Modelling of sea surface temperature by using generalized additive mixed models in risk detection

R A Humaira¹, Miftahuddin¹, and I Setiawan²

¹Department of Statistics, Faculty of Mathematics and Science, Syiah Kuala University
²Department of Marine Science, Faculty of Marine and Fisheries, Syiah Kuala University, Banda Aceh, Indonesia

Abstract. Rising temperatures from year by year caused an extreme weather which could result in all kinds of natural disasters. One of methods that can analyze the factors which affect sea surface temperature (SST) is Generalized Additive Mixed Model (GAMM). The GAMM is able to analyze more complex, especially those related to random effect and data which have not normal distribution. The purposes of this study is to construct a model of SST dataset by interacting with several variables, then the factors that affect SST be discovered. The type of data that used in this research is secondary data obtained through the National Oceanic and Atmospheric Administration website. The data has been taken from 2 locations, namely points 4°N90°E and 1.5°N90°E. The data used is daily unitin periods from September 17, 2006 to June 14, 2017. The variables used are SST, wind speed, air temperature, dynamic height, heat content, rainfall, relative humidity, shortwave radiation, sea surface density, and sea surface salinity, and a random effect namely location. Based on the results obtained the best model is the third model with Akaike Information Criterion, Bayesian Information Criterion and log-likelihood values respectively are -986,021, -894,911, and 515,752. Then the factors that significantly affect SST are wind speed (X₁), air temperature (X₂), dynamic height (X₃), heat content (X₄), sea surface density (X₅), and sea surface salinity (X₆). Furthermore, there are several interacting predictor variables that affect SST including interactions between air temperature and dynamic height, dynamic height and sea surface density, wind speed and rainfall, and air temperature with sea surface density.

1. Introduction
Climate change has a significant impact on Indonesia region. The effects of climate change can be felt by people who living in rural, urban and coastal areas. Rising temperatures from year by year caused an extreme weather which could result in all kinds of natural disasters. According to Meteorology, Climatology and Geophysics [1], in the last few years, it’s been five times that happened tropical cyclone in indonesia.

The other effects that caused by climate change are drought, forest fires and storms that occurred on the high seas. In 2017, one of the districts in Aceh Province, namely Nagan Raya Regency, experienced a peat fires and West Aceh Regency experienced a forest fires, sixty five-acre lot. Not only forest fires, on February 2018 there are thousands of acres of rice fields in several districts in Aceh Province experienced drought thus effect to rice plants being threatened corp failure. There are 8 districts that experienced drought on paddy fields, namely Southeast Aceh, Aceh Tamiang, Lhokseumawe, Langsa, North Aceh, Aceh Besar, Banda Aceh, and Nagan Raya. The most extensive
area of wetland that experienced drought is North Aceh Regency, 2,002 acres and Aceh Besar Regency as vast as 1,882 acres. Then some time ago there was a storm in the waters of the Indian Ocean and The Malacca Strait in the northern tip of Aceh which caused a high level of concern. This is because of two boats have sunk on the Banda Aceh - Sabang route and in the waters of North Aceh [2].

There are several parameters that affect the climate system on the Earth, especially for regions with tropical climates that are very influential in determining the relationship between the atmosphere and oceans that affect climate change. One of the interactions between the atmosphere and the ocean is the interaction between SST and rainfall. The sea is affected by surface currents, solar heat, cloud conditions, and upwelling especially in estuaries and along coastlines. In addition, several meteorological factors also affect SST, such as rainfall, evaporation, humidity, air temperature, wind speed, and the intensity of solar radiation.

According to previous research on SST in the Indian Ocean region at 4°N90°E, [3] the data period used from 2011 to 2015 using Generalized Additive Models method. Based on the optimum model obtained a few significant variables namely wind speed, SSS, shortwave radiation, salinity, humidity, prec. rain, dynamic height, density, temperature water, isoterm, conductivity, current velocity, zonal, and merid.

Additive models are comprehensive, the meaning is able to analyze more complex, especially those related to random effect and data that have not normal distribution. Generalized Additive Model (GAM) is an extension of the ordinary linear regression model but replaces the linear function into an additive function, thus it can be used in cases where there is a nonlinear relationship between the response variable and the predictor variable. Analysis by using the GAM method is not only carried out on the response variable which have not normal distribution, but also it can be used in distributions that are included in the exponential family. However, GAM cannot be used if there are two effects in a model, namely random effects and fixed effects. This study uses a random effect namely location. There are 2 locations that are used as a random effect, these are location A which is the point of data retrieval point 4°N90°E and location B is point 1.5°N90°E. Therefore the Generalized Additive Mixed Model (GAMM) approach is used to detect both effects to analyze SST data by interacting with several variables such as rainfall, wind speed, air humidity, and etc.

