Study on Temporal Effects of Urban Malaria Incidences

Krishnendra S. Ganguly[1], Soumita Modak[2,*], Krishna S. Ganguly[3] and Asis K. Chattopadhyay[2]

[1]*Birla Institute of Technology, Mesra, India*

[2]*Department of Statistics, University of Calcutta, Kolkata, India*

[3]*Kolkata Municipal Corporation, Kolkata, India*

Abstract: In Africa and Asia Malaria is considered to be the most widespread vector-borne disease taking lives of many people and specially affecting children. Many parts of India are significantly affected by malaria over a long period of time. Kolkata is one of the Metropolitan cities in India where the seasonal effect of malaria is very common. In the present work attempts have been made to study temporal variation of urban malaria incidences using time series model on the basis of a large survey conducted by the Kolkata Municipal Corporation. It is found that the proposed time series model can be used successfully for prediction purpose.

Keywords: Malaria, Spatio-temporal variation, Time series model, Urban.

1. INTRODUCTION

Malaria is one of the most widespread vector-borne diseases of our time, taking the lives of almost one million people every year, most of them in sub-Saharan Africa and in children under 5 years of age [1]. It is the fifth leading cause of death worldwide and almost half the world's population (3.3 billion) [1] is at the risk of contracting malaria.

Children and pregnant women are among the most vulnerable. The disease is not only a major killer in Africa but a primary cause of poverty [2,3]. Malaria traps people in poverty and undermines the development of some of the poorest countries in the world. Though the majority (85%)[x] of the cases as well as deaths from malaria are found in sub-Saharan Africa, malaria is a global problem and a killer disease, affecting countries as an endemic in large parts of Asia and Latin America. Malaria imposes great socio-economic burden on humanity, and with six other diseases (Diarrhoea, HIV/AIDS, Tuberculosis, Measles, Hepatitis-B, and Pneumonia), accounts for 85% of global infectious disease burden [4,5]. In addition, the estimated annual mortality attributed to malaria ranges from 700 000 to 2.7 million globally and more than 75% of them are African children and expectant mothers. Doubts have been expressed about reliability of these estimates because most of the hyper- and holo-endemic countries [6,7]. World Health Organization reports on the South-East Asian (SEA) Region show that there is a land area of 8466600 sq. km. for 11 countries, i.e., 6% of global area, where amongst around 1.4 billion people 1.2 billion are exposed to the risk of malaria, with the majority living in India [8]. However, SEA contributed only 2.5 million cases to the global burden of malaria, while India alone contributed 76% of the total cases. Epidemiologic models, geographical and demographic data, in consideration of clinical episodes, point to the fact that Plasmodium falciparum [PF] estimates outside Africa, especially in SEA, are 200% higher than reported by the WHO (i.e., 118.94 million of global estimates out of 515 million cases) [7]. Burden of P. vivax [PV] malaria in the world has been calculated as 71-80 million cases, of which Southeast Asia and Western pacific countries contributed 42 million cases [9].

URBAN Malaria, i.e., Malaria cases in urban settings, has several unique features compared to RURAL Malaria cases, particularly when these cases are considered for designing their control programme. Incidence of Rural malaria lowered drastically In India by 1970s, i.e., from 0.1 to 0.15 million cases per year, while urban towns reported rising trend. Urban Malaria is contributed by large scale urban-rural migrations triggered by urban "push" (for earning livelihood in suburban and rural areas) and urban "pull" (to avail of both medicare / healthcare and educational opportunities in urban areas) phenomenon. Unplanned urbanization, different construction-projects with complete disregarding of health impact assessment and/or eco-friendly technologies, demographic and societal changes, etc. day by day are contributing to enhanced potentials for breeding of malaria and other diseases Vectors. Factors, like, inability of the civic bodies to supply with pure water to every household,

*Address correspondence to this author at the Department of Statistics, University of Calcutta, Kolkata, India;
E-mail: soumitamodak2013@gmail.com

E-ISSN: 1929-6029/16

incapability to dispose of sewage and/or solid-wastes regularly and hygienically, etc., all lead to an all-round disruptions in Vector ecology. Infrequent water-supply lead to increased water-storage practices, resulting in extensive breeding of An. stephensi, the main vector of urban malaria in the study-area. Urban malaria control was thus thought of as a separate strategy, i.e., UMS (Urban Malaria Scheme) for National Vector Borne Disease Control Programme (NVBDCP) in India.

Though malaria-epidemics are usual phenomenon alike other vector-borne diseases, malaria typically shows temporal and spatial variations, based on different climatic, ecological and human factors [10-14], which can be used to correlate the spatio-temporal distribution of disease-burden, guiding the management aspect of cost-effectiveness of resource-allocation for disease-intervention in space and time. Also, if any time- series surveillance-data can be used correlated with different climatic factors [15-17], forecasting some Warning signals of malaria, institution of intervention to control malaria will be much better before any epidemic takes place [18,19]. Similar efforts were done in Ethiopia [20], in Kenya [21,22], in Southern Asia [23] and in China [24]. In the current study, monthly patterns of malaria in Kolkata (a metropolitan city in India) are explored on reported PF and PV clinical events over a short 6 year-period showing both Spatio-Temporal variation, only temporal variations being statistically analyzed herein.

The Kolkata Municipal Corporation (K.M.C.) is working in this urban malaria control programme through a multidisciplinary approach. The K.M.C. deployed a poly-pronged system of controlling Urban Malaria of which the present study is just a trifling observation of the mainframe work-galore, where data (as are regularly being collected by the K.M.C.) have been analyzed to get some nature about the occurrence of malaria in a geographical region of the city of Kolkata as well as to make some comparison study on time-scale and spatial scale, plus some correlation studies about the disease-incidence also. Spatio-temporal correlation study of urban malaria cases in Wards of a multi-racial, multi-lingual, cosmopolitan city can provide a basic standard for effective planning and evaluation of Malaria-control programme. To understand the spatio-temporal variation of urban malaria incidences in Kolkata, a trend analysis has been planned to be effected. However, as a pilot study, in this paper a temporal trend analysis has been done for Kolkata.

The paper is organized as Introduction (Section 1), Data Description (Section 2) and Descriptive Data Analysis (Section 3). Section 4 and 5 discuss time series analysis and modeling of the time series, respectively. Results and discussion are given in Section 6.

2. DESCRIPTIVE DATA ANALYSIS

The Vector-Borne Disease control project, as has been discussed, is being conducted in an old and renowned city of India under the administration of the K.M.C. Kolkata consists of 141 Wards (Figure **1**), distributed within 15 Boroughs (administrative zones). The administrative area of Borough-V (Figure **2**) is a conglomeration of eleven (11) Wards of Kolkata, where this study has been. More elaborately, it is a study based on the population of Ward-36, Ward-37, Ward-40, Ward-41, Ward-42, Ward-43, Ward-44, Ward-45, Ward-48, Ward-49 and Ward-50 of Kolkata. But more correctly, the study-design is based on those persons, who are being tested plus treated for febrile illnesses of Malaria (through active and/or passive surveillance), at twelve Ward-level Malaria Clinics (cum Treatment-Centres) of 11 Wards of Borough-V of the K.M.C. between January 2008 and December 2014. It is worth mentioning here that Borough-V is chosen, because geographically it is the heart of the Kolkata, which includes main gateway Railway Station (Sealdah) gateway Bus-stand (esplanade), principle business-areas (Burrabazar), main office-places (Esplanade) of the city, while it economically includes comparable number of slums with high-rises, socio-culturally involves maximum open play-grounds of cricket, football, etc. (including famous Eden Gardens), historical Fort-areas (e.g., Fort Williams), maximum schools, colleges, educational institutions with books/ magazines and other publications houses (College Street areas), one old renowned medical college & hospital (Calcutta Medical College & Hospitals) and Borough-V is situated on the bank of the famous Ganges river. Borough-V was particularly infamous for its malaria-prone nature.

