Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 1 (2015), pp. 519-522

http://www.ripublication.com

Estimation of Malaria Cases With MLR Method In Prajuab Kirikhan, Thailand

J. Mekparyup[1], S. Khamrod[2] and K. Saithanu[3*]
[1,2,3]Department of Mathematics, Faculty of Science, Burapha University
169 Muang, Chonburi, Thailand
[1]jatupat@buu.ac.th, [2]usbenz@hotmail.com, corresponding author:
[3]ksaithan@buu.ac.th*

Abstract

The objective of this study was to estimate the number of Malaria cases in Prajuab Kirikhan using the multiple linear regression (MLR) method. The fitted MLR equation was then chosen with the best subset method. The data was collected since 2009 to 2011 from the Ministry of Public Health, Thailand. The results found that variables influencing the estimation of the number of Malaria cases (y) were maximum temperature (x1), minimum temperature (x2), rain fall (x3), relative humidity (x4) and fitted the equation was $\hat{y}' = -12.40+0.257x1 +0.509x2+0.0952x3 -0.075x4$ with adjusted coefficient of determination 0.404 and standard error of estimation 1.345.

Mathematics Subject Classification: 62J05

Keywords: MLR method, best subset method

Introduction

Malaria is caused by a parasite, called Plasmodium, which is transmitted by the bites of Anopheles. In a few years ago, there are many studies on the factors affecting the rate of incidence of Malaria in countries around the world. These resulting factors are important to the casuse of Malaria depending on climatic factors such as maximum temperature, minimum temperature, rainfall [1][2][3][4][5][6] and relative humidity [1][2][4][5]. Also, there is a research about epidemiology and control of Malaria in South Asia which depicts the rising outbreak of Malaria in South Asia [7]. The South of Thailand located between the Andaman Sea and the Gulf of Thailand has a tropical climate so it is favorable on transmitting Malaria disease of Anopheles breeding. Although the Ministry of Public Health, Thailand, is planning to prevent the epidemic of Malaria continually, Malaria is one of the problems required to be solved urgently especially in Prajuab Kirikhan [8]. This presented research pointed to study an appropriate model to predict the number of Malaria cases in Prajuab Kirikhan for a more effective guideline of preventing the outbreak of Malaria.

Materials and Methods

The number of Malaria cases (y), Maximum temperature (x1), Minimum temperature (x2), Rainfall (x3) and Humidity relative (x4), were collected since 2009 to 2011 from the Ministry of Public Health, Thailand. Firstly, correlation coefficient (R) was used to monitor relationship among air pollutant concentrations and meteorological factors. Then MLR method was used to fit the MLR equation using a dependent variable, y, and four independent variables, x1, x2, x3 and x4, following Equation 1.

$$y = \beta_0 + \beta_1 x1 + \beta_2 x2 + \beta_3 x3 + \beta_4 x4 + \varepsilon \qquad (1)$$

where β_i = the regression coefficient $(i = 0,1,2,...,4)$ and ε = error of the regression model. Selecting the MLR equation, Mallows' C_p [9], standard error of estimation (S) and adjusted coefficient of determination (R^2_{adj}) were considered using the best subset method. Then the selected equation was tested by F statistic of analysis of variance (ANOVA). After receiving the appropriate MLR equation, the assumptions of model were verified in accordance with (I) Normal distribution of the error was tested by Anderson-Darling statistic [10], (II) Independence of the errors was confirmed by Durbin-Watson statistic [11], (III) Homoscedasticity of the errors was checked by Breusch-Pagan statistic [12] and (IV) Multicollinearity among independent variables was examined by Variance Inflation Factor (VIF) as of Equation 2

$$VIF_j = \frac{1}{1 - R^2_{j|others}} \qquad (2)$$

where $R^2_{j|others}$ is the coefficient of multiple determination with independent variable x_j on the $p-2$ other independent variables x in the multiple linear regression model (p is the number of independent variables). Box-Cox transformation [13] was used when any of these four assumptions was violated. Lastly, validation of the fitted MLR equation was illustrated by comparison time series graphs between the observed data (OBS) and the estimated data (EST).

