J. Appl. Environ. Biol. Sci., 5(7)76-83, 2015
© 2015, TextRoad Publication
ISSN: 2090-4274
Journal of Applied Environmental and Biological Sciences
www.textroad.com

Parameter Estimation Model Multivariate Adaptive Regression Splines Data Series In The number of HIV / AIDS in EAST JAVA

Herlina Jusuf[1*], Bambang Widjanarko Otok[2]

[1]Faculty Science Health and Sportmanship, University Country of Gorontalo, Gorontalo,Indonesia
[2]Laboratory of Environmental and Health Statistic, Sepuluh Nopember Institute of Technology (ITS), Surabaya, Indonesia

Received: March 9, 2015
Accepted: May 22, 2015

ABSTRACT

Modeling Multivariate Adaptive Regression Splines (MARS) often uses cross-sectional data because it involves a predictor variable. If it involves the data series, then instead of predictor variables a significant lags of the response variable.Predictor variables determination obtained of the optimal input through the identification of ACF and PACF. The purpose of this study to obtain estimates of the MARS model parameters for the data series. The results show if $\mathbf{B}$ *non singular matrix and the smoothing parameter* $\delta^2 = 0$ *then the estimator of* $\hat{\underset{\sim}{\theta}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T Y$ *, for* $\delta^2 > 0$ *then the estimator of* $\hat{\underset{\sim}{\theta}}$ *is* $\hat{\underset{\sim}{\theta}} = \left(\mathbf{B}^T\mathbf{B} + \delta^2 D\right)^{-1}\mathbf{B}^T\tilde{Y}$.

With

$$\mathbf{B} = \begin{pmatrix} 1 & \prod_{k=1}^{K_1} s_{1m}\left(y_{(t-1)(1,m)} - \xi_{1m}\right) & \cdots & \prod_{k=1}^{K_M} s_{Mm}\left(y_{(t-1)(M,m)} - \xi_{Mm}\right) \\ 1 & \prod_{k=1}^{K_1} s_{1m}\left(y_{(t-2)(1,m)} - \xi_{1m}\right) & \cdots & \prod_{k=1}^{K_M} s_{Mm}\left(y_{(t-2)(M,m)} - \xi_{Mm}\right) \\ \vdots & \vdots & \vdots\!\vdots & \vdots \\ 1 & \prod_{k=1}^{K_1} s_{1m}\left(y_{(t-p)(1,m)} - \xi_{1m}\right) & \cdots & \prod_{k=1}^{K_M} s_{Mm}\left(y_{(t-p)(M,m)} - \xi_{Mm}\right) \end{pmatrix}$$

KEYWORDS:*MARS, HIV-AIDS, GCV, PACF*

1. INTRODUCTION

Modeling of time series data is usually on the condition of the data with stationary and linear fluctuations. Autoregressive Integrated Moving Average (ARIMA) is very effective and will be obtained models with a high degree of conformity if the above preconditions are met. If the data does not meet the conditions of the assumption of the model is obtained be incomplete in describing patterns of behavior that are less precise system used for forecasting. Data with the condition of fluctuations that is not linear and time series modeling to get the value of the next period MARS proposed a new approach that is better than the stochastic model [1] [2][3].

Modeling of time series data using ARIMA and MARS among other research conducted by [4] using multiple seasonal ARIMA, [5] research results turned out better than the MARS method with soft computing models in predicting rain, [6] research on the recession prediction using MARS method gives better results than ARIMA. [1], hybrid modeling using MARS and ARIMA to predict the number of HIV-AIDS.

Election time series forecasting model that generates value that is valid and accurate predictions are able to resolve the major problems in the application of methods of forecasting is so important that it needs to study to get the forecasting model which combines the advantages of MARS with Autoregressive Integrated Moving Average (ARIMA), Expected by the merger of two methods: MARS ARIMA can improve the accuracy of forecasting, especially in forecasting the number HIV-AIDS. [9][11][19]

Hybrid models MARS with ARIMA forecasting combined because each method has advantages and disadvantages in analyzing the data, which can be obtained from the best model with an error rate (error) the smallest in predicting a case [8]. [7] That the forecasting accuracy can be improved by combining multiple linear combination model rather than choosing a solution based on the best model.