The study has been planned to be conducted based on the data already collected by the Kolkata Municipal Corporation Ward-level Malaria Clinics (cum-Treatment-Centres) over a six-year period from January 2008 to December 2013. The secondary data, as will be used for this study, are an extract from the daily-collected database of the K.M.C., which have been arranged and grouped to form a month-wise database. The month-wise data, collected over six-year

Figure 1: Kolkata in West Bengal in India.

period in cases of every individual Ward of Borough-V, are only considered for necessary analysis, as mentioned in the objectives of the study. High-resolution raster maps of the study-area are imported into a geographical information system (GIS) and linked spatially [25] to a digitized boundary map of the 11 Wards across Borough-V of the city of Kolkata. Data from national censuses [26] conducted in 2011 were used to impute inter-censual growth rates at the county level and to estimate population counts for each month between January 2008 and December 2014.

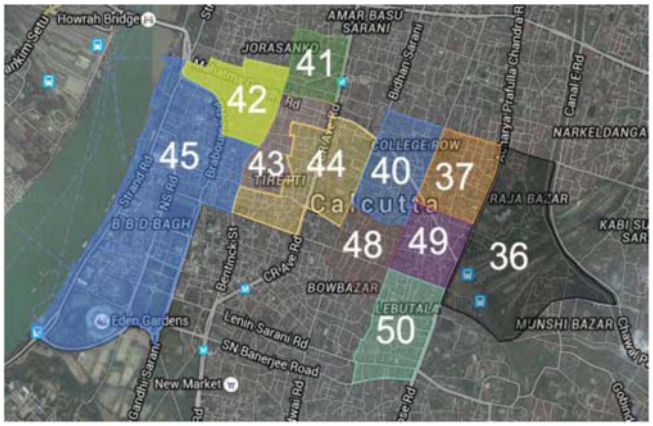

Figure 2: Borough-V of Kalkata.

Fever patients (known by active and passive surveillance), belonging to all age-groups & sex, are subjected to Blood-Slide Testing (stained with Giemsa stain) for detection of malarial parasite and the species, under monitoring & supervision of Medical Officers of Wards plus Laboratory Coordinator with cross-checking for quality-issues of prepared slides and diagnosis of positive-cases by random selection of 10% of slides which were labeled positive and 10% of slides which were labeled negative. Generated data are recorded and tabulated month-wise and Ward-wise for Kolkata.

Annual Blood-slide Examination Rate (ABER =all slides examined x 100 / population), Slide Positivity Rate (SPR =total positive slides x 100 / total slides),

Species-specific SPR and Plasmodium falciparum Ratio (PFR% =total falciparum-Malaria x 100 / total positive slides) and Annual Parasite Incidence (API =total positive cases x 100 /population). Slide Positivity Rate (SPR), as described by WHO, is a Ratio between the number of laboratory- confirmed Malaria cases per 100 clinically suspected cases (whose Blood Slides are examined), indicating quantity & quality of detection of patients with malaria (increasing Rate indicating beginning of a malaria epidemic or a seasonal outbreak). Proportion of fever caused by malaria is almost a similar Ratio where number of Fever cases with confirmed parasitaemia divided by total number of fever, suggests how important a health priority is Malaria. SPR is thus considered as a substitute to estimate temporal changes in malaria-incidence. Moreover, SPR, an easy and inexpensive way to know malaria-burden in a population utilizing data from health-care facilities, is quite accurate as it considers laboratory-confirmed malaria-cases only, in cross-sectional studies, to know malaria endemicity [27,28] and malaria-control interventions [29,30]. However, there is little work to know quantitative relationship between the metrics of temporal SPR-changes and malaria-incidence changes using empiric data. But Trevor P Jensen, in Africa [31], conducted a study to compare the results, showing observed changes in the malaria-incidence with changes in SPR and concluded that SPR is a surrogate measure of malaria-burden. After interventions to control malaria, one should follow secular trends in malaria-incidence based on surveillance-data like SPR.

As has been described above, SPR has some advantages as a surrogate measure to follow changes in Malaria-incidence, as it uses normal HMIS data, without involvement of any special cost for measuring actual malaria-incidence. SPR, in spite of being a very useful measure of malaria-control interventions, do not change corresponding to proportional/ linear change in actual malaria-incidence. This is because fever is used as the criterion for laboratory testing for SPR, while population is the denominator for malaria-incidence; thus any change in the incidence of non-malaria fevers may result in a change in SPR which is not reflected in actual incidence of malaria. Thus, SPR is an estimate of relative-change for malaria-incidence over time, failing to estimate actual malaria-incidence in a target population.

Many factors, viz., quality aspects of laboratory diagnosis, age & sex of patients, seasonality of climatic variables affect malaria-incidence, which may maximize or shroud the impacts of interventions of control programme. In practice, these factors also affect the ability of SPR to estimate surveillance-effects on modifying actual malaria-incidences. Simultaneously majority of malaria surveillance systems gather only a subset of fever-cases in a target population, i.e., those are limited by incomplete and/or inaccurate laboratory testing.

In this longitudinal study, a descriptive study of Malaria has been conducted Year-wise and Ward-wise. Slide Positivity Rates (SPR) of Malaria for each Ward is calculated by dividing the observed number of slide-positive malaria cases across 2008-2015 by the number of BSE for Malaria. Table **1** shows Microscopic Blood-Slide Test Results in Borough-V from 2008-2015, while Table **2** depicts comparison of 2008-2015 parameters of ABER, SPR, PFR and API. Table **1** shows there is gradual rise of PV and PF amongst BSP slides from 2008 to 2010, but due to anti-malaria intervention there is gradual fall in BSP from 2011 to 2015. Simultaneously, there is similar rise and fall of both PV and PF, but PF-incidences remarkably got reduced. Pair-wise Pearson cross-correlations for each

Table 1: Microscopic Blood-Slide Test Result in Borough-V from 2008-2015

Years	BSE	BSP	PV	PF
2008	35305	6830	5764	1054
2009	41248	12672	10139	2515
2010	42878	14532	11722	2756
2011	25606	6622	5677	934
2012	32420	5915	5058	849
2013	19307	3072	2769	299
2014	16153	1265	1187	76
2015	18590	1672	1576	95

Table 2: Comparison of Eight Years Parameters

Years	ABER (%)	SPR (%)	PFR (%)	API (%)
2008	13.37	19.35	15.43	2.59
2009	15.62	30.72	19.85	4.80
2010	16.23	33.89	18.97	5.50
2011	9.69	25.86	14.10	2.51
2012	12.27	18.24	14.35	2.24
2013	7.31	15.91	9.73	1.16
2014	7.89	7.83	6.01	0.62
2015	9.08	8.99	5.68	0.82
8 Year-Average	11.43	20.10	13.01	2.53

