS) to Multi Drug-Resistant Tuberculosis (MDR-TB) modeling in Lamongan Regency [11].

One of the categorical data that attracts special attention is count data. The count data type is often found in a variety of fields, especially the health sector. The method commonly used to model count data is Poisson regression. However, there are still limitations in using the Poisson regression method. It motivated the theory development and application of the MARS method, namely the Multivariate Adaptive Poisson Regression Spline (MAPRS). This method is a combination of the MARS method and the Poisson Regression. The MAPRS method is expected to be able to model data according to its type and distribution.

In this study, the MAPRS method has been applied to model the number of Tuberculosis (TB) data. Moreover, the number of TB data has been classified as the count data type. The WHO report states that in 2018, there were 44% of TB cases occurring in Southeast Asia, 24% in Africa, 18% in the Western Pacific, 3% in America, and 3% in Europe. Meanwhile, 8% of them occur in Indonesia, so Indonesia ranks third after India and China [12]. The TB in Indonesia is under concern based on the commitment of the Indonesian government at the Global Sustainable Development Goals (SDGs) meeting to eliminate TB by 2030 [12]. Therefore, as another effort to accelerate the identification of TB extension chains, modeling is necessary. Mathematical and statistical modeling can explain the relationship between the number of TB and its predictor variables.

2. Methods

2.1. Multivariate Adaptive Regression Spline (MARS)

Multivariate Adaptive Regression Spline (MARS) is a nonparametric method that combines the truncated spline method with Recursive Partitioning Regression (RPR) [1]. The MARS model as follows:

\[ y_i = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} \left[ s_{km}(x_{v(k,m)i} - t_{km}) \right] + \epsilon_i, \]

where,

- \( a_0 \): a constant parameter of basis function
- \( a_m \): a non-constant parameter of \( m \)-th basis function
- \( M \): the number of maximum basis function
- \( K_m \): the maximum interaction of \( m \)-th basis function
- \( s_{km} \): the sign of basis function in the \( k \)-th interaction and \( m \)-th basis function, where \( s_{km} \) is (+1) or (-1)
- \( x_{v(k,m)i} \): the \( v \)-th predictor variable, where \( v \) is an index of predictor variables related to \( k \)-th interaction and \( m \)-th basis function in MARS function.
- \( t_{km} \): the value of knot in \( k \)-th interaction and \( m \)-th basis function.
The MARS model in equation (1) has been rewritten in matrix form as in equation (2).

\[ y = Ba + \varepsilon \]  

(2)

where,

\[ y_{m+1} = (y_1, \ldots, y_n)^T, \]

\[ a_{(M+1)} = (a_0, \ldots, a_M)^T, \]

\[ \varepsilon_{m+1} = (\varepsilon_1, \ldots, \varepsilon_n)^T, \]

\[
B_{m(M+1)} = \begin{bmatrix}
1 & \prod_{k=1}^{K_1} s_{k1}(x_{(k,1)_1} - t_{k1}) & \cdots & \prod_{k=1}^{K_M} s_{kM}(x_{(k,M)_1} - t_{kM}) \\
1 & \prod_{k=1}^{K_1} s_{k1}(x_{(k,1)_2} - t_{k1}) & \cdots & \prod_{k=1}^{K_M} s_{kM}(x_{(k,M)_2} - t_{kM}) \\
\vdots & \vdots & \ddots & \vdots \\
1 & \prod_{k=1}^{K_1} s_{k1}(x_{(k,1)_n} - t_{k1}) & \cdots & \prod_{k=1}^{K_M} s_{kM}(x_{(k,M)_n} - t_{kM})
\end{bmatrix}
\]

2.2. Poisson Regression

Poisson regression is an approach to count data analysis. Furthermore, the model formed is non-linear [13]. If \( y \) is the number of events that occur within a certain period or area, then the Poisson regression assumes \( y \) is a random variable with the Poisson distribution.

\[ \Pr(Y = y | \mu) = \frac{e^{-\mu} \mu^y}{y!} \]  

(3)

The Poisson regression model belongs to the Generalized Linear Model (GLM). Moreover, The GLM consists of three components, i.e., the random component, the systematic component, and the link function [14]. Furthermore, the link function is a component that connects random components with systematic components, \( E(y) = \eta \). Meanwhile, the link function of the Poisson regression model has obtained in the following ways:

a) Perform the logarithms of the two sides of the equation (3).

