Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 7, Issue, 12, pp.24157-24160, December, 2015

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

ROLE OF ENVIRONMENTAL FACTORS ON INCIDENCE OF MALARIA CASES: A CASE STUDY USING POLYNOMIAL REGRESSION

[1,*]Naba Jyoti Saikia, [2]Hazarika, J., [3]Das, P. K. and [4]Hussain, S.

[1,3]Department of Mathematics, North Eastern Regional Institute of Science & Technology (NERIST), Nirjuli, India-791109
[2]Department of Statistics, Dibrugarh University, Dibrugarh, Assam
[4]System Administrator, Dibrugarh University, Dibrugarh, Assam

ARTICLE INFO

Article History:
Received 26th September, 2015
Received in revised form
24th October, 2015
Accepted 19th November, 2015
Published online 30th December, 2015

Key words:
SPR,
Climatic Factors,
Lakhimpur District,
Polynomial Regression.

ABSTRACT

The measure of malaria incidence on the basis of stratification of the population under study with respect to host and environmental factors is essential to know about the impact of causative/predictive factors. Therefore, to develop a suitable model based on causative/predictive factors on malaria incidence is most important. Here, we developed polynomial regression model to trace the impact of environmental factors on incidence of malaria cases in Lakhimpur district of Assam. Here, LogSPR is considered as explained variables whereas population, maximum temperature, minimum temperature, humidity and average rainfall are considered as explanatory variables. However, lag values of LogSPR at lag one is also considered as one of the explanatory variables. The study reveals that variables minimum temperature, maximum temperature, average rainfall and population play an important role in predicting malaria incidence cases.

Citation: Naba Jyoti Saikia, Hazarika, J., Das, P.K. and Hussain, S., 2015. "Role of environmental factors on incidence of malaria cases: a case study using polynomial regression", *International Journal of Current Research,* 7, (12), 24157-24160.

INTRODUCTION

Malaria is one of the oldest recorded diseases in the world. Now, it is becoming an even greater problem than before. Malaria is endemic in 91 countries with about 40% of the world's population at risk. Each year malaria infects 300-500 million people and kills 1.5-2.7 million people in a year. Out of these more than 1 million children kills (Park, 1997). Malaria in man is causes by four species of the malaria parasites namely Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, and Plasmodium malariae and those are transmitted by the bite of infective female mosquitoes of the genus Anopheles. This was discovered by Ronald Ross, while he was working in Secunderabad (Andhra Pradesh, India).

Ross developed Mathematical model to predict the dynamics of malaria transmission. After developing this model we need another model to estimate input parameters, which is called statistical model. Using this model one can measure malaria incidence on the basis of stratification of the population under study with respect to host (age, sex etc.) and environmental factors (minimum temperature, maximum temperature, humidity and average rainfall etc.). India's geographic position and climatic conditions had been, for long, favourable to the transmission of malaria. Malaria is a seasonal disease. In most parts of India, the maximum prevalence is from July to November. Gomez- Elipe *et al.* (2007) developed a model to predict malaria incidence in an area of unstable transmission in Burundi by studying the association between environmental variables and disease dynamics.

Also, Chattopadhyay *et al.* (2004) proposed a regression model on Plasmodium falciparum malaria death in Kolkata Corporation, India, considering different environment as well as social factors. Further, Chetterjee *et al.* (2009) have developed a simple non linear regression methodology in modeling and forecasting malaria incidence in Chennai city (India), and predicted future disease incidence with high confidence level. Wattanavadee *et al.* (2008) use of general linear regression and generalised linear regression model such as Poisson and negative binomial to identity the patterns of hospital-diagnosed malaria incidences by month, district and age-group for two North-western border provinces in Thailand. Among the models fitted, the best were chosen based on the analysis of deviance and the negative binomial generalized linear model was clearly preferable.