Table 3: Monthly Slide Positivity Rate (X_t), of Borough-V Malaria Cases in Kolkata from January 2008 to December 2015

Year	January	February	March	April	May	June	July	August	September	October	November	December
2008	0.076449	0.069351	0.08642	0.094378	0.140588	0.140891	0.2	0.235177	0.196334	0.229087	0.212941	0.167568
2009	0.113285	0.079918	0.121279	0.237934	0.155544	0.302343	0.33255	0.358866	0.332957	0.333333	0.335996	0.2715
2010	0.18481	0.145095	0.241769	0.309524	0.30504	0.281837	0.331989	0.423395	0.375956	0.347832	0.320859	0.288827
2011	0.19709	0.191242	0.252895	0.365642	0.379242	0.263343	0.317151	0.339789	0.236282	0.222278	0.183855	0.178826
2012	0.086735	0.095395	0.149502	0.162592	0.157895	0.104704	0.099198	0.196131	0.217778	0.185358	0.214412	0.143013
2013	0.104294	0.079572	0.096216	0.132444	0.180907	0.196667	0.194006	0.186756	0.19569	0.156463	0.134586	0.070892
2014	0.030979	0.038076	0.04757	0.058352	0.044882	0.07329	0.09603	0.145178	0.114246	0.068984	0.057993	0.049577
2015	0.015692	0.012881	0.014493	0.030516	0.042254	0.037756	0.052914	0.126808	0.13495	0.118246	0.132773	0.058964

month (January-December) are calculated to investigate temporal correlation in the Malaria positive data, separately for P. vivax or P. falciparum cases, over the span of years 2008-2015, in Borough V.

In Borough-V of Kolkata, in eight years (2008-2015) total slides examined are 231507 in all Wards, where total positive malarial parasites are 52580, out of which P. vivax is found in 43892 and P. falciparum in 8578 cases, as depicted in Table **1**. Table **3** shows Monthly Slide Positivity Rate (X_t), of Borough-V Malaria Cases in Kolkata from January 2008 to December 2015. Based on total slides examined and positive cases for malarial parasites, data are subjected to determine Annual Blood Examination Rate (ABER), Slide Positivity Rate (SPR), Plasmodium Falciparum Ratio (PFR%) and Annual Parasite Incidence (API), as are shown in Table **2**. In eight years, ABER has an average of 11.43%. During this period, SPR is 20.1%. Unlike SPR, PFR increased from 2008 to 2011, but there is gradual fall from 2012 to 2015. PFR ranges from 5.68% to 19.85% with an average of 13.01%. Annual Parasite Incidence was being 0.62 to 5.50 with an average of 2.53 during this period. The month wise data revealed that SPR, PFR and API increased in post-monsoon season, maximally between August-September and November-December.

GIS Maps of crude SPRs for the 11 Wards of Borough-V for 2008-2015 are presented in Figure **3**. It is observed from the GIS maps that there is a noticeable change of SPR over the period from 2008 to 2015 for each Ward of Borough-V. Here, the Color Gradient Scale is used where yellow indicates the lowest range of SPR values and Red indicates the Highest SPR values in Borough-V. Now from the GIS maps over the period, we can tell nature of changes of SPR-values is different for each Ward. Some Wards, e.g., 36, 41, 43, 44 & 48 do not ever cross the values of 35% (Sky Blue), whereas Ward 50 only goes beyond the value 65% (Red). It can also be seen from the GIS Maps that for every Ward there is an increase of SPR from year 2008 to 2010 but from 2011 to 2015 the SPR values decreases; SPR for some wards like 36, 37, 40, 41, 43, 44, 48, 49 and 50 come below 25% (Green) during this period 2012-2015. It is also noticeable that the SPR values for Ward 50 both increases and also decreases sharply within 2008-2015.

3. ANALYSIS OF TIME SERIES

We consider a time series Y_1, Y_2,…,Y_N having length N, where study variable SPR is a monthly series of length 96 observed over eight years. Time series plot (Figure **4**) shows an upward trend from the start to late 2010, which after 2010 changes to downward till the finish. There is a pattern of a peak in summer and rains, i.e., from July to September with peak in August and a trough in winter, i.e., from December to January, whereas a smaller peak in November is present for most years. Yearly seasonality in the series increases at the same sort of rate as the yearly mean levels, so we choose a multiplicative seasonal model for our data.

Figure **5** shows the decomposition of time series Y_t into seasonal (S_t), trend (T_t) and irregular (e_t) components using moving averages [32] wherein an upward trend from the year 2008 to 2010 and then downward upto the end of 2015 and seasonality of period 12 months are prominently displayed. For the decomposition we use the multiplicative model $Y_t=T_t*S_t*e_t$, where a moving average uisng a symmetric window with equal weights is used to comupte T_t. Then T_t is removed from Y_t and S_t is estimated averaging over all periods and recycled. Finally e_t is obtained by eliminating S_t from Y_t using the above considered model.