This problem would require a scientific study to obtain forecasting model which combines the advantages of MARS and ARIMA in predicting the number of people with HIV-AIDS.

***Corresponding author:**Herlina Jusuf, Faculty Science Health and Sportmanship, University Country of Gorontalo, Gorontalo,Indonesia. Emails:herlina_jusuf@yahoo.co.id,

2. METHODOLOGY

Theoretical study starts from the determination of the model MARS series of data, determining the response and predictor variables, obtaining a model parameter estimator approach Penalized least square and least square. While the applied study, the data obtained through observation / survey is secondary data the number of HIV-AIDS from East Java Provincial Health Office. Once the data is acquired Pre-Processing is done, then the data is divided into two parts: a data-in-sample and out-sample of data. In-sample data is the data that is used to construct the best model, ARIMA and MARS. While the out of sample of data is the data that is used to validate / testing the best models derived from the data in-sample. Predictor variables is the number of AIDS patients in the period prior to that derived from determining the optimal input through PACF plot. The response variable (dependent variable) is the number of AIDS patients in 2000-2011. In-sample data is number of AIDS patients from 2000 to 2010 (120 months). Data out-sample is the number of AIDS patients in 2011. [1].

3. RESULTS AND DISCUSSION

3.1 PARAMETER MODEL ESTIMATION

The prediction MARS Model is a model MARS on a continuous response variable. MARS is a complex combination of spline and recursive partitioning. Spline regression model gives an equation form that represents the piecewise polynomial parametric form. The basic idea of parametric modeling piecewise (divided into several regions) this is a function $f(y)$ that is approached by some parametric function (usually a low-order polynomial) defined in each region in the domain D. Each region is separated by dots knots, and parametric functions defined in each region commonly referred to as basis functions. Knots is the end of a beginning for the region and other regions. Spline regression modeling is implemented by formingset of basis functions that can reach spline approach to order-q and estimate the coefficients of the basis functions using a least-squares.[10][13][16]

MARS models produce a continuous model in knots with basis functions:[14][15][16]

$$B_m^{(q)}(y_t) = \prod_{k=1}^{K_m} \left[s_{km} \cdot (y_{(t-p)(k,m)} - \xi_{km}) \right]_+^q$$

Estimation of the regression curve $f(y)$ generally obtained through Penalized least-squares (PLS) which minimizes the following equation [18]:

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - f\left(y_i\right)\right)^2 + \delta^2 \int_a^b \left(f^{(m)}\left(u\right)\right)^2 du \qquad (1)$$

After the modification of the model RPR and combined with the splines, estimators MARS predictive models for the data series can be written as follows:

$$\hat{f}(y_t) = \theta_0 + \sum_{m=1}^{M} \theta_m \prod_{k=1}^{K_m} [s_{km}(y_{(t-p)(k,m)} - \xi_{km})] \qquad (2)$$

with

θ_0 = the parent basis function

θ_m = coefficients of the basis functions to-m

M = The maximum basis function (nonconstant basis functions)

K_m = the degree of interaction

s_{km} = value±1

$y_{(t-p)(k,m)}$ = independent variable

ξ_{km} = knots value of the independent variable $y_{(t-p)(k,m)}$

By using the estimator MARS, the MARS models presented as follows:

$$y_i = \theta_0 + \sum_{m=1}^{M} \theta_m \prod_{k=1}^{K_m} [s_{km}(y_{(t-p)(k,m)} - \xi_{km})] + \varepsilon_i$$