2. Literature Review

2.1. Generalized Additive Model
[4] The additive model is the development of a linear model in which the sum predictor component of the smoothing function. Suppose there are a number of data \( \{y_i, x_{i1}, x_{i2}, ..., x_{ip}\}_{i=1}^{n} \) where \( n \) is the number of observations. The form of the additive model can be explained as follows.

\[
\begin{align*}
\mathbf{y} &= \mathbf{X}'\boldsymbol{\alpha} + \sum_{j=1}^{p} f_j(x_j) + \epsilon \\
\end{align*}
\]

\( y \) = vector of response variable

\( X' \) = tranpose of fixed effects matrix, start the observation to the \( i \)

\( \alpha \) = vector of fixed effects coefficients

\( f_j(x_j) \) = smoothing function of covariate \( x_j \)

\( x_j \) = predictor variable to the \( j \)

\( \epsilon \) = random error

[5] The generalized Additive Model (GAM) is an extension of the additive model which assumes that the mean of the response variable depends on the additive predictor variable. GAM also assumes that the response variable has an exponential distribution. GAM is also an extension model of Generalized
Linear Model (GLM) by replacing the linear predictor function $\sum_{j=0}^p x_j \beta_j$ to be an additive predictor function of $\sum_{j=0}^p f_j(x_j)$. The general form of the GAM model is defined as follows:

$$g(\mu) = X^T \alpha + \sum_{j=1}^p f_j(x_j)$$

(2)

$g(\mu)$ = link function

$X^T$ = tranpose of fixed effects matrix, start the observation to the $i$

$\alpha$ = vector of fixed effects coefficients

$f_j(x_j)$ = smoothing function of covariate $x_j$

$x_j$ = predictor variable to the $j$

2.2. Generalized Mixed Additive Model (GAMM)

Generalized Additive Mixed Model (GAMM) is used when there is a nonlinear relationship between the response variable and the predictor variable. The use of GAMM for quantitative variable data with the estimated smoothing function used is smoothing spline. GAMM is an extension of the Generalized Linear Mixed Model (GLMM) model where GAMM replaces linear functions in the GLMM to additive functions. The general form of the GAMM model is as follows.

$$g(\mu) = X^T \alpha + \sum_{j=1}^p f_j(x_j) + Zb$$

(3)

$g(\mu)$ = link function

$X^T$ = tranpose of fixed effects matrix, start the observation to the $i$

$\alpha$ = vector of fixed effects coefficients

$f_j(x_j)$ = smoothing function of covariate $x_j$

$x_j$ = predictor variable to the $j$

$Z$ = tranpose of random effects matrix, start the observation to the $i$

$b$ = vector of random effects coefficients

Generalized Additive Mixed Model (GAMM) is an extension of the GAM and GLMM methods, only replacing linear functions into additive functions. The GAMM model is able to identify and explain the effect of random components more effectively and efficiently in a model. The GAMM method includes random effects and fixed effects in the model [5]. Random effects are randomized individual-experimental retrieval of a unit from the population, while the fixed effect (fixed effect) are all related parameters taken from the population in an experiment. A model that is influenced by a fixed effect and a random effect is called a mixed effect [6].

2.3. Link function

[7] The link function $g(\cdot)$ is a function that connects a linear predictor $\eta$ with the expectation value of the response $y$, which is $\mu$. The link function or the inverse contact function is determined by the variable category of the response variable as seen in the table 1.

| Response Variable | Distribution | Link function | Reverse connection function |
|-------------------|--------------|---------------|----------------------------|
| Continuous        | Normal       | Identity      | $\eta$                     |
|                   | Gamma        | Invers        | $1/\eta$                   |
|                   |              | Log           | $e^\eta$                   |
|                   | Invers       | $1$           | $1/\theta^2$               |
|                   | Gaussian     | $\mu^2$       | $1/\theta^2$               |
| Proportion        | Binomial     | Logit         | $e^\eta = (1 + e^\eta)$    |

Table 1. Connect function of several distributions
2.4. **GAMM parameter estimator**

The estimation is the estimated value of the population parameter based on the available data, to estimate the parameters of GAMM, the first step is to describe the probability density function of the exponential family as the response variable as follows:

\[
(y; \theta) = \exp[a(y)b(\theta) + c(\theta) + d(y)]
\]  

(4)

GAMM uses the maximum likelihood estimator (MLE) approach to obtain the \( \beta \) and estimator parameters \( b \). The estimated value obtained in GAMM is obtained by maximizing the log-likelihood function [6].