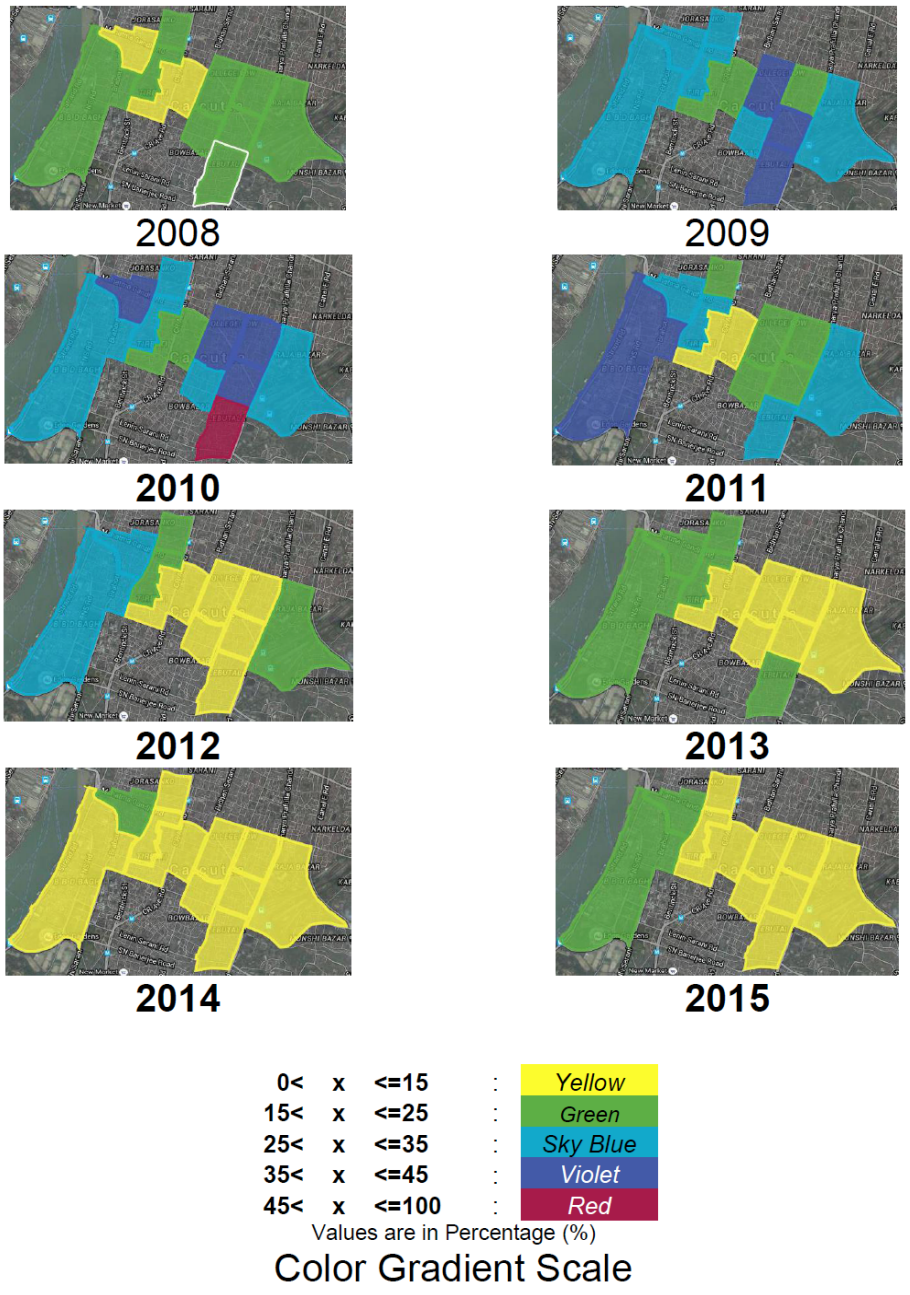

Figure 3: GIS Maps of crude SPRs for the 11 Wards of Borough-V 2008-2015.

3.1. Analysis of Correlogram

For time series Y_t, the sample version of the autocorrelation coefficient (ac.f.) at lag k is given by

$$r_k = \frac{\sum_{t=1}^{N-k}(Y_k - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{N}(Y_t - \bar{Y})^2} \text{ , k=1, 2, …, M and M<N.} \quad (1)$$

Here we take M=48. Figure **6** shows the correlogram, plot of r_k versus k, of Y_t along with the dotted lines at ±2/√(N)= 0.204 , wherein outside these lines values are considered to be significantly different from zero [33]. Correlogram clearly suggests the existing seasonality pattern in the SPR series and the large values of r_k at consecutive lags on both sides of r_k= 0 indicates the trend present in the SPR series.

4. MODELLING OF TIME SERIES

We consider a general multiplicative seasonal ARIMA (SARIMA) model [34] of order $(p,d,q)\times(P,D,Q)s$ as

$$\phi_p(B)\Phi_P(B^s)W_t = \theta_q(B)\Theta_Q(B^s)Z_t \text{ ,}$$

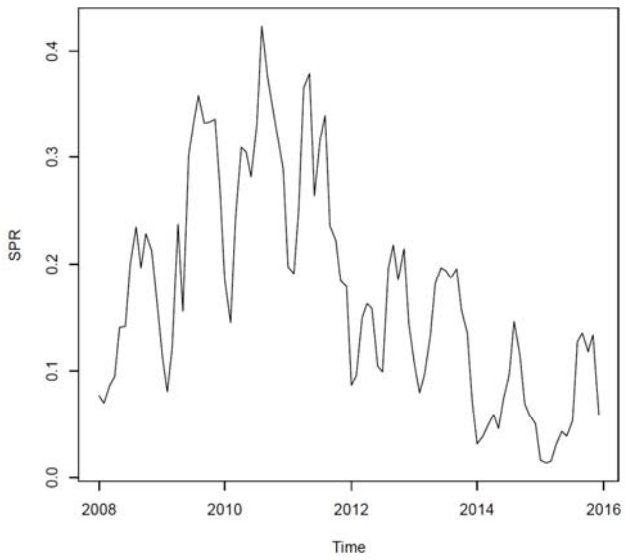

Figure 4: Time Series Plot.

where B denotes the backward shift operator and s is the seasonal period. $\phi_p, \Phi_P, \theta_q, \Theta_Q$ are polynomials of order p (autoregressive (AR) order), P (seasonal autoregressive (SAR) order), q (moving-average (MA) order), Q (seasonal moving-average (SMA) order), respectively and Z_t denotes a purely random process. $W_t = \nabla^d \nabla_s^D Y_t$ is the differenced series, obtained by applying difference operator of appropirate order on Y_t to remove the non-stationary elements from Y_t, where d, D are called difference order and seasonal difference order, respectively.

From the correlogram of $\nabla_{12} Y_t = \tilde{Y}_t$ (Figure **7**) we see that seasonality is removed from Y_t by taking the s^{th} (here s=12) order difference on Y_t which gives $D = 1$. Existing trend in $\tilde{Y}_t$ is eliminated by taking the 1st order difference on $\tilde{Y}_t$ as follows $W_t = \nabla \nabla_{12} Y_t = \nabla_{12} Y_t - \nabla_{12} Y_{t-1} = (Y_t - Y_{t-12}) - (Y_{t-1} - Y_{t-13})$, suggesting $d = 1$ to obtain the sationary series W_t. The ac.f. plot of W_t (Figure **8**) shows significantly large value at only lag 12 among the seasonal lags, hence we choose Q=1 to take into account the present one seasonal MA term. Figure **8** also has significant value at non-seasonal lag 1 indicates one MA term present, i.e., q=1. There is slightly significant value at unusual lag 38 which is ignored as it is meaningless under the study. The partial ac.f. plot of W_t (Figure **9**) shows large value only at seasonal lag 12 implying P=1, i.e., one seasonal AR term present and at non-seasonal lag 1 suggesting p=1. Also notice Figure **9** has significant value by a slight amount at lag 9, which is not