In matrix form can be written into:

$$\underset{\sim}{Y} = \mathbf{B}\underset{\sim}{\theta} + \underset{\sim}{\boldsymbol{\varepsilon}} \qquad (3)$$

where,

$\underset{\sim}{Y}$ = response variable,

$\mathbf{B} = [1,\ (y_{(t-p)(k,m)} - \xi_{km})_1^K]$ = basis functions,

$\underset{\sim}{\theta}$ = coefficients of basis functions

$\underset{\sim}{\varepsilon}$ = error

with,

$$\tilde{Y} = (y_1, y_2, \ldots, y_n)^T \text{ and } \underset{\sim}{\theta} = (\theta_0, \theta_1, \ldots, \theta_M)^T$$

$$\mathbf{B} = \begin{pmatrix} 1 & \prod_{k=1}^{K_1} s_{1m}(y_{(t-1)(1,m)} - \xi_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(y_{(t-1)(M,m)} - \xi_{Mm}) \\ 1 & \prod_{k=1}^{K_1} s_{1m}(y_{(t-2)(1,m)} - \xi_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(y_{(t-2)(M,m)} - \xi_{Mm}) \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 1 & \prod_{k=1}^{K_1} s_{1m}(y_{(t-p)(1,m)} - \xi_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(y_{(t-p)(M,m)} - \xi_{Mm}) \end{pmatrix}$$

To obtain the estimator $\underset{\sim}{\theta}$ through Penalized least-squares is to minimize equation (1) and use the following assumptions.

Assumption 1: *If* $(Y_{(t-p)i}, \xi_{km})$ *the points on the line is fixed, and there is a continuous function* $h_j(.)$ *defined on [0,1] such that each element* $Y_{(t-p)i}$ *meet,* $Y_{(t-p)i} = h_j(\xi_i) + u_{ij}$, $1 \le i \le n$, $1 \le j \le p$ *with* $\{u_{ij}\}$ *is a real lineup that meets,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} u_i^T u_i = \mathbf{C} \text{ and, } \limsup_{n\to\infty} \frac{1}{a_n} \max_{1\le k\le n} \left| \sum_{i=1}^{n} u_{ij} \right| < \infty$$

with, $u_{ij} = (u_{i1}, \cdots, u_{ip})^T$, $a_n = n^{1/2} \log n$, *C = positive definite matrix of order p*

Theorem1. *If the assumption 1, is met,* $\mathbf{B}$ *non singular matrix and the smoothing parameter* $\delta^2 = 0$ *then estimator of* $\hat{\underset{\sim}{\theta}}$ *is to minimize the value* Z, *with which* $Z = (\tilde{Y} - \mathbf{B}\underset{\sim}{\theta})^T (\tilde{Y} - \mathbf{B}\underset{\sim}{\theta})$ *namely* $\hat{\underset{\sim}{\theta}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T Y$

Proof:

Note the equation (1):

$$\frac{1}{n} \sum_{i=1}^{n} \left(y_i - f(y_i) \right)^2 + \delta^2 \int_a^b \left(f^{(m)}(u) \right)^2 du$$

with $\delta^2 = 0$, the equation becomes:

$$\frac{1}{n} \sum_{i=1}^{n} \left(y_i - f(y_i) \right)^2 \qquad (4)$$

From equation (3), then $\hat{f}(y_i) = \mathbf{B}\underset{\sim}{\theta}$, so that the equation (4) becomes:

$$\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \mathbf{B}\underset{\sim}{\theta} \right)^2 = (\tilde{Y} - \mathbf{B}\underset{\sim}{\theta})^{\mathrm{T}} (\tilde{Y} - \mathbf{B}\underset{\sim}{\theta}) = Z$$