2.5. **Significant Parameter Test**

One test that can be used to test the significance of the \( \beta \) coefficient in the model is a partial test. Partial test (T-test) is used to determine the effect of an individual predictor variable on the response variable. The hypothesis used is

\[ H_0 : \beta_j = 0 \]

- The \( j \) predictor variable does not significantly influence the response variable

\[ H_1 : \beta_j \neq 0 \]

- The \( j \) variable predictor has a significant effect on the response variable

Test statistics used for partial test is:

\[
t_{\text{calculate}} = \frac{\hat{\beta}_n}{SE(\hat{\beta}_n)}
\]  

(5)

2.6. **Model Selection**

2.6.1. **Akaike’s Information Criterion (AIC)**

[8] Choosing the goodness of a model used a criterion in measuring it. One type of criteria used is AIC.

\[
AIC = -2LL + (2 \times K)
\]  

(6)

where, LL is the Likelihood log model and \( K \) is the number of parameters estimated in the model. The model with the smallest AIC value is the chosen model.

2.6.2. **Schwarz’s Bayesian Information Criterion (BIC)**

BIC is an asymptotic result derived from the assumption that the distribution of data in exponential families. Formula in calculating the BIC value is

\[
-2 \cdot \ln p(x|k) \approx BIC = -2 \cdot \ln L + k \ln(n)
\]  

(7)

to determining the optimal model, the model chosen is the model that has the smallest BIC value. This is because BIC improves the function of \( k \). That is, variations that cannot be explained in the response variable and the number of predictor variables increase the BIC value. Therefore, small BIC values explain that fewer predictor variables, more appropriate or both [9].

3. **Research methodology**

The type of data used in this study is secondary data obtained through the National Oceanic and Atmospheric Administration website [10] at points 4°N90°E and 1.5°N90°E. Data retrieval is through the points 4°N90°E and 1.5°N90°E because this point is closest to Aceh Province. The data used is daily unit in periods from September 17, 2006 to June 14, 2017.
4. Results and Discussion
In this study, before carrying out further analysis, a descriptive statistical analysis was first carried out to obtain an overview of the data as shown in the Table 3 and Table 4.

| Table 2. Summary statistic data point 4°N90°E (location A) |
|----------------------|---------|---------|---------|---------|---------|---------|
| Variabel             | Min     | Kuartil 1 | Median   | Mean    | Kuartil III | Max     | Range   |
| Y                    | 28,57   | 29,05    | 29,27   | 29,39   | 29,57       | 30,99   | 2,42    |
| X1                   | 0,40    | 3,50     | 5,10    | 4,98    | 6,52        | 11,00   | 10,60   |
| X2                   | 26,75   | 28,17    | 28,62   | 28,57   | 29,01       | 29,94   | 3,19    |
| X3                   | 104,00  | 116,50   | 121,30  | 121,70  | 125,80      | 140,80  | 36,80   |
| X4                   | 35,08   | 35,24    | 35,32   | 35,32   | 35,38       | 35,58   | 0,50    |
| X5                   | 0,00    | 0,02     | 0,035   | 0,453   | 0,44        | 0,99    | 13,16   |
| X6                   | 73,10   | 78,70    | 80,20   | 80,22   | 81,92       | 87,70   | 14,60   |
| X7                   | 19,75   | 20,57    | 21,02   | 20,98   | 21,23       | 21,63   | 1,88    |
| X8                   | 24,04   | 176,23   | 236,04  | 214,97  | 271,05      | 315,18  | 291,14  |
| X9                   | 3,20    | 3,34     | 5,10    | 4,98    | 6,52        | 11,00   | 7,80    |