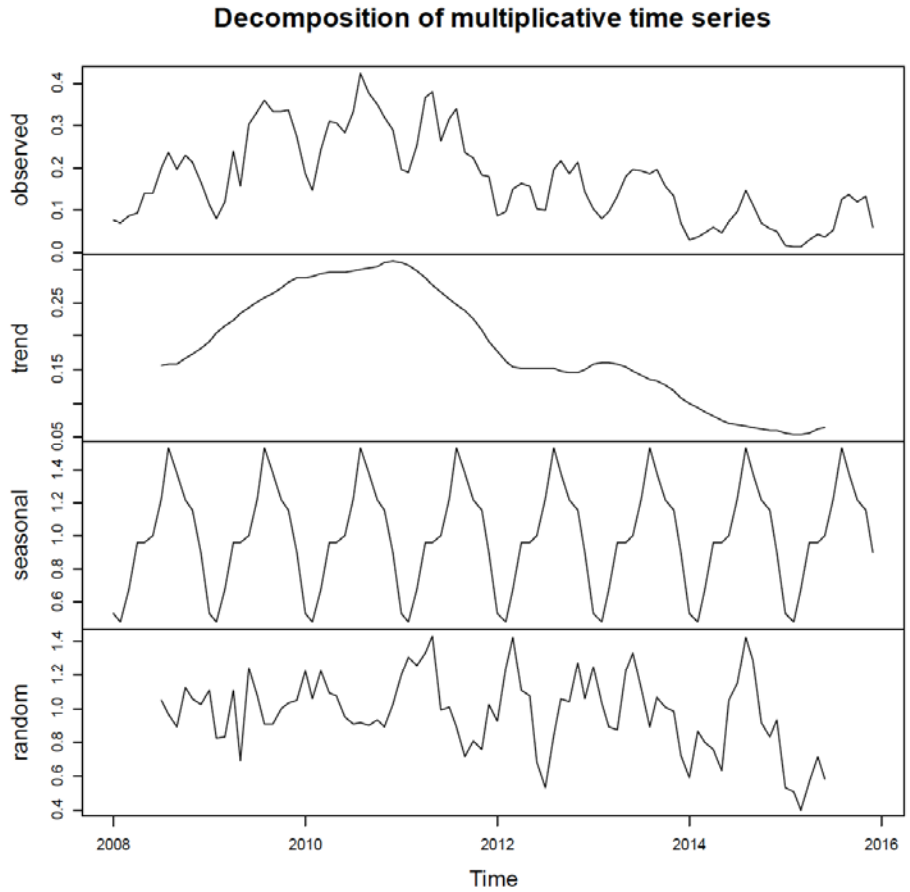

Figure 5: Components of Time Series.

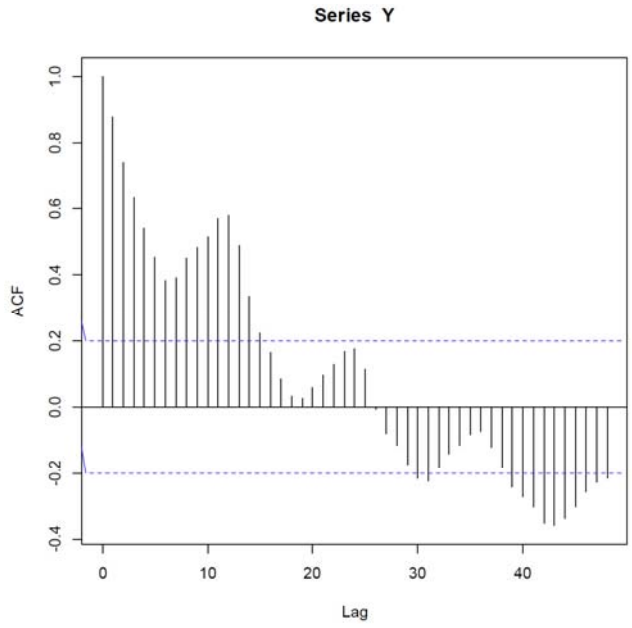

Figure 6: The correlogram of of Y_t .

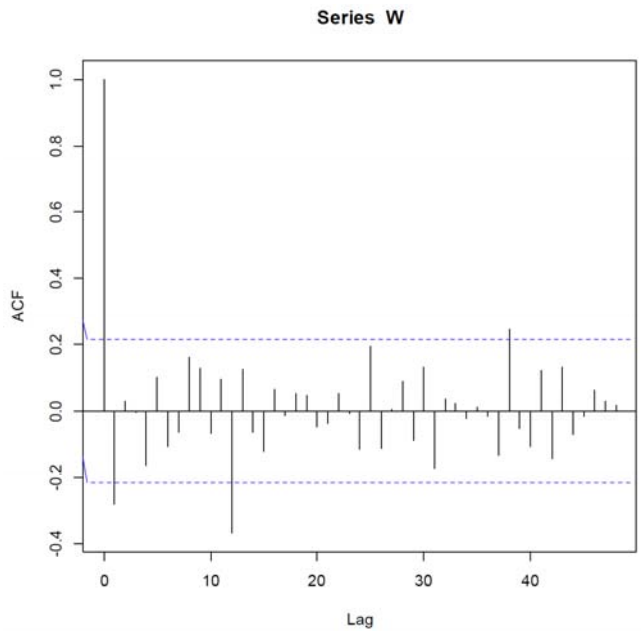

Figure 8: The correlogram of W_t .

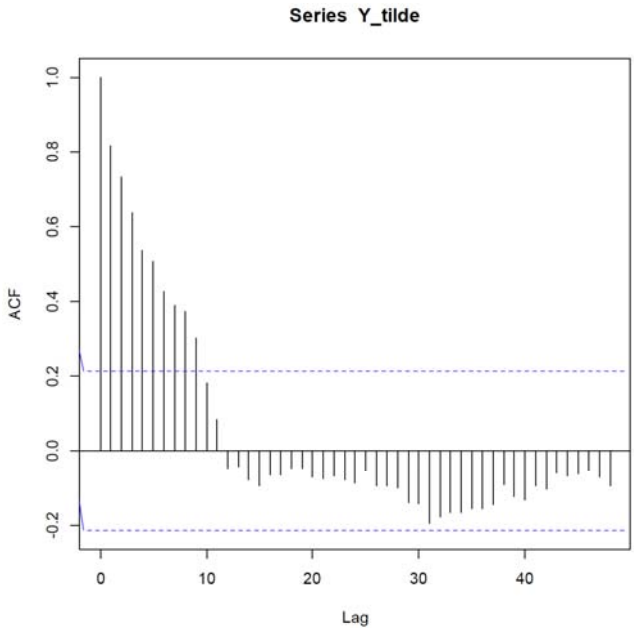

Figure 7: The correlogram of $\tilde{Y}_t$.

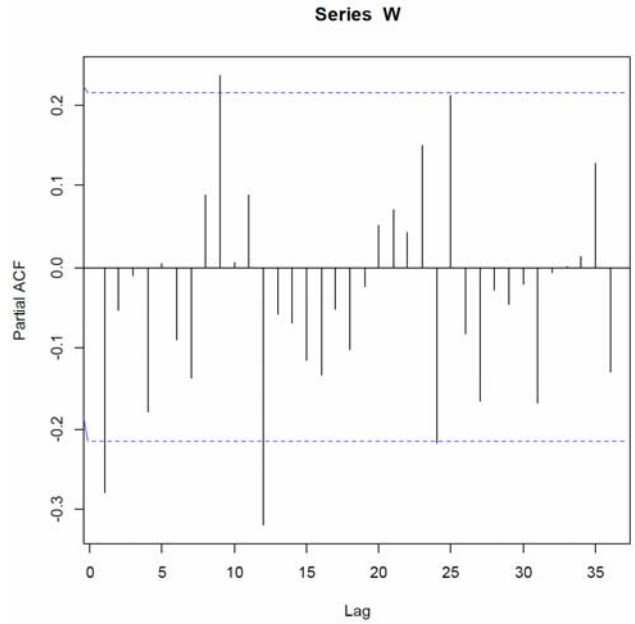

Figure 9: The partial ac.f. plot of W_t .

considerable in this context as this lag has no information to be physically interpreted. Also there is just significant value at the 2nd seasonal lag 24. Though the value is not large, the lag is very important, so we also consider P=2.