To obtain estimator $\hat{\underset{\sim}{\theta}}$ used a least squared method, which in principle minimizes Z, stated as follows:

$$Z = (\tilde{Y} - \mathbf{B}\underset{\sim}{\theta})^{\mathrm{T}} (\tilde{Y} - \mathbf{B}\underset{\sim}{\theta})$$

$$Z = (\tilde{Y}^T\tilde{Y} - \underset{\sim}{\theta}^T\mathbf{B}^{\mathrm{T}}\tilde{Y} - \tilde{Y}^T\mathbf{B}\underset{\sim}{\theta} + \underset{\sim}{\theta}^{\mathbf{T}}\mathbf{B}^{\mathbf{T}}\mathbf{B}\underset{\sim}{\theta})$$

$$Z = (\tilde{Y}^T\tilde{Y} - 2\underset{\sim}{\theta}^T\mathbf{B}^T\tilde{Y} + \underset{\sim}{\theta}^T\mathbf{B}^T\mathbf{B}\underset{\sim}{\theta})$$

To obtain the normal equation is done by lowering the partially $\underset{\sim}{\theta}$ on the results as follows:

$$\frac{\partial Z}{\partial \underset{\sim}{\theta}} = -2\mathbf{B}^T\tilde{Y} + 2\mathbf{B}^T\mathbf{B}\underset{\sim}{\theta} = 0$$

$$-\mathbf{B}^T\tilde{Y} + \mathbf{B}^T\mathbf{B}\hat{\underset{\sim}{\theta}} = 0$$

$$\mathbf{B}^T\mathbf{B}\hat{\underset{\sim}{\theta}} = \mathbf{B}^T\tilde{Y} \quad (5)$$

because $\mathbf{B}$ *non singular matrix,* then

$$\hat{\underset{\sim}{\theta}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\tilde{Y} \quad (6)$$

So it proved Theorem 1.

MARS parameters in addition to the base amount of the parent function, can be obtained by substituting equation (6) into the equation (3) as follows:

$$\begin{aligned}\tilde{Y} &= \mathbf{B}[(\mathbf{B}^T\mathbf{B})^{-1}(\mathbf{B}^T\tilde{Y})] \\ &= \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\tilde{Y} \\ &= \mathbf{P}\tilde{Y}\end{aligned} \quad (7)$$

Thus, $\mathbf{P} = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ the size $(M+1)\times(M+1)$. Because $\mathbf{P}$ is the symmetric matrix and idempotent then Trace of a matrix $\mathbf{P}$ equal to the rank of $\mathbf{P}$ which is the number of basis functions in addition to the constant parameter and the number of parameters to be estimated, denoted as: $C(\tilde{M}) = \text{Trace}(\mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T) + 1$.

Theorem2. *If the rank of* $\mathbf{P}$ *is (M+1) then* $\mathbf{P}$ *has (M+1) unit eigenvalues are not zero and as many as [N-(M+1)] unit eigenvalues are zero.*

Proof:

Suppose λ are eigen values and Γ eigenvectors, then:

$$\mathbf{P}\Gamma = \lambda\Gamma \qquad (\Gamma \neq 0)$$

if multiplied Γ^T become $\Gamma^T\mathbf{P}\Gamma = \lambda\Gamma^T\Gamma$ and $\mathbf{P}$ is idempotent matrix ($\mathbf{P} = \mathbf{P}^2$) then $\Gamma^T\mathbf{P}^T\mathbf{P}\Gamma = \lambda\Gamma^T\Gamma$, so that $(\lambda\Gamma)^T\lambda\Gamma = \lambda\Gamma^T\Gamma$ or $\lambda(\lambda-1)\Gamma^T\Gamma = 0$. This shows that the eigen values of $\mathbf{P}$ valuable 1 as much as (M+1) and the value 0 as [N-(M+1)], so that the rank of the matrix $\mathbf{P}$ is equal to (M+1).