| Table 3. Summary statistic data point 1.5°N90°E (location B) |
|----------------------|---------|---------|---------|---------|---------|---------|
| Variabel             | Min     | Kuartil 1 | Median   | Mean    | Kuartil III | Max     | Range   |
| Y                    | 28,06   | 28,78    | 29,03   | 29,14   | 29,41       | 30,60   | 2,54    |
| X1                   | 0,60    | 2,90     | 4,70    | 4,59    | 6,37        | 10,30   | 9,70    |
| X2                   | 26,52   | 27,91    | 28,23   | 28,28   | 28,75       | 30,22   | 3,70    |
| X3                   | 9,99    | 115,19   | 122,51  | 119,75  | 127,70      | 139,37  | 129,38  |
| X4                   | -9,99   | 35,27    | 35,36   | 34,26   | 35,43       | 35,59   | 25,60   |
| X5                   | 0,00    | 0,01     | 0,035   | 0,453   | 0,44        | 0,99    | 9,99    |
| X6                   | 73,80   | 79,62    | 81,40   | 81,68   | 84,05       | 89,20   | 15,40   |
| X7                   | 13,04   | 161,40   | 234,71  | 222,06  | 273,66      | 999,99  | 32,05   |
| X8                   | 9,99    | 21,02    | 21,35   | 21,18   | 21,18       | 21,65   | 11,66   |
| X9                   | 33,05   | 33,66    | 34,14   | 34,07   | 34,40       | 35,23   | 2,18    |

Based on Table 3 for data in location A and Table 4 for the data in location B above, it is known that the minimum, quartile I, median, mean, quartile III, maximum and data range for each variable.
The correlation matrix provides information about relationship patterns, forms of distribution, and correlations between both dependent and independent variables and between independent variables. Figure 1 shows that of the nine variables, only one variable had a symmetrical histogram, namely the $X_5$ (relative humidity). Y (SST), $X_2$ (dynamic height) and $X_3$ (heat content) have a bimodal histogram. For variables $X_1$ (air temperature), $X_6$ (wind speed), and $X_7$ (short wave radiation) have the shape of a histogram sticking to the left. While the variables $X_4$ (rainfall), $X_8$ (sea level density), and $X_9$ (sea surface salinity) are not symmetrical because they have different data that goes up or down.

Based on Figure 1 also known the pattern of the relationship between the response variable (Y) with all predictor variables (X) based on the distribution of data. There are several variables that have a formless distribution pattern so that smoothing is done in the modeling. Then there is one independent variable that has a perfect correlation (strong correlation) that is the correlation between variables $X_2$ and $X_3$ with a correlation value of 1. Besides that there are also several variables that are linearly related and have a strong correlation, variables $X_7$ and $X_6$, $X_3$ and $X_8$, $X_2$ and $X_8$ with a correlation value of 0.89; 0.74, and 0.74. In addition to seeing the pattern of relationships and distribution between response variables and explanatory variables, other requirements of using GAMM method are the response variables included in the exponential family. Some distributions included in the exponential family are Normal, Gamma, Inverse Gaussian, Exponential and so on. Here are visualizations and comparison tables of several distributions of data.

![Normal Distribution](image1)

(a). Normal Distribution

![Eksponential Distribution](image2)

(b). Eksponential Distribution

![Gamma Distribution](image3)

(c). Gamma Distribution

![Inverse Gaussian Distribution](image4)

(d). Inverse Gaussian Distribution

**Figure 1.** Correlation matrix of SST dataset

**Figure 2.** Visualization for the selection of data distribution

**Table 4.** Comparison of distribution

| Distribution    | AIC     |
|-----------------|---------|
| a. Normal       | 629,852 |

Here are visualizations and comparison tables of several distributions of data.
Based on Figure 2, it can be seen that the distributions (a), (c), and (d) have similar forms of distribution visualization. So that, it is difficult to determine the appropriate distribution. Next in Table 5. can be seen the comparison of data distribution based on the AIC value. From Table 5 it is known that the distribution that matches the data is the Gaussian Inverse distribution because it has the smallest AIC value. Then, based on table 1. for the inverse gaussian distribution, the link function that used is $\frac{1}{\mu^2}$.