Results and Discussion

The highest positive correlation coefficient was between y and x2 with R=0.559 (P-value=0.000) while y and x1 was the second highly one with R=0.505 (P-value=0.000). Then the MLR equation was fitted by the best subset method with Mallows' C_p= 5.0, S=10.24 R^2_{adj}=0.395 (F=26.28, P-value=0.000) following Equation 3.

$$\hat{y} = -94.580 + 2.0497x1 + 3.6941x2 + 0.7022x3 - 0.6928x4 \tag{3}$$

After generating this equation, the assumptions was monitored; (I) the error of the MLR normally distributed (AD=2.353, P-value<0.005) so Box-Cox transformation was used. Then, the adjusted MLR equation was displayed as of Equation 4 with S=1.345 R^2_{adj}=0.404 (F=27.29, P-value=0.000)

$$\hat{y}' = -12.4 + 0.257x1 + 0.509x2 + 0.0952x3 - 0.0754x4 \tag{4}$$

where $\hat{y}' = \sqrt{y}$. Reconsidering all assumptions of Equation 4, it was found that (I) the error of the MLR equation was normally distributed (AD=0.614, P-value=0.108), (II) the errors were significant independence (DW=1.641 with a critical value 1.378), (III) the variance of errors was significant constant (BP=5.673 with a critical value 9.49) and (IV) there was no correlation among the independent variables (VIF_{x1}=1.884, VIF_{x2}=2.060, $VIF_{x3=1.398}$, VIF_{x4}=1.388). After the obtaining appropriate MLR equation, the validation of the equation was then plotted by time series plot between the OBS and the EST values demonstrated as of Figure 1. The values of OBS and EST were really closed so the graphical validation of the MLR equation was satisfied. That meant, the resulting MLR equation could well predict the number of Malaria cases in Prajuab Kirikhan.

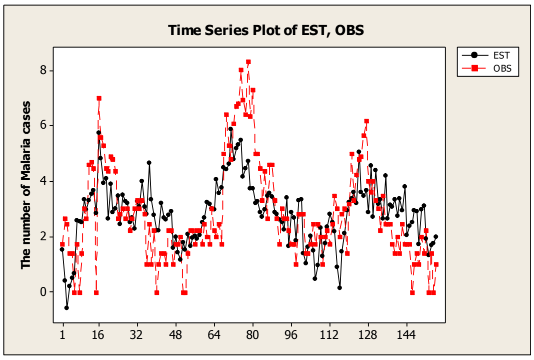

Figure 1: Time series plot between OBS and EST values in Prajuab Kirikhan

Acknowledgement

The authors wish to thank to the Ministry of Public Health, Thailand for kind supporting all data.

References

[1] Al-Mansoob, M. A. K., & Al-Mazzah, M. M., 2005, "The role of climate on malaria incidence rate in four governorates of Yemen," Medical Journal of Malaysia, 60(3), 349.

[2] Ceccato, P., Connor, S. J., Jeanne, I., & Thomson, M. C., 2005, "Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk," Parassitologia, 47(1), 81-96.

[3] Craig, M. H., Snow, R. W., & Le Sueur, D., 1999, "A climate-based distribution model of malaria transmission in sub-Saharan Africa," Parasitology today, 15(3), 105-111.

[4] Githeko, A. K., Ototo, E. N., & Guiyun, Y., 2012, "Progress towards understanding the ecology and epidemiology of malaria in the western Kenya highlands: opportunities and challenges for control under climate change risk," Acta tropica, 121(1), 19-25.

[5] Amin, M. R., Tareq, S. M., & Rahman, S. H., 2011, "Impacts of climate change on public health: Bangladesh perspective," Global J of Enviro Res, 5, 97-105.

[6] Zhou, G., Minakawa, N., Githeko, A. K., & Yan, G., 2004, "Association between climate variability and malaria epidemics in the East African highlands," Proceedings of the National Academy of Sciences of the United States of America, 101(8), 2375-2380.

[7] Kumar, A., Chery, L., Biswas, C., Dubhashi, N., Dutta, P., Dua, V. K., ... & Rathod, P. K., 2012, "Malaria in South Asia: prevalence and control," Acta tropica, 121(3), 246-255.

[8] University Social Resposibility, 2015, "Malaria," Retrieved February 4, 2015, from Mahidol University Web site: http://www.sc.mahidol.ac.th/usr/?p=51

[9] Hocking, R.R., & Leslie, R.N., 1967, "Selection of the best subset in regression analysis," Technometrics, 9(4), 531-540.

[10] Lewis, P.A.W., 1961, "Distribution of the Anderson-Darling Statistic," The Annals of Mathematical Statistics, 32(4), 1118-1124.

[11] Durbin, J., & Watson, G.S., 1951, "Testing for Serial Correlation in Least Squares Regression II," Biometrika, 38(2), 159-177.

[12] Breusch T.S., & Pagan, A.R., "A Simple Test for heteroscedasticity and Random Coefficient Variation," Econometrica, 47(5), 1287-1294.

[13] Sakia, R.M., 1992, "The Box-Cox transformation technique: a review," The statistician, 169-178.