Theorem 3. *If the assumption 5.1.1. fulfilled,* $\mathbf{B}$ *non singular matrix and the smoothing parameter* $\delta^2 > 0$ *then the estimator of* $\hat{\underset{\sim}{\theta}}$ *is to minimize the value of* $ASR(\underset{\sim}{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(\underset{\sim}{x}_i, \underset{\sim}{\theta}))^2 + \delta^2\int|\nabla f(\underset{\sim}{x}_i, \underset{\sim}{\theta})|^2 d\underset{\sim}{y}$ is

$$\hat{\underset{\sim}{\theta}} = \left(\mathbf{B}^T\mathbf{B} + \delta^2 D\right)^{-1}\mathbf{B}^T\tilde{Y}$$

Proof:

If the base function in equation (3) bringing to an average value of zero, $\mathbf{B}^T\mathbf{B}$ proportional to the following the covariance matrix basis functions:

$$V^*\underset{\sim}{\theta} = c \quad (8)$$

with,

$$V_{ij} = \sum_{t=1}^{N} B_j(y_{t-p})[B_i(y_{t-p}) - \overline{B}_i] \quad (9)$$

$$c_i = \sum_{t=1}^{N}(y_t - \overline{y})B_i(y_{t-p}) \quad (10)$$

$\overline{B}_i$ and $\overline{y}$ the average of all the data. Equation (9) were completed for each location knots ξ, for each variable υ, for all the basis functions m and for all interactions M.

Every new basisfunction, enter into rows and columns R and V, so that the completion of the normal equation (5), updated with Chelosky Decomposition with the following modifications,

$$(V+\varepsilon\mathbf{D})\theta = c \tag{11}$$

where $\mathbf{D}$ is a diagonal matrix of size (M+1)(M+1) of the elements diagonal matrixof the matrix V.

$$ASR + RP \tag{12}$$

or

$$\frac{1}{N}\sum_{t=1}^{N}(y_t - f(y_{t-p}, \underset{\sim}{\theta}))^2 + \delta^2 \int |\nabla f(y_{t-p}, \underset{\sim}{\theta})|^2 dy \tag{13}$$

Consider, $f(y_t, \underset{\sim}{\theta}) = \sum_{j=1}^{m} \underset{\sim}{\theta}_j \mathbf{B}_t(y)$ and a roughness penalty (RP), $RP = \underset{\sim}{\theta}^T R \underset{\sim}{\theta}$ is a positive definite quadratic function of the parameters, with $\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_m)^T$, and R is the matrix size $m \times m$, stated as follows:

$$R_{jk} = \delta^2 \frac{1}{N}\sum_{t=1}^{N}\sum_{d=1}^{n} \frac{\partial B_j(y_t)}{\partial y_d}\frac{\partial B_k(y_t)}{\partial y_d} \tag{14}$$

So that the equation (5.1.12) can be expressed in the form of a matrix, namely:

$$\begin{aligned} ASR + RP &= \frac{1}{N}|\tilde{Y} - \mathbf{B}\underset{\sim}{\theta}|^2 + \underset{\sim}{\theta}^T R \underset{\sim}{\theta} \\ &= \frac{1}{N}|\tilde{Y}|^2 - \frac{1}{N}2\mathbf{B}\underset{\sim}{\theta}\tilde{Y} - \frac{1}{N}\underset{\sim}{\theta}^T \mathbf{B}^T \mathbf{B}\underset{\sim}{\theta} + \underset{\sim}{\theta}^T R \underset{\sim}{\theta} \end{aligned} \tag{15}$$

Suppose $V = \frac{1}{N}\mathbf{B}^T\mathbf{B}$ dan $c = \frac{1}{N}\mathbf{B}^T\tilde{Y}$ where $\tilde{Y} = (y_1, \cdots, y_N)^T$. It is assumed that the sample average of variable response after deduction is, $\bar{\mathbf{y}} = \frac{1}{N}\sum_{t=1}^{n} y_t = 0$ and the sample variance of y is $\hat{\sigma}_y^2 = \frac{1}{N}|y|^2$. Then the equation (15) can be expressed as,