4.1. Modeling using GAMM

The response variable of sea surface temperature ($Y$) is modeled with nine predictor variables namely wind speed ($X_1$), air temperature ($X_2$), dynamic height ($X_3$), heat content ($X_4$), rainfall ($X_5$), relative humidity ($X_6$), short wave radiation ($X_7$), sea surface density ($X_8$), and sea surface salinity ($X_9$). There are several alternative models for modeling sea surface temperature data. Modeling is done by adding locations as random influences, location interactions with each predictor variable, and interactions between predictor variables that have the greatest correlation value. Then the best model is chosen which has the lowest AIC, BIC, and log-likelihood values. Based on the data analysis conducted, 3 GAMM models were obtained as follows.

a. The 1st GAMM model is a model with random influence that interacts with all predictor variables with the addition of interactions between variables $X_2$ and $X_1$, $X_1$ and $X_5$, $X_2$ and $X_6$, $X_3$ and $X_8$, and $X_7$ and $X_9$. Interaction between predictor variables is determined through the correlation value obtained based on Figure 1.

$$g(\mu) = 1.160 \times 10^{-3} + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_5(x_5) + f_6(x_6) + f_7(x_7) + f_8(x_8) + f_9(x_9) + x_2x_3 + x_1x_5 + x_2x_8 + x_3x_8 + x_7x_8 + (5.808 \times 10^{-7})$$ \hspace{1cm} (8)

b. The 2nd GAMM model is a model with random influence that interacts with all predictor variables.

$$g(\mu) = 1.160 \times 10^{-3} + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_5(x_5) + f_6(x_6) + f_7(x_7) + f_8(x_8) + f_9(x_9) + (1.57 \times 10^{-5})$$ \hspace{1cm} (9)

c. The 3rd GAMM model is a model with a random influence and interaction between predictor variables are correlated, between variables $X_2$ and $X_1$, $X_1$ and $X_5$, $X_2$ and $X_6$, $X_3$ and $X_8$, $X_7$ and $X_9$.

$$g(\mu) = 1.163 \times 10^{-3} + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_5(x_5) + f_6(x_6) + f_7(x_7) + f_8(x_8) + f_9(x_9) + x_2x_3 + 2.886x_1x_5 + x_2x_8 + x_3x_8 + x_7x_8 + (-7.633 \times 10^{-7})$$ \hspace{1cm} (10)

4.2. Optimal Model Selection and Interpretation

Based on the results in Table 6. it is known that the alternative GAMM model that has the AIC, BIC, and the smallest log-likelihood values is the 3rd GAMM model with a random influence and the interaction between predictor variables that are strongly correlated between the variables $X_2$ and $X_5$, $X_1$ and $X_6$, $X_2$ and $X_8$, $X_3$ and $X_6$, and $X_7$ and $X_9$. Based on Table 6. it is known that the model has the AIC, BIC and the smallest log-likelihood is 3rd GAMM model and then the 3rd GAMM model is the best model. The AIC, BIC, and log-likelihood values obtained for 3rd GAMM model are -981,0602; -873,853 and 517,2894.

| Distribution | AIC | BIC | Log-Likelihood |
|--------------|-----|-----|----------------|
| Eksponensial  | 3557,64 | 626,256 | 517,2894 |
| Gamma        | 624,511 | 626,256 | 517,2894 |
| Inverse gaussian | 624,511 | 626,256 | 517,2894 |

Table 5. Comparison of models based on AIC, BIC, and Log-likelihood values
That is known that there are some variables that are not significant, so that the parameter significance test is determined and determine the best model using the backward elimination method.

4.3. Parameter Significance Test

Parameter significance test is conducted to determine whether the predictor variable has a significant effect on the response variable. Here are the hypotheses used.

$H_0$: The $i$-predictor variable has no significant effect on response variable

$H_1$: The $i$-predictor variable has a significant effect on the response variable

The rejection area used is reject $H_0$ if the $p$-value is $<\alpha$ (0.05).

| Parameter                   | Estimated value | $P$-value |
|-----------------------------|-----------------|-----------|
| Intersep                    | $1.160 \times 10^{-3}$ | $<2e-16^{***}$ |
| Location B                  | $-7.366 \times 10^{-7}$ | 0.371 |
| $s(X_1)$                    | $4.11 \times 10^{-10}$ | 0.000407$^{***}$ |
| $s(X_2)$                    | $5.825 \times 10^{-12}$ | $8.83 \times 10^{-5}^{**}$ |
| $s(X_3)$                    | $-1.655 \times 10^{-10}$ | $7.04 \times 10^{-6}^{***}$ |
| $s(X_4)$                    | $5.423 \times 10^{-12}$ | 0.031796$^*$ |
| $s(X_5)$                    | $8.619 \times 10^{-4}$ | $<2e-16^{***}$ |
| $s(X_6)$                    | $-8.657 \times 10^{-11}$ | $<2e-16^{***}$ |
| $ti(X_2 X_3)$               | $2.354 \times 10^{-6}$ | 0.030021$^*$ |
| $ti(X_1 X_5)$               | $2.691 \times 10^{-8}$ | 0.057082$^*$ |
| $ti(X_2 X_6)$               | $-4.302 \times 10^{-6}$ | 0.054772$^*$ |
| $ti(X_3 X_8)$               | $-2.683 \times 10^{-4}$ | 0.024302$^*$ |