We compare the two plausible SARIMA models for SPR series as 1) SARIMA of order (1,1,1) x $(1,1,1)_{12}$ and 2) SARIMA of order (1,1,1) x $(2,1,1)_{12}$ using model selection criteria: the Akaike information criterion (AIC), Second-order Akaike Information Criterion (AIC_c) and Bayesian information criterion (BIC), where the model with the lowest value is the best among the compared ones [35-37]. Hence model (1) selected (see Table **4**) for the present study is given by

$$(1+\phi B)(1+\Phi B^{12})W_t = (1+\theta B)(1+\Theta B^{12})Z_t\,, \tag{2}$$

or

$$\begin{aligned} X_t = \; & (1-\phi)X_{t-1} + \phi X_{t-2} + (1-\Phi)X_{t-12} + (\Phi - \phi\Phi + \phi - 1)X_{t-13} \\ & + \phi(\Phi - 1)X_{t-14} + \Phi X_{t-24} + \Phi(\phi - 1)X_{t-25} - \phi\Phi X_{t-26} \\ & + Z_t + \theta Z_{t-1} + \Theta Z_{t-12} + \theta\Theta Z_{t-13}\,, \end{aligned} \tag{3}$$

where $W_t = \nabla\nabla_{12}Y_t$ and $\phi, \Phi, \theta, \Theta$ are constant parameters to be estimated. Table **5** shows the estimates for the parameters using maximum likelihood method with conditional-sum-of-squares estimates as intial values for the parameters.

Table 4: Model Choosing Criteria

Criteria	Model (1)	Model (2)
AIC	-280.80	-278.96
AIC_c	-280.02	-277.85
BIC	-268.71	-264.44

Table 5: Estimates for Parameters of SARIMA (1,1,1) x $(1,1,1)_{12}$

	ϕ	θ	Φ	Θ
estimate	0.3777	-0.6296	0.1439	-0.9999
Standard error	0.2690	0.2218	0.1223	0.2186

4.1. Cross-Validation of Chosen Model

For a completely random time series consisting of independent and identically distributed observations of length N, large, $r_k \approx 0$ for all $k \neq 0$ and r_k follows $N(0, \frac{1}{N})$. The chosen model is considered to be describing the time series under study well if the residual series, i.e., the difference series from observartion of fitted series, is random. Plot of residuals (Figure **10**) and corresponding correlogram (Figure **11**) indicate that our fitted model describes the observed SPR series quite well. Also non-parametric Kolmogorov-Smirnov (K-S) test to test the null hypothesis that r_k of residuals follows $N(0, \frac{1}{N})$ is performed and the null hypothesis is accepted with a large p-value at 5% level of significance (see Table **6**). Test based on the modified Ljung-Box-Pierce statistic [38] for examining the null hypothesis that the fitted model is appropriate, has quite large p-value at all the lags from 1 to 48 get the null hypothesis accepted at 5% level of significance (see Table **7**).

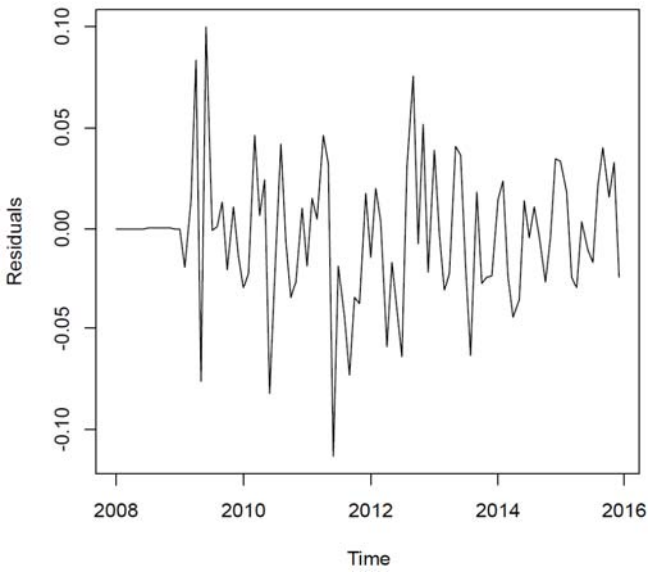

Figure 10: Time Series plot of Residuals.

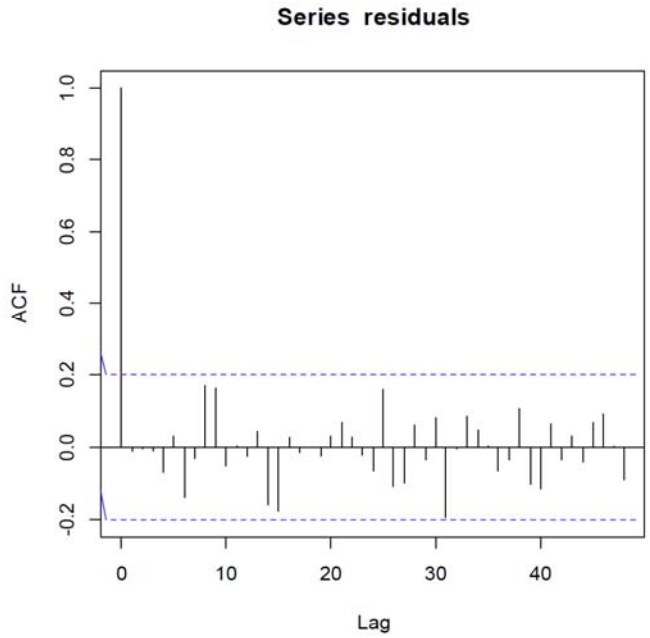

Figure 11: The correlogram of residuals.

Table 6: K-S Test on r_k of Residuals

Test statistic	p-value
0.10194	0.6914

5. RESULTS AND DISCUSSION

In this paper time series analysis includes the total SPR series for 11 wards of Kolkata city observed over each month from the year 2008 to 2015 collected at the Kolkata Municipal Corporation Ward-level Malaria Clinics. By analyzing the data we choose model SRIMA $(1,1,1)x(1,1,1)_{12}$ and fit it to the data, which describes the data quite well. We also forecast the SPR series for the year 2016 with the help of equation (3) and the fitted values of model parameters from Table **5**. Point and interval forecast of Y_t for the year 2016 are shown in Figure **12** and Table **8**. As forecast says, SPR will reach a higher peak in this year compared to last year, which needs to be taken care of.

Table 7: p-Value Based on the Modified Ljung-Box-Pierce Test Statistic

Lag	p-value	Lag	p-value
1	0.8851	25	0.9689
2	0.8945	26	0.9740
3	0.9723	27	0.775
4	0.9924	28	0.9842
5	0.9959	29	0.9891
6	0.8442	30	0.9916
7	0.8813	31	0.9834
8	0.7545	32	0.9881
9	0.7067	33	0.9897
10	0.7593	34	0.9926
11	0.8277	35	0.9949
12	0.8772	36	0.9963
13	0.8922	37	0.9975
14	0.8732	38	0.9979
15	0.8098	39	0.9980
16	0.8477	40	0.9981
17	0.8846	41	0.9986
18	0.9121	42	0.9991
19	0.9287	43	0.9994
20	0.9484	44	0.9996
21	0.9576	45	0.9997
22	0.9707	46	0.9998
23	0.9779	47	0.9999
24	0.9843	48	0.9999

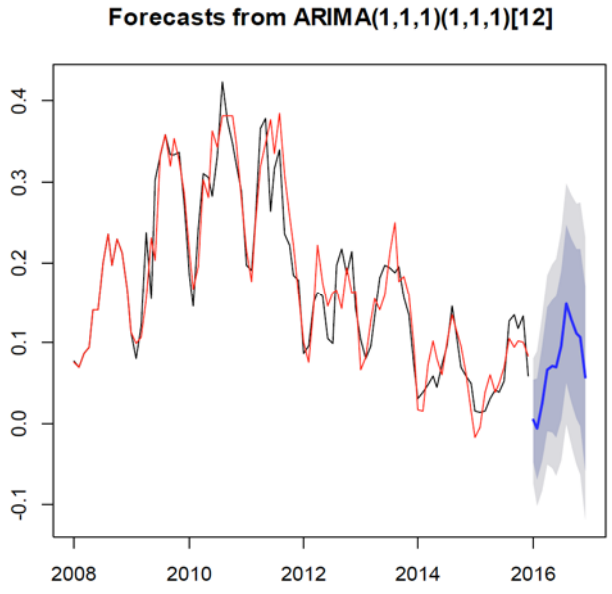

Figure 12: Black line: Observed Time Series, Red line: Fitted Time Series, Blue line: Forecast Time Series, Dark Grey region: 80% Confidence Interval for Forecast, Light Grey region: 95% Confidence Interval for Forecast.

Knowledge of disease-burden & trends, as well as intervention-effects [39] is mandatory to implement an effective malaria-control programme. After fitting a SARIMA time-series model on Malaria SPR data collected at the KMC Ward-level Malaria Clinics (cum-Treatment-Centres) over a six-year period from January 2008 to December 2015, a forecasting of Malaria SPR for the year 2016 for a part of Kolkata-population, has been tried.

A unique feature of this model is that it is based on SPR, rather than Malaria-Incidence Rate, as is generally used in many concurrent studies. Malaria Incidence Rate (WHO) is Number of New Malaria Cases out of total population per week in each area, i.e., number of confirmed malaria cases per person-time [40], which points out whether Malaria is increasing and necessary interventions when Malaria increases. As to monitor and to evaluate a general malaria-case, i.e., “fever with malarial parasitaemia”, in a control programme all patients are necessary, who require anti-malarial treatment [40]. Thus, in longitudinal studies, the total population of a geographical area is scrutinized for finding out of suspected malaria-cases (from all fever cases) and subjecting them to a diagnostic-test (having high sensitivity and specificity) are necessary, demanding considerable human and monitory resources. It is rarely followed in usual malaria surveillance programme.

Malaria-incidence is thus usually estimated based on the number of reported malaria-cases recorded in HMIS (Health Management Information System) of a city, state and country. When one tries to determine control-interventions, based on malaria-incidence rate using HMIS-data, different risk-factors are noticed leading to high bias and confounding factors, viz., a significant lag-time before their availability, incomplete reporting, temporal variations in reporting, varied utilization of health care services, lack of a proper denominator (for countries varying in population), and genuineness / accuracy of laboratory-confirmation of these data [38]. However proper statistical procedures can improve the accuracy-factors of derived estimates of malaria-incidence.

The nature of malaria transmission is usually estimated using different indicators such as ABER, API, SPR and Incidence of malaria [42-46]. Annual malaria incidence includes numbers of laboratory-confirmed malaria cases in a year as numerator and local population as denominator. As census is only

Table 8: Forecast for the Year 2016

Time	Point forecast	80% Confidence interval	95% Confidence interval
Jan 2016	0.012948	(-0.043790, 0.069686)	(-0.073826, 0.099721)
Feb 2016	-0.004012	(-0.073218, 0.065194)	(-0.109853, 0.101829)
Mar 2016	0.038264	(-0.040391, 0.116919)	(-0.082028, 0.158556)
Apr 2016	0.092490	(0.005537, 0.179444)	(-0.040494, 0.225474)
May 2016	0.103698	(0.009188, 0.198208)	(-0.040843, 0.248239)
Jun 2016	0.102282	(0.000778, 0.203786)	(-0.052955, 0.257518)
Jul 2016	0.128294	(0.020249, 0.236339)	(-0.036947, 0.293535)
Aug 2016	0.163900	(0.049687, 0.278112)	(-0.010774, 0.338573)
Sep 2016	0.139362	(0.019298, 0.259426)	(-0.044260, 0.322983)
Oct 2016	0.122742	(-0.002900, 0.248385)	(-0.069412, 0.314897)
Nov 2016	0.109085	(-0.021902, 0.240071)	(-0.091242, 0.309411)
Dec 2016	0.059436	(-0.076697, 0.195570)	(-0.148762, 0.267635)

carried out once 10 years in India, local population size may be under or overestimated. Moreover, population movement is enormous, due to economic reasons, in a city like Kolkata. Thus Malaria-incidence may be inaccurate due to limited health-care resources [43] or wrong population-size [47]. To plan public-health interventions, estimation of malaria burden is a necessary requirement, for which SPR may be used as a surrogate measure for malaria-incidence [43,45,48,49], to express the quantum of malaria endemicity in any region [47], plus to identify high risk areas [50]. In annual reporting system, Annual Parasite Index (i.e., malaria-incidence rate) is the principal monitoring indicator in malaria control programme in Kolkata for several years [50,52]. The changes in malaria incidences may be estimated from the SPR trend as well [43]. Studies showed SPR steadily decreased with the decline in malaria-incidence [44,48], while some showed that the annual parasite index (API) increased, but SPR-increase remained somewhat slow at the same level [44]. In our study SPR-trend has been subjected to statistical analysis.

Statistical Analysis points out that magnitude of the seasonal variation increases at the same rate as the yearly mean-values, indicating that a multiplicative seasonal model is appropriate. From the empirical analysis we find that between projected Model and data fitted in that Model, there are few apparent dissimilarities, for example, the usual bimodal rise of Malaria in 2008-2013 graph is modified to a unimodal peak in 2014 (somewhat like 2010 data, but of much lowered SPR-value). The plateau-area of highest malaria-incidences, is converted to a single incidence peak. Instead of stepped monthly rise with stepped monthly fall of malaria-incidences, there is sharp rise and sharp fall in malaria-incidences in this study-population (some what like 2010 data). High-degree seasonal pattern (mainly, climatological) is noted in the data all throughout the period, but clearly there is a upward trend from 2008 - 2010 followed by a downward trend gradually since 2011 up to 2014, which may be as a result of different anti-malaria and anti-mosquito interventions and malaria control strategy, adopted by The K.M.C. from 2011 onwards, plus many other covariates, influencing the Model. All these factors combined are subjected to a multivariate study-analysis, which is going on, to better understand the influences of those factors on the Forecasting system, as adopted in this study.

Interactive Agent-Host-Environment Factors of Epidemiological Triad are under observation to explain the causes of SPR changes, though it is assumed that the study population did not change during these years from 2008 to 2016. However, societal changes due to political changes during these years obviously have impacts on the result. This study is a significant step to attain a base-level study for the future works ahead.

REFERENCES

[1] "The Global Malaria Action Plan For a Malaria Free World" - The Roll Back Malaria Partnership, World Health Organization, Geneva, Switzerland.

[2] Pierre-Louis A-M, Qamruddin J, Espinosa I, Challa S. The Malaria Control Success Story. Chapter 23: 417-432.

[3] Gallup JL, Sachs J. The Economic Burden of Malaria. American Journal of Tropical Medicine and Hygiene 2001; 64(Suppl. 1): 8596.

[4] Murray CJL, Lopez AD. Evidence-based health policy lessons from the Global Burden of Disease Study. Science 1996; 274: 740743. http://dx.doi.org/10.1126/science.274.5288.740

[5] Murray CJL, Lopez AD. The Global Burden of Disease 1990-2020: alternative projections of mortality and disability by cause for eight regions. Lancet 1997; 349: 14981504.

[6] Breman JG. The ears of the hippopotamus: manifestations, determinants and estimation of the malaria burden. Am J Trop Med Hyg 2001; 64(Suppl 1): 11.

[7] Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of Plasmodium falciparum malaria. Nature 2005; 434: 214217.

[8] Kondrachine AV. Malaria in WHO Southeast Asia Region. Indian J Malariol 1992; 29: 129160.

[9] Mendis K, Sina BJ, Marchesini P, Carter R. The neglected burden of Plasmodium Vivax malaria. Am J Trop Med Hyg 2001; 64: 97106.

[10] Nihei N, Hashida Y, Kobayashi M, Ishii A. Analysis of malaria endemic areas on the Indochina Peninsula using remote sensing. Jpn J Infect Dis 2002; 55: 160-166.

[11] Klinkenberg E, HoekW, Amerasinghe FP. A malaria risk analysis in an irrigated area in Sri Lanka. Acta Trop 2004; 89: 215-225. http://dx.doi.org/10.1016/j.actatropica.2003.08.007

[12] Leonardo LR, Rivera PT, Crisostomo BA, Sarol JN, Bantayan NC, Tiu WU, Bergquist NR. A study of the environmental determinants of malaria and schistosomiasis in the Philippines using Remote Sensing and Geographic Information Systems. Parassitologia 2005; 47: 105-114.

[13] Zhou G, Sirichaisinthop J, Sattabongkot J, Jones J, Bjornstad ON, Yan G, Cui L. Spatio-temporal distribution of Plasmodium falciparum and P. vivax malaria in Thailand. Am J Trop Med Hyg 2005; 72: 256-262.

[14] Childs DZ, Cattadori IM, Suwonkerd W, Prajakwong S, Boots M. Spatiotemporal patterns of malaria incidence in northern Thailand. Trans R Soc Trop Med Hyg 2006; 100: 623-631. http://dx.doi.org/10.1016/j.trstmh.2005.09.011

[15] Ceccato P, Connor SJ, Jeanne I, Thomson MC. Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. Parassitologia 2005; 47: 81-96.

[16] Rogers DJ, Randolph SE, Snow RW, Hay SI. Satellite imagery in the study and forecast of malaria. Nature 2002; 415: 710-715. http://dx.doi.org/10.1038/415710a

[17] Abeku TA, Hay SI, Ochola S, Langi P, Beard B, de Vlas SJ, Cox J. Malaria epidemic early warning and detection in African highlands. Trends Parasitol 2004; 20: 400-405. http://dx.doi.org/10.1016/j.pt.2004.07.005

[18] Thomson MC, Connor SJ. The development of Malaria Early Warning Systems for Africa. Trends Parasitol 2001; 17: 438-445. http://dx.doi.org/10.1016/S1471-4922(01)02077-3

[19] Cox J, Abeku TA. Early warning systems for malaria in Africa: from blueprint to practice. Trends Parasitol 2007; 23: 243-246. http://dx.doi.org/10.1016/j.pt.2007.03.008

[20] Abeku TA, de Vlas SJ, Borsboom G, Teklehaimanot A, Kebede A, Olana D, van Oortmarssen GJ, Habbema JD. Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: a simple seasonal adjustment method performs best. Trop Med Int Health 2002; 7: 851-857. http://dx.doi.org/10.1046/j.1365-3156.2002.00924.x

[21] Hay SI, Were EC, Renshaw M, Noor AM, Ochola SA, Olusanmi I, Alipui N, Snow RW. Forecasting, warning, and detection of malaria epidemics: a case study. Lancet 2003; 361: 1705-1706. http://dx.doi.org/10.1016/S0140-6736(03)13366-1

[22] Thomson M, Indeje M, Connor S, Dilley M, Ward N. Malaria early warning in Kenya and seasonal climate forecasts. Lancet 2003; 362: 580. http://dx.doi.org/10.1016/S0140-6736(03)14135-9

[23] Bouma MJ, Kaay HJ. The El Nino Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics? Trop Med Int Health 1996; 1: 8696. http://dx.doi.org/10.1046/j.1365-3156.1996.d01-7.x

[24] Archie CA Clements, Adrian G Barnett, Zhang W Cheng, Robert W Snow and Hom N Zhou: Space-time variation of malaria incidence in Yunnan province, China, Malaria Journal 2009; 8 180. http://dx.doi.org/10.1186/1475-2875-8-180

[25] Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 2005; 25: 1965-1978. http://dx.doi.org/10.1002/joc.1276

[26] Census of India [http://censusindia.gov.in/] website.

[27] Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW. The global distribution and population at risk of malaria: past, present, and future. Lancet Infect Dis 2004; 4(6): 327-336. http://dx.doi.org/10.1016/S1473-3099(04)01043-6

[28] Joshi PL, Chandra R, Bhattacharya M, Vaish HC. Validity of using slide positivity rate (SPR) in identification of high risk malarious segments in rural areas. The Journal of Communicable Diseases 1997; 29(1): 41-45.

[29] Ceesay SJ, Casals-Pascual C, Erskine J, Anya SE, Duah NO, Fulford AJ, Sesay SS, Abubakar I, Dunyo S, Sey O, *et al*. Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis. Lancet 2008; 372(9649): 1545-1554. http://dx.doi.org/10.1016/S0140-6736(08)61654-2

[30] O'Meara WP, Bejon P, Mwangi TW, Okiro EA, Peshu N, Snow RW, Newton CR, Marsh K. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. Lancet 2008; 372(9649): 1555-1562. http://dx.doi.org/10.1016/S0140-6736(08)61655-4

[31] Jensen TP, Bukirwa H, Njama-Meya D, Francis D, Kamya MR, Rosenthal PJ, Dorsey G. Use of the slide positivity rate to estimate changes in malaria incidence in a cohort of Ugandan children. Malaria Journal 2009; 8: 213. http://dx.doi.org/10.1186/1475-2875-8-213

[32] Kendall M, Stuart A. The Advanced Theory of Statistics, 1983; Vol. 3: Griffin. pp. 410-414.

[33] Chatfield C. The Analysis of Time Series: An Introduction (Fourth Edition ed.). New York, NY: Chapman & Hall 1989.

[34] Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis: Forecasting and Control, 4th Edition 2008.

[35] Aho K, Derryberry D, Peterson T. Model selection for ecologists: the worldviews of AIC and BIC. Ecology 2014; 95: 631-636.

[36] Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in Model Selection. Sociological Methods & Research 2004; 33: 261-304.

[37] Burnham KP, Anderson DR. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. New York, Springer-Verlag, 2002.

[38] Ljung GM, Box GEP. On a measure of lack of fit in time series models. Biometrika 1978; 65: 297-303. http://dx.doi.org/10.1093/biomet/65.2.297

[39] Breman JG, Holloway CN. Malaria surveillance counts. Am J Trop Med Hyg 2007; 77(6 Suppl): 36-47.

[40] World Health Organization: World malaria report. Technical document 2008. WHO/HTM/GMP/2008.1.

[41] Cibulskis RE, Bell D, Christophel EM