$$\begin{aligned} ASR + RP &= \hat{\sigma}_y^2 + \underset{\sim}{\theta}^T V \underset{\sim}{\theta} - 2c^T\underset{\sim}{\theta} + \underset{\sim}{\theta}^T R \underset{\sim}{\theta} \\ &= \hat{\sigma}_y^2 + \underset{\sim}{\theta}^T \underset{\sim}{\theta}(V + R) - 2c^T\underset{\sim}{\theta} \\ &= \hat{\sigma}_y^2 + \underset{\sim}{\theta}^T V^* \underset{\sim}{\theta} - 2c^T\underset{\sim}{\theta} \\ &= \hat{\sigma}_y^2 - 2\underset{\sim}{\theta}(c^T - \underset{\sim}{\theta}^T V^*) \end{aligned} \tag{16}$$

or,

$$\begin{aligned} ASR + RP &= \tfrac{1}{N}\sum_{k=1}^{N}(y_k - \bar{y})^2 - \tfrac{1}{N}\sum_{i=1}^{M+1}\theta_i(2c^T - \delta^2 D_{ii}\theta_i) \\ &= \tfrac{1}{N}\left[\sum_{k=1}^{N}(y_k - \bar{y})^2 - \sum_{i=1}^{M+1}\theta_i(2c^T - \delta^2 D_{ii}\theta_i)\right] \end{aligned} \tag{17}$$

Suppose $ASR(\underset{\sim}{\theta}) = ASR + RP$, then the coefficients of the basis functions $\hat{\underset{\sim}{\theta}}$ obtained by minimizing the equation (16) or (17).

$ASR(\underset{\sim}{\theta})$ minimum order then

$$\frac{\partial ASR(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = 0,$$

$$\frac{\partial ASR(\underset{\sim}{\theta})}{\partial \hat{\underset{\sim}{\theta}}} = \frac{\partial}{\partial \hat{\underset{\sim}{\theta}}}(\tilde{Y}^T\tilde{Y} + \hat{\underset{\sim}{\theta}}^T V^* \hat{\underset{\sim}{\theta}} - 2c^T\hat{\underset{\sim}{\theta}}) = 2V^*\hat{\underset{\sim}{\theta}} - 2c^T = 0$$

So that,

$$2V^*\hat{\underset{\sim}{\theta}} = 2\mathbf{c}$$
$$V^*\hat{\underset{\sim}{\theta}} = \mathbf{c}$$
$$(V+R)\hat{\underset{\sim}{\theta}} = \mathbf{c} \quad (18)$$
$$\left(\frac{1}{N}\mathbf{B}^T\mathbf{B} + \frac{1}{N}\delta^2 D\right)\hat{\underset{\sim}{\theta}} = \frac{1}{N}\mathbf{B}^T\tilde{Y}$$
$$\frac{1}{N}\left(\mathbf{B}^T\mathbf{B} + \delta^2 D\right)\hat{\underset{\sim}{\theta}} = \frac{1}{N}\mathbf{B}^T\tilde{Y}$$
$$\hat{\underset{\sim}{\theta}} = \left(\mathbf{B}^T\mathbf{B} + \delta^2 D\right)^{-1}\mathbf{B}^T\tilde{Y} \quad (19)$$

where $\mathbf{D}$ is a positive definite diagonal matrix of $\mathbf{R}$.
Completion is given in equation (19) is to minimize $ASR(\underset{\sim}{\theta})$ and an estimator MARS. So Theorem 3 proved.

3.2. Comparison MARS, MARS Hybrid ARIMA and ARIMA models

[1] Comparison of the GCV and the value of $R^2$and RMSEP on each ARIMA models, MARS and MARS HYBRID ARIMA results are presented in the following table.

Table 3. Criteria for the Best Model Based on MSE

Model	Data			
	In Sample		Out Sample	
	MSE	RMSE	MSE	RMSE
MARS	10.36	3.218	20.29	4.505
MARS HYBRID ARIMA	1.01	1.005	18.33	4.282
ARIMA	22.55	4.748	31.47	5.610

MSE values in Table 3 shows that the in-sample data which has the smallest MSE value is HYBRID MARS ARIMA models. So that the model is the best model to predict the number of HIV-AIDS in the province of East Java, and are presented in the following figure.