***Corresponding author: Naba Jyoti Saikia,**
Department of Mathematics, North Eastern Regional Institute of Science & Technology (NERIST), Nirjuli, India-791109.

Also, he developed regression models to identify the patterns of hospital-diagnosed malaria incidences in district and quarterly periods in the North-western region of Thailand in 1999-2004. Regression models based on principal components described these patterns. The models showed trends and spatial variations in disease incidence. Graphical displays showing both regional and period effects are presented. The results of this study show that malaria incidence rates decreased substantially in most districts during the study period, but remained very high in border districts with Myanmar (Watanavadee *et al*., 2009). Malaria affects in all age of both sexes. At present malaria is a major public health problem in Assam. Forest-fringed areas of Assam are one of the most affected areas for malaria transmission. Also, hill Districts are considered as worst affected for malaria Transmission (Dev *et al*., 2006). Amongst the four species of the malaria parasite, Plasmodium vivax (PV) has the widest geographic distribution throughout the world. In India, about 70% of the infections are reported to be due to Plasmodium vivax. But in Assam Plasmodium falciparum (the killer parasite) cases outnumbered the Plasmodium vivax cases (Das *et al*., 2007).

Inspite taking various steps by the appropriate authorities, still malaria remains as the major public health problems in Assam and it contributes more than 5% of the total cases recorded in the country annually. Based on the annual parasite incidence (API), defined as number of confirmed cases per thousand of population, ten districts reported less than two cases. Most of the districts in Assam (44% of the total population), API was greater than 2, a criterion which is considered to be sensitive malariometric indicator for residual spray interventions against vector populations. We have observed that as a whole the API of Assam is 2.75 (Dev *et al*., 2006). Also, for the year 2003 we have show that among the districts of upper Assam the API is the highest in Lakhimpur, which is 3. Dev *et al*. (2004) performed a study in Assam based on malaria incidence cases in three population years (1990-1992) to model annual parasitic incidence (API) rates.

Logarithm of incidence rate was modeled as a linear function of the risk factor. Multivariate analysis with backward elimination method was used to develop the most parsimonious model for the risk factors. A Chi-square test for independence was performed to check for any significant differences in the proposition of malaria cases across different age groups and sexes. Stratified analysis was done to avoid confounding effect between these two factors. Malaria cases were observed in all age groups of both sexes. There was no significant difference in proportions of positive cases (P >0 .144) between the two sexes across age groups except in those 5- 15 years old. The multivariate analysis of risk factors (final model selected using backward elimination procedure) indicated that among the risk factors investigated, distance from breeding habitat and geographic location of human settlements were the most significant risk factors in order of importance (with both the P values < .0001 in the final model), and the risk due to distance from health care facility was confounded with the later. The corresponding adjusted relative risks and the 95% confidence intervals were 7.28, 3.23-16.41 and 2.22, 1.49-3.32, respectively.

Very recently, Nath *et al*. (2013) considered malaria incidence rates in Kokrajhar district of Assam over the period 2001-2010 for analyzing temporal correlation between malaria incidence and climatic variables. Associations between the two were examined by Pearson correlation analysis. Cross-correlation tests were performed between pre-whitened series of climatic variable and malaria series. Linear regressions were used to obtain linear relationships between climatic factors and malaria incidence, while weighted least squares regression was used to construct models for explaining and estimating malaria incidence rates. Annual concentration of malaria incidence was analyzed by Markham technique by obtaining seasonal index. Keeping all these points in view, the objective of our study is to fit an appropriate regression model for describing the relation between malaria incidence, environmental factors and population and period factors based on routinely collected data available from block PHCs in Lakhimpur district of Assam. The district cover an area of 2277 sq.km out of which 2257 sq. Km is rural and 20 sq. Km is urban with about 12,09,825 population (2,11,098 no's population are tribal). The climate of the district is warm, subtropical, the temperatures range lies between 14^0-34^0. The relative humidity varies from 60% to 85%. Most areas in the state have heavy rainfall, and floods occur annually. Many villages are located in the foothills/forest fringe.