Description: significant at $\alpha$ 0.1% (***) , 1% (**), 5% (*) and 10% (•)

Based on the results in Table 7, all significant parameters were obtained with different levels of significance. There are 5 predictor variables and 2 significant predictor variable interactions on $\alpha$ 5%, respectively, variables $X_1$, $X_2$, $X_3$, $X_8$ and $X_9$ and $X_2 X_3$ and $X_3 X_8$. Then obtained 1 predictor variable and 2 significant predictor variable interactions on $\alpha$ 10%, respectively, $X_4$ variables and $X_1 X_5$ and $X_2 X_8$ interactions. The following are the AIC, BIC, and log-likelihood values obtained for the significant 3rd GAMM model, which are -986.0211; -894.911 and 515.7519. Here are the best models obtained.

$$g(\mu) = 1.163 \times 10^{-3} + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_6(x_8) + f_8(x_9) + x_2 x_3 + x_3 x_8 + 2.918 x_1 x_5 + x_2 x_8 + (-7.633 \times 10^{-7})$$

5. Conclusion

Based on data analysis and discussion can be concluded that:

1. The best model of sea surface temperature data using the generalized additive mixed models approach is the 3rd GAMM model. The 3rd GAMM model has the smallest AIC, BIC, and log-likelihood values with values of -986.0211; -894.911 and 515.7519. Following is the construction of the 3rd GAMM model:

$$g(\mu) = 1.163 \times 10^{-3} + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) + f_6(x_8) + f_8(x_9) + x_2 x_3 + x_3 x_8 + 2.918 x_1 x_5 + x_2 x_8 + (-7.633 \times 10^{-7})$$
2. Factors that significantly influence (α = 0.5%) on sea surface temperature (Y) are wind speed (X₁), air temperature (X₂), dynamic height (X₃), sea surface density (X₄), and sea surface salinity (X₉). Then there are several interacting predictor variables that affect sea surface temperature including interactions between air temperature with dynamic height and dynamic height with density of sea surface temperature.

References
[1] Badan Meteorologi, Klimatologi, dan Geofisika 2018 Iklim. www.bmkg.go.id Accessed on February 20, 2018 (Indonesia)
[2] Serambi Indonesia 2017 Bencana Perubahan Iklim www.acehtribunnews.com Accessed on April 24, 2018
[3] Ananda S, Ariska R, Humaira R.A, and Miftahuddin 2017 Pemodelan Suhu Permukaan Laut Menggunakan Generalized Additive models (GAM) Dalam Efek Waktu. Seminar Nasional dan Rapat Tahunan (SEMIRATA) (Indonesia)
[4] Hastie, T and Tibshirani, R 1990 Generalized Additive Models, Statistical Science Chapman and Hall Washington USA
[5] Wood S.N 2017 Generalized Additive Models An Introduction with R CRC Press, United Kingdom
[6] Jamilatuzzahro, Caraka, R.E, and Herliansyah, R 2018 Aplikasi Generalized Linear Model pada R. Innosain Yogyakarta
[7] Codd C.L 2014 A review and comparison of models and estimation methods for multivariate longitudinal data of mixed scale type Disertasion The Ohio State University USA
[8] Shadish W.R, Zuur A.F, and Sullivan K.J 2014 Using Generalized Additive (Mixed) Models to Analyze Single Case Designs Journal of School Psychology 52, 149-178
[9] Dziak, John J, Coffman, Donna L, Lanza, Stephanie T, and Li, R 2017 Sensitivity and Specificity of Informaion Criteria Peerj Preprints
[10] National Oceanic and Atmospheric Administration 2017 Global Tropical Moored Buoy Array Pacific Marine Environmental Laboratory. www.noaa.gov Accessed on February 11, 2018