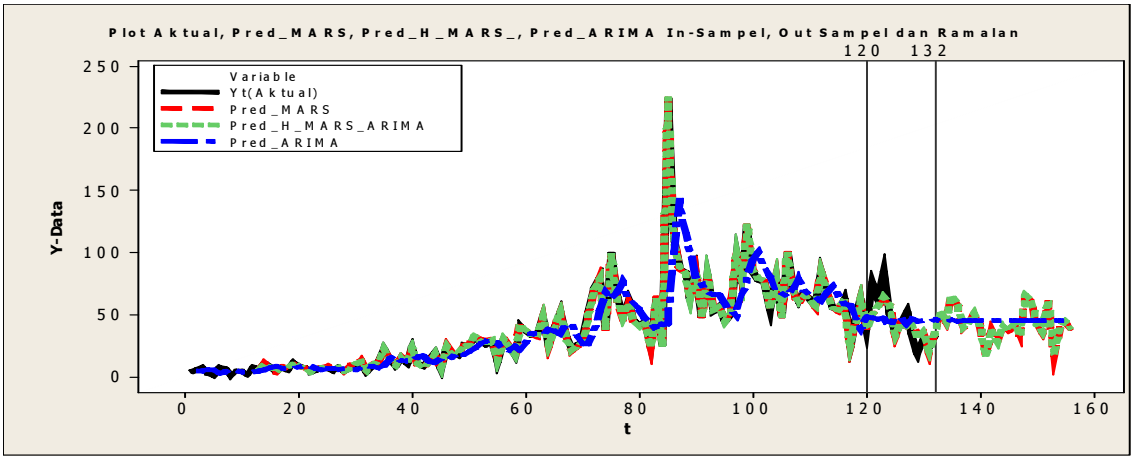

Figure 5
Plot the data in-sample Actual, MARS and MARS ARIMA, ARIMA
On the number of HIV-AIDS in East Java

4. CONCLUSION

MARS model parameter estimators of data series, the first decisive step further predictor variable were estimated by the least squares method has resulted if $\mathbf{B}$ non singular matrix and the smoothing parameter $\delta^2 = 0$ then the estimator of $\hat{\underset{\sim}{\theta}}$ $\hat{\underset{\sim}{\theta}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T Y$, for $\delta^2 > 0$ then the estimator of $\hat{\underset{\sim}{\theta}}$ is $\hat{\underset{\sim}{\theta}} = \left(\mathbf{B}^T\mathbf{B} + \delta^2 D\right)^{-1}\mathbf{B}^T\tilde{Y}$.
With

$$\mathbf{B}=\begin{pmatrix} 1 & \prod_{k=1}^{K_1} s_{1m}(y_{(t-1)(1,m)}-\xi_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(y_{(t-1)(M,m)}-\xi_{Mm}) \\ 1 & \prod_{k=1}^{K_1} s_{1m}(y_{(t-2)(1,m)}-\xi_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(y_{(t-2)(M,m)}-\xi_{Mm}) \\ \vdots & \vdots & \vdots\vdots & \vdots \\ 1 & \prod_{k=1}^{K_1} s_{1m}(y_{(t-p)(1,m)}-\xi_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(y_{(t-p)(M,m)}-\xi_{Mm}) \end{pmatrix}$$

[1] Estimation of parameters on the data the number of HIV / AIDS consists of four (4) phases, namely the determination of the optimal input, modeling MARS, MARS with ARIMA modeling error, MARS Hybrid ARIMA and ARIMA Selection of the best model, MARS and MARS hybrid ARIMA out-sample data based RMSE. Lag- lag significantly to PACF an optimal input in the modeling MARS, because it produces the smallest GCV value. Prediction of the number of HIV-AIDS sufferers are affected by the amount of one, seven and twelve previous period. HIV-AIDS predictions with ARIMA modeling MARS provides out-sample MSE value of 18.33 is better than the MARS and ARIMA modeling respectively by 20.29 and 31.47.