MATERIALS AND METHODS

Data Management

The study area comprised of six block level primary health care centres (PHCs) – Dhalpur, Bihpuria, Nowboicha, Boginadi, Ghilamora and Dhakuakhana in Lakhimpur district and the PHC's are consisting with 156 number of sub centres. The data collected were mainly of three types- malaria incidence, environmental factors and population from these PHC's. Malaria incidence rates were calculated by monthly basis (January 2008-January 2011) and block PHC's. Incidence rate on malaria incidence such as slide positivity rate (SPR) was calculated by the following formula:

Slide Positivity Rate (SPR): Total positive x 100/Total slide examined.

Here we consider four environmental factors, namely, minimum temperature, maximum temperature, humidity and average rainfall for January 2008 to January 2011 collected from regional meteorological centre, Guwahati.

Models and its estimation

In our present study, polynomial regression model has been adopted to trace the impact of environmental factors on incidence of malaria cases. Here, LogSPR is considered as explained variables whereas population, maximum temperature, minimum temperature, humidity and average rainfall are considered as explanatory variables. However, lag values of LogSPR at lag one is also considered as one of the explanatory variables.

The proposed model has been finalized using multi-step procedure as explained below:

At the very outset, the variable having optimum correlation with the explained variable (LogSPR) has been selected as one of the independent or explanatory variable and we keep inducing from lower order functional form of that variable moving upwards to the higher order depending on the coefficients of determination, R^2 of the model at each step. After all forms of the first variable are induced, we then repeat the procedure with the functional form of second variable based on pair wise correlation in the same way from lower to higher order. The above procedures have been repeated until all the proposed variables are exhausted.

RESULTS

In the following sections, we discuss the results obtained from our method and model for longer time series data consisting of malaria incidences (SPR) over different time periods (January 2008 to December 2011) in Lakhimpur district along with minimum temperature, maximum temperature, humidity, average rainfall and population.

Correlation

Pearson's correlation analysis demonstrated that climatic variables of the district were correlated with SPR (malaria incidence rate). When the association between Lakhimpur district and climatic variable was analyzed month wise, all the climatic variables, except humidity, exhibited significantly positively correlated at 95% confidence interval with SPR. Here, we also consider the correlation of SPR with Population and SPR values at lag- one. Both of these two variables were highly significantly correlated at 95% confidence interval with SPR, while SPR was negatively correlated with Population (coefficient= -0.622483), for other we got positive correlation. The nature of correlations of SPR in Lakhimpur district of Assam with minimum temperature, maximum temperature, humidity, average rainfall and population and SPR values at lag-one are given in Table 3.1.1 and these are ranked depend on their increasing correlation. It is observed that the highest correlation is found between log SPR and SPR at lag-one (coefficient=0.6582116) and the lowest correlation is found between log SPR and humidity (coefficient=-0.0145514).

Table3.1.1. Correlation between SPR and proposed variables

Variable	Coefficient	95% Confidence Interval	Rank
SPR at lag-one	0.6582116	0.4548193 to 0.7963458	1
Population	-0.622483	(-)0.7730810 to (-)0.4054525	2
Maximum temperature	0.4246463	0.1532386 to 0.6364826	3
Minimum temperature	0.4184169	0.1458325 to 0.6319542	4
Rainfall	0.2430772	(-)0.05080594 to 0.49821855	5
Humidity	-0.0145514	(-)0.3035671 to 0.2769164	6

Polynomial Regression Model

Here, we consider SPR in Lakhimpur district over a time period 48 months from January 2008 to December 2011 as a dependent variable and the explanatory variables are minimum temperature, maximum temperature, humidity and average rainfall.