\[ \log[ \Pr(Y = y | \mu) ] = \log \left( \frac{e^{-\mu} \mu^y}{y!} \right) = -\mu + y \log \mu - \log y! \]

b) Perform the exponential of the two sides of the equation obtained in step (a).

\[ \exp \left\{ \log[ \Pr(Y = y | \mu) ] \right\} = \exp \left\{ -\mu + y \log \mu - \log y! \right\} \]

\[ \Pr(Y = y | \mu) = \exp \left\{ -\mu + y \log \mu - \log y! \right\} \]

c) Perform the mathematical manipulations
\[
f(y; \theta, \phi) = \exp \left\{ \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \right\} \\
= \exp \left\{ -\mu + y \log \mu - \log y! \right\} \\
= \exp \left\{ y \log \mu - \mu - \log y! \right\}
\]

where,
\[
y = y \\
\theta = \log(\mu), \mu = e^\theta \\
b(\theta) = \mu = e^\theta \\
\phi = 1 \\
a(\phi) = \phi \\
c(y, \phi) = -\log(y!)
\]

Furthermore, the link function for the Poisson regression model is \( \log(\mu) \), so \( \log(\mu) = X^T \beta \) or \( \mu = \exp(X^T \beta) \) [15].

### 2.3. Multivariate Adaptive Poisson Regression Spline (MAPRS)

Multivariate Adaptive Poisson Regression Spline (MAPRS) is a combination of Poisson Regression and MARS. The MAPRS model is as follows:

\[
Y_i \sim \text{Poisson}(\mu)
\]

\[
\ln(\mu_i) = f(x_i) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K} \left[ s_{km}(x_{v(k,m)i} - t_{km}) \right] \\
= a_0 + \sum_{m=1}^{M} a_m B_{mi}(x_i)
\]

\[
\mu_i = \exp \left( a_0 + \sum_{m=1}^{M} a_m B_{mi}(x_i) \right)
\]

The parameter estimation of the MAPRS model in equation (4) uses the Weighted Least Square (WLS).

\[
\psi = \sum_{i=1}^{n} V_i^{-1} e_i^2 = \sum_{i=1}^{n} \left( y_i - \mu_i \right)^2 \frac{1}{\mu_i}
\]

then,

\[
\psi = \sum_{i=1}^{n} V_i^{-1} e_i^2 = \sum_{i=1}^{n} \left( y_i - e^{a_0 + \sum_{m=1}^{M} a_m B_{mi}(x_i)} \right)^2 \frac{1}{e^{a_0 + \sum_{m=1}^{M} a_m B_{mi}(x_i)}}
\]

The estimate of parameter \( \mathbf{a} \) is obtained by deriving equation (6) with respect to \( \mathbf{a} \), then equating it to zero.
2.4. Research Variable

The application of this research used secondary data from the Health Profile of Lamongan Regency in 2017, where Lamongan is one of the regencies in East Java Province, Indonesia. Furthermore, Table 1 shows the research variables in this study. The unit of observation in this research is the subdistrict.

| Code | Variables | Scale |
|------|-----------|-------|
| Y    | The number of TB | Ratio |
| X1   | Population density (people/km²) | Ratio |
| X2   | HIV/AIDS prevalence (per 10,000 population) | Ratio |
| X3   | Percentage of households with PHBS (%) | Ratio |
| X4   | Percentage of healthy house (%) | Ratio |
| X5   | Ratio of primary health facilities (per 10,000 population) | Ratio |
| X6   | Ratio of health workers (per 10,000 population) | Ratio |
| X7   | Percentage of the population enrolled in school (%) | Ratio |

2.5. Steps of Analysis

The steps of analysis in this study are as follows:

1. Parameter estimation of MAPRS model
2. Data exploration
3. Analyze the data with the MAPRS method
   - Determine the possibility of the maximum number of basis functions (BF)
   - Determine the maximum number of interactions (MI)
   - Determine the minimum observation (MO) between knots by trial and error
   - Determine the best model based on the minimum GCV value
4. Testing the significance of basis function coefficients
5. Interpreting the best model
6. Drawing the conclusions and suggestions

3. Results and Discussion

3.1. Parameter Estimation of Multivariate Adaptive Poisson Regression Spline (MAPRS)

The Multivariate Adaptive Poisson Regression Spline (MAPRS) model have shown by equation (8), which is:

\[
\ln \mu_i = f(x_i) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K} \left[ s_{km} \left( x_{i(k,m)} - t_{km} \right) \right] = a_0 + \sum_{m=1}^{M} a_m B_{mi}(x_i)
\]

\[
\mu_i = \exp \left( a_0 + \sum_{m=1}^{M} a_m B_{mi}(x_i) \right)
\]

Next, the Poisson distribution has

\[ E[Y] = \text{Var}[Y] = \mu \]

so a weighting matrix can be formed as:

\[
W = \begin{bmatrix}
\frac{1}{\mu_1} & 0 & 0 & 0 \\
0 & \frac{1}{\mu_2} & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \frac{1}{\mu_n}
\end{bmatrix}
\]
The parameter estimation of the MAPRS model uses the Weighted Least Square (WLS) method, which minimizes the following functions:

\[
\psi = \varepsilon^T W \varepsilon
\]

\[
= (y - \exp(Ba))^T W (y - \exp(Ba))
\]

\[
= y^T W y - \exp(a^T B^T) W y - \exp(a^T B^T) W \exp(Ba)
\]

\[
= y^T W y - 2\exp(a^T B^T) W y + \exp(a^T B^T) W \exp(Ba)
\]

Then, find the first derivative of the function \( \psi \) with respect to \( a \), then equalize zero.

\[
\frac{\partial (\psi)}{\partial (a)} = \frac{\partial \left[y^T W y - 2\exp(a^T B^T) W y + \exp(a^T B^T) W \exp(Ba)\right]}{\partial (a)}
\]

\[
0 = -2B^T \exp(a^T B^T) W y + 2B^T \exp(a^T B^T) W \exp(Ba)
\]

Next,

\[
2B^T \exp(a^T B^T) W \exp(Ba) = 2B^T \exp(a^T B^T) W y
\]

\[
B^T \exp(a^T B^T) W \exp(Ba) = B^T \exp(a^T B^T) W y
\]

\[
\left(B^T \exp(a^T B^T) W \right)^{-1} (B^T \exp(a^T B^T) W) \exp(Ba) = \left(B^T \exp(a^T B^T) W \right)^{-1} B^T \exp(a^T B^T) W y
\]

\[
\exp(Ba) = \left(B^T \exp(a^T B^T) W \right)^{-1} B^T \exp(a^T B^T) W y
\]

\[
Ba = \ln \left( \left(B^T \exp(a^T B^T) W \right)^{-1} B^T \exp(a^T B^T) W y \right)
\]

\[
(B)^{-1} Ba = (B)^{-1} \ln \left( \left(B^T \exp(a^T B^T) W \right)^{-1} B^T \exp(a^T B^T) W y \right)
\]

\[
\hat{a} = (B)^{-1} \ln \left( \left(B^T \exp(\hat{a}^T B^T) W \right)^{-1} B^T \exp(\hat{a}^T B^T) W y \right)
\]

So,

\[
\hat{a}_{WLS} = (B)^{-1} \ln \left( \left(B^T \exp(\hat{a}^T B^T) W \right)^{-1} B^T \exp(\hat{a}^T B^T) W y \right)
\]

Therefore,

\[
\hat{f}(X) = \exp(B \hat{a}) = \exp \left(B (B)^{-1} \ln \left( \left(B^T \exp(\hat{a}^T B^T) W \right)^{-1} B^T \exp(\hat{a}^T B^T) W y \right) \right)
\]

\[
= \exp \left( \ln \left( \left(B^T \exp(\hat{a}^T B^T) W \right)^{-1} B^T \exp(\hat{a}^T B^T) W y \right) \right)
\]

\[
= \left(B^T \exp(\hat{a}^T B^T) W \right)^{-1} B^T \exp(\hat{a}^T B^T) W y
\]
3.2. Application of Multivariate Adaptive Poisson Regression Spline (MAPRS) Model

3.2.1. Conditions of Tuberculosis in Lamongan Regency
The spread of TB count in the Lamongan Regency at each subdistrict is shown in Figure 1.

![Figure 1. The Map of TB Count Spread in Lamongan Regency](image)

Figure 1 shows that the color on the map represents the count number. If the color in an area gets darker, the count of TB in that area gets higher. Meanwhile, the color is getting lighter, then the number of TB in that area is getting less. The Lamongan subdistrict, as the capital of Lamongan regency, is the area with the highest number of TB.

3.2.2. Modeling the Number of Tuberculosis using MAPRS
Table 2 shows the results of modeling the number of TB in Lamongan regency using the MAPRS method based on the combinations of BF, MI, and MO.

| BF | MI | MO | GCV   | $R^2$  | BF | MI | MO | GCV   | $R^2$  | BF | MI | MO | GCV   | $R^2$  |
|----|----|----|-------|-------|----|----|----|-------|-------|----|----|----|-------|-------|
| 14 | 1  | 0  | 1570.2687 | 0.7582 | 21 | 1  | 0  | 1248.7060 | 0.8077 | 28 | 1  | 0  | 1172.4234 | 0.8194 |
| 14 | 1  | 1  | 254.8554 | 0.9608 | 21 | 1  | 1  | 53.9073 | 0.9917 | 28 | 1  | 2  | 15.3046 | 0.9976 |
| 14 | 1  | 2  | 477.1446 | 0.9265 | 21 | 1  | 2  | 82.8560 | 0.9872 | 28 | 1  | 3  | 32.1159 | 0.9951 |
| 14 | 1  | 3  | 518.4397 | 0.9202 | 21 | 1  | 3  | 63.5218 | 0.9902 | 28 | 1  | 5  | 6.3719 | 0.9990 |
| 14 | 1  | 5  | 718.9700 | 0.8908 | 21 | 1  | 5  | 75.6308 | 0.9884 | 28 | 1  | 10 | 2098.1169 | 0.6769 |
| 14 | 10 | 10 | 708.8970 | 0.6769 | 21 | 1  | 10 | 2098.1169 | 0.6769 | 28 | 1  | 10 | 2098.1169 | 0.6769 |
| 14 | 2  | 0  | 365.0884 | 0.9438 | 21 | 2  | 0  | 124.2536 | 0.9809 | 28 | 2  | 0  | 29.5933 | 0.9954 |
| 14 | 2  | 1  | 434.859 | 0.9902 | 21 | 2  | 1  | 7.9043 | 0.9985 | 28 | 2  | 1  | 7.9043 | 0.9985 |
| 14 | 2  | 2  | 163.7818 | 0.9748 | 21 | 2  | 2  | 9.7880 | 0.9985 | 28 | 2  | 2  | 1.5328 | 0.9998 |
| 14 | 2  | 3  | 230.9739 | 0.9644 | 21 | 2  | 3  | 16.0775 | 0.9975 | 28 | 2  | 3  | 3.9959 | 0.9994 |
| 14 | 2  | 5  | 284.9160 | 0.9561 | 21 | 2  | 5  | 84.8330 | 0.9869 | 28 | 2  | 5  | 5.5103 | 0.9992 |
| 14 | 2  | 10 | 762.3131 | 0.8826 | 21 | 2  | 10 | 288.4635 | 0.9556 | 28 | 2  | 10 | 122.3855 | 0.9812 |
| 14 | 3  | 0  | 350.0375 | 0.9461 | 21 | 3  | 0  | 118.6704 | 0.9817 | 28 | 3  | 0  | 17.2576 | 0.9973 |
| 14 | 3  | 1  | 58.9519 | 0.9909 | 21 | 3  | 1  | 2.4053 | 0.9996 | 28 | 3  | 1  | 2.4053 | 0.9996 |
| 14 | 3  | 2  | 163.7818 | 0.9748 | 21 | 3  | 2  | 5.6609 | 0.9991 | 28 | 3  | 2  | 5.6609 | 0.9991 |
According to the estimates of the parameter summarized in Table 2, the best model has a GCV = 0.1145, which is attained at a combination of BF = 28, MI = 3, and MO = 3.

\[
\hat{f}(x) = 2.851 + 0.004BF_1 - 0.001BF_2 + 0.328BF_3 - 0.097BF_4 - 0.012BF_5 - 0.034BF_6 + 0.479BF_7
\]
\[
+ 0.010BF_8 - 0.434BF_9 + 0.018BF_{10} + 0.001BF_{11} - 0.001BF_{12} + 0.0002BF_{13} + 0.015BF_{14}
\]
\[
+ 0.007BF_{15} + 0.006BF_{16} + 0.007BF_{17} + 0.004BF_{18} - 0.007BF_{19} + 0.00005BF_{20} - 0.002BF_{21}
\]
\[-0.002BF_{22}
\]

where,

\[
BF_1 = h(834.23-x1); \quad BF_2 = h(0.294-x2); \quad BF_3 = h(0.294-x2); \quad BF_4 = h(0.294-x2); \quad BF_5 = h(78.6-x3);
\]

\[
BF_6 = h(0.294-x3); \quad BF_7 = h(14.91-x5); \quad BF_8 = h(14.91-x5); \quad BF_9 = h(79.14-x7); \quad BF_{10} = h(x7-79.14);
\]

\[
BF_{11} = h(x1-834.23) * h(x2); \quad BF_{12} = h(605.88-x1) * h(x7-79.14); \quad BF_{13} = h(1-x1-605.88) * h(x7-79.14);
\]

\[
BF_{14} = h(3.23-x2) * h(x7-79.14); \quad BF_{15} = h(x2-3.23) * h(x7-79.14); \quad BF_{16} = h(88.02-x4) * h(x5-14.91);
\]

\[
BF_{17} = h(x4-88.02) * h(x5-14.91); \quad BF_{18} = h(x5-14.91) * h(x6-11.88);
\]

\[
BF_{19} = h(x5-14.91) * h(x11-11.88); \quad BF_{20} = h(605.88-x1) * h(x5-79.14);
\]

\[
BF_{21} = h(88.02-x4) * h(x5-14.91) * h(x6-11.88); \quad BF_{22} = h(x4-88.02) * h(x5-14.91) * h(x6-11.88)
\]

Next, interpretation is given for one of the basis functions from the best model, which is attained at \(BF_7 = h(14.91-x5)\). If the ratio of primary health facilities is less than 14.91, then the \(BF_7\) coefficient will have a significant effect. Furthermore, if there is an increase in \(BF_7\) by one unit, whereas the other basis functions is considered constant, then the number of TB will increase by 0.479. Then, Table 3 shows the significance test for each of the basis functions has been performed. According to Table 3, all basis function coefficients have a significant effect on the model.

### Table 3. Testing of Function Basis Coefficients

| Coefficients of Basis Function | Estimate | Std. Error | Pr(>|t|) |
|--------------------------------|----------|------------|----------|
| (Intercept)                    | 2.8510   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x1-834.23)              | -0.0015  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(834.23-x1)              | 0.0042   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x1-834.23)*x2           | 0.0012   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x2-2.94)                | -0.0973  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x2-2.94)                | 0.3286   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x7-79.14)               | 0.0190   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x7-79.14)               | -0.4341  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x5-14.91)               | 0.0102   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x5-14.91-x5)            | 0.4791   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x1-605.88)*h(x7-79.14)  | 0.0002   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(605.88-x1)*h(x7-79.14)  | -0.0015  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x5-14.91)*h(x6-11.88)   | 0.0048   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x5-14.91)*h(x11-11.88)  | -0.0074  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x3-78.6)                | -0.0348  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x7-78.6-x3)             | -0.0121  | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(605.88-x1)*x5*h(x7-79.14)| 0.0001 | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x4-88.02)*h(x5-14.91)   | 0.0075   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(88.02-x4)*h(x5-14.91)   | 0.0062   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x4-88.02)*h(x5-14.91)*h(x6-11.88) | -0.0029 | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x4-88.02)*h(x5-14.91)*h(x6-11.88) | -0.0026 | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(88.02-x4)*h(x5-14.91)*h(x6-11.88) | 0.00480 | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(x2-3.23)*h(x7-79.14)    | 0.0080   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |
| bx1,1h(3.23-x2)*h(x7-79.14)    | 0.0152   | \(\approx 0.0000\) | \(\approx 0.000000^*\) |

Note: * Significant at 0.05 level
3.3. Comparison of MAPRS and Poisson Regression

In this study, the proposed MAPRS model is compared with the Poisson regression model. The criteria used for comparison are $R^2$ and RMSE, as summarized in Table 4. The empirical results show that the MAPRS model outperforms the Poisson regression model. The MAPRS model has a higher $R^2$ value and a smaller RMSE value than those produced by the Poisson regression model.

| No | Methods               | $R^2$     | RMSE     |
|----|-----------------------|-----------|----------|
| 1  | MAPRS                 | 0.9999824 | 112.2235 |
| 2  | Poisson Regression    | 0.6751645 | 112.2957 |

4. Conclusion

In this study, the proposed model is Multivariate Adaptive Poisson Regression Spline (MAPRS), a combination of Poisson Regression and MARS. The parameter estimation of the MAPRS model uses the Weighted Least Square (WLS) method. On applying MAPRS to model the number of TB in Lamongan regency, the best model has a GCV of 0.1145, which is attained at the combination of BF = 28, MI = 3, and MO = 3. Furthermore, based on the $R^2$ and RMSE criteria, the proposed MAPRS model outperforms the Poisson regression model. The MAPRS model has a higher $R^2$ value and a smaller RMSE value than those generated by the Poisson regression model.

Acknowledgments

We would like to express our gratitude to the Ministry of Research and Technology/National Research and Innovation Agency of Republic Indonesia (RISTEK-BRIN) for the financial support of this research. The authors would also like to thank the anonymous reviewers for their valuable comments and suggestions for improving the quality of this paper.

References

[1] Friedman, J. 1991 Multivariate Adaptive Regression Splines The Annals of Statistics 19 1-141
[2] Ampulembang, A. P. 2017 Pengembangan Model Regresi Nonparametrik Birespon Kontinu Menggunakan MARS PhD Thesis Institut Teknologi Sepuluh Nopember, Surabaya
[3] Hastie, T., Tibshirani, R. and Friedman, J.H. 2008 The Element of Statistical Learning: Data Mining, Inference and Prediction (New York: Springer Series in Statistics)
[4] Hiola, R., Otok, B.W., and Hiola, R. 2016 Model Kelelahan Mata Pengrajin Kerawang Berdasarkan Pengukuran Visus Menggunakan Multivariate Adaptive Regression Spline (MARS) Jurnal Statistika Universitas Muhammadiyah Semarang 4 1-7
[5] Ampulembang, A.P., Otok, B.W., Rumiati, A.T., and Budiasih 2016 Modeling of Welfare Indicators in Java Island Using Biresponses MARS International Journal of Applied Mathematics & Statistics 54 66-75
[6] Ayub Parlin Ampulembang et al 2016 AIP Conf Proc. 1707 080003
[7] Ampulembang, A.P., Otok, B.W., and Rumiati, A.T. 2015 Biresponses Nonparametric Regression Model Using MARS and its Properties Applied Mathematical Sciences 9 1417-1427
[8] Otok, B.W., Putra, R.Y., Sutikno, and Yasmirullah, S.D.P. 2020 Bootstrap Aggregating Multivariate Adaptive Regression Spline for Observational Studies in Diabetes Cases Systematic Reviews in Pharmacy 11 406-413
[9] Nidhomuddin, and Otok, B.W. 2015 Random Forest dan Multivariate Adaptive Regression Spline (MARS) Binary Response untuk Klasifikasi Penderita HIV/AIDS di Surabaya Jurnal Statistika Universitas Muhammadiyah Semarang 3 49-57
[10] Arleina, O.D., and Otok, B.W. 2014 Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) untuk Mengklasifikasikan Rumah Tangga Miskin di Kabupaten Jombang Jurnal Sains dan Seni ITS 3 D91-D96
[11] S D P Yasmirullah et al 2021 J. Phys.: Conf. Ser. 1752 012017
[12] WHO 2019 Global Tuberculosis Report 2019
[13] Retrieved from https://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf?ua=1
[14] Cameron, A. C., and Trivedi, P. 2013 Regression Analysis of Count Data. in Second Edition (New York: Cambridge University Press)
[15] McCullagh, P., and Nelder, J. A. 1989 Generalized Linear Models Second Edition (New York: Chapman and Hall)
[16] Caecilia Bintang Girik Allo et al 2019 IOP Conf. Ser.: Mater. Sci. Eng. 546 052050