REFERENCES

[1] *H. Jusuf, Bambang W. Otok.* Optimal input of data series in predicted the number patient of HIV-AIDS in East Java province using multivariate adaptive regression splines*. Journal of Health Sport and Tourism 2014; 5(1), 35-42.* ***DOI:*** *10.7813/jhst.2014/5-1/6*

[2] Friedman,J.H. (1991),Multivariate Adaptive Regression Splines (with discussion)*.Ann. Statist*. 19:1–141.

[3] Buja, A., Duffy, D., Hastie, T. and Tibshirani, R. (2001), *Multivariate Adaptive Regression Splines* (with discussion).

[4] Taylor,J.W., Menezes,L.M and McSharry,P.E (2006),*A Comparison of Univariate methods for forecasting electricity demand up to a day ahead, 22 1-16*

[5] Abraham, A. (2002),*Analysis of Hybrid Soft and Hard Computing Techniques for Forex Monitoring Systems*. School of Business Systems, Monash University, Clayton, Victoria, 3800, Australia.

[6] Sephton, P.(2001), *Forecasting Recessions: Can We Do Better on MARS*. The Federal Reserve Bank of St. Louis

[7] Bates, J.M., Granger,C.W.J (1969),The combination of forecasts*. Operational Research* Quarterly 20, 451-468

[8] Abraham, A. and Steinberg, D. (2001,*MARS: Still an Alien Planet in Soft Computing?*. School of Computing and Information Technology, Salford System.Inc, USA.

[9] Ong'ech. J and Maranga, O(2007),*Integrated HIV/AIDS Prevention, Treatment And Care*. Kenya: African Medical Research Foundation (AMREF)

[10] Otok, B. W. (2008), *Pendekatan Bootstrap pada Model Multivariate Adaptive Regression Splines* (MARS), Disertasi, Universitas Gadjah Mada,Yogyakarta.

[11] Richard, P and Peter Aggleton (2009),*HIV and AIDS-related stigma and discrimination*: a Conceptual Framework and Implications for Action*Social Science & Medicine*, Volume 57, Issue 1, Pages 13-24

[12] Rhodes,T., Merrill Singer, Philippe Bourgois, Samuel R. Friedman, Steffanie A. Strathdee (2009), *The social structural production of HIV risk among injecting drug users Social Science & Medicine*, Volume 61, Issue 5, Pages 1026-1044

[13] Breiman, L. (1991), *Discussion* of *Multivariate Adaptive Regression Splines*, by J.H. Freidman, *Annals of Statistics*, Vol. 19, 82-90.

[14] Box, G.E.P., Jenkins, G.M., and Reinsel. G.C., (1994), *Time Series Analysis Forecasting and Control*, 3^{d} edition, Englewood Cliffs : Prentice Hall.

[15] Chattfield, C. (1997), *Time Series, Theory and Practice and Forecasting*, Chapman Hall, London.

[16] Friedman, J.H. (1990),*Estimating Functions Of Mixed Ordinal And Categorical Variables Using Multivariate Adaptive Regression Splines*. Technical Report LCS 107, Statistics Department, Stanford University.

[17] Friedman, J.H. (1990),*Estimating functions of mixed ordinal and categorical variables using Multivariate Adaptive Regression Splines*. Technical Report LCS 107, Statistics Department, Stanford University.

[18] Lewis, P.A.W ; J.G. Stevens. (1991), Nonlinier Modeling of time Series Using Multivariate Adaptive Regression Splines (MARS). *Journal of the American Statistical Assosation*, 86, No. 416.(Dec.,1991), pp.864-877.

[19] Musoke, R (2007), *HIV/AIDS in Children.* African Medical Research Foundation (AMREF) Allan and Nesta Ferguson Trust : Kenya