Table 3.2.1. Estimated results of the polynomial regression model

Coefficient	Estimate	Std. Error	t value	Pr (>|t|)
Intercept	1.27E+01	3.48E+00	3.654	0.000759 ***
SPR-at-lag-one	2.60E-01	3.21E-01	0.81	0.422982
I(Population^2)	-1.12E-11	2.48E-12	-4.523	5.57e-05 ***
I(Max_temp^2)	2.84E-03	8.34E-04	3.407	0.001535 **
I(Min_temp^9)	-3.34E-13	1.43E-13	-2.336	0.024723 *
Humidity	-4.07E-03	1.32E-02	-0.309	0.75879
Average_rainfall	3.19E-02	1.33E-02	2.4	0.021243 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Adjusted R-squared: 0.6209 and p=3.913e-08

Also we consider population and SPR values at lag-one as additional independent factors in our study due to the high correlation with log SPR. Then, we identify the initial relationship through the scatter plots between the dependent variable and selected independent variables (Fig.1). According to the procedure mentioned in Materials and Methods we arrive the estimated model equation as shown in Table 3.2.1.

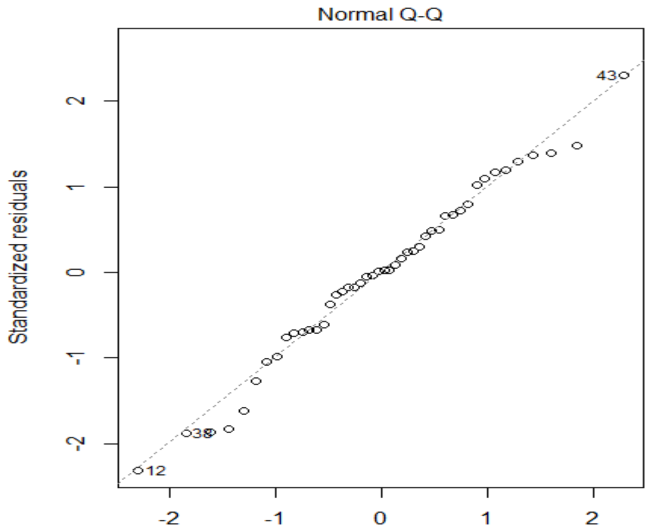

Fig. 1. Q-Q plot

From our model equation, we predict the values of SPR for the given time points and also we forecasts of SPR values based on the independent variable values (Population, Climatic factors) of the corresponding time points as well as the SPR values at lag-one. The model could explain 62.09% of the observed variability in SPR. From our model statistics given in table 3.2.1 it is observed that all climatic variables, except humidity are given significance values. For other independent variables only population is significance at 0.1% level of significance. F-statistics as well as Q-Q plot show quite good fitting of the selected polynomial regression model.

DISCUSSION

Malaria is become a very serious disease Assam, particularly in Lakhimpur district of Assam. Environmental factors play an important role for occurrence of this disease. For development plasmodium vivax and plasmodium falciparum parasite of malaria, generally, favourable temperature range lies between 15-30 and 20-35 respectively (Bhattacharya *et al*., 2006). According to Park *et al*. (1997), a relative humidity of 60% is considered necessary for mosquitoes to live their normal span of life.

When the relative humidity is high, mosquitoes are more active. If the humidity is low, mosquitoes do not live long. For study the effect of rainfall on malaria transmission is complicated. Different studies have given various results. Srimath-Tirmula-Peddinti *et al*. (2015) mentioned that rainfall as the significant factor for malaria transmission. Which is different from Nath *et al*. (2013), according to him rainfall did not bear any significant relationship with malaria incidence. Similarly, Chatterjee *et al*. (2009) told that total rainfall influential only for few study area. However, after tableating our data we are observed that the environment of the district remains favourable for transmission of malaria.

Mainly we can take two major ways for control malaria incidence. One is the parasite control through the control of vector on a geographical basis and the other is to understand the causative factors and underlying transmission mechanics of the disease and on the basis of this a prior information one can map and predict the risk of disease in terms of the causative factors. Here we are study the second approach and observed the effects of environmental factors on malaria incidence in lakhimpur district. In this paper, first we are determined the relationship of each independent variable with the malaria incidence. After that we model the malaria incidence on the basis of large time series data during the period January 2008 to December 2011 with various climatic factors. In earlier, linear and multiple regression analysis on malaria incidence are study by various researchers for find out which independent variable has more significantly impact on dependent variable (Chatterjee and Sarkar, 2009; Nath and Mwchahary, 2013; Srimath-Tirumula-Peddinti *et al*., 2015). However, we consider polynomial regression model derived through a step-by-step process and also with high predictive power of malaria incidence. In our model, it is clear that variables minimum temperature, maximum temperature, average rainfall and population play an important role in predicting malaria incidence cases. Verves, the factor humidity has no significant role as an impact factor of malaria cases. The regression model for our study area and its predictions act as an important indicator for future occurrence of the disease considering the causative factors. For obtained better prediction about the disease incidence from regression model, only we need the previous occurrence of the disease and corresponding climatic factors as well as other factors, do not required any prior assumptions about the disease or knowledge of parameters. It is also mentioned that this can be generalized to any volume of data over any time period with respect to causative factors in the respective study, where we get a significance degree of predictability of malaria incidence. Thus, we can take necessary action for prevention of disease and control the disease for a particular area or as a whole.

REFERENCES

Bhattacharya, S., Sharma, C., Dhiman, R. C. and Mitra, A. P. 2006. "Climate change and malaria in India". *Current Science*, 90(3): 369-375

Chatterjee, C. and Sarkar, R.R. 2009. "Multi Step Polynomial Regression Method to Model and Forecast Malaria Incidence". Plos One., 4(3): e4726.

Chattopadhyay, J., Sarkar, R.R., Chaki, S. and Bhattacharya, S. 2004. "Effects of environmental fluctuations on the occurrence of malignant malaria - a model based study". *Ecological Modeling,* 177:179–192.

Das, N.G., Talukdar, P.K., Kalita, J., Baruah, I. and Sribastava, R.B. 2007. "Malaria situation in forest-fringed villages of Sonitpur district (Assam), India bordering Arunachal Pradesh during an outbreak". *J. Vect. Borne Dis.,* 44, 213-218.

Dev, V., Dash, A.P. and Khound, K. 2006 ."High- risk areas of malaria and Prioritizing interventions in Assam". *Current Science*., 90: 32-36.

Dev, V., Phookan, S., Sarmah, V.P. and Anand, S.P. 2004. "Physiographic and Entomologic Risk Factors of malaria in Assam, India". *Am. J. Trop. Med. Hyg*., 71(4): 451-456.

Gomez-Elipe, A., Otero, A,. Herp van, M. and Aguirre- Jaime A. 2007. "Forecasting malaria incidence based on quarterly case report and environmental factors in Karuzi, Burundi, 1997-2003". *Malaria Journal*, 6: 129.

Nath, D.C., Mwchahary, D.D. 2013. "Association between climatic variables and malaria incidence: a study in Kokrajhar district of Assam, India". *Glob J Health Sci.,* 5: 90–106.

Park, K. 1997. "Preventive and Social Medicine". M/S Banarsidas Bhanot Publishers, 15-Edition.

Srimath-Tirumula-Peddinti, R.C.P.K., Neelapu, N.R.R. and Sidagam, N. 2015. "Association of Climatic Variability, Vector Population and Malaria Disease in District of Visakhapatram, India: A Modeling and Prediction Analysis". Plose ONE. 10(6): e0128377.

Watanavadee, S. and Metta, K. 2009. "Modeling Malaria Incidence in north western Thailand". *Chiang Mai. J. Sci.,* 36(3): 403-410.

Watanavadee, S., Metta, K. and Naratip, J. 2008. "Malaria in north western Thailand". *Songklanakarin Journal*, 207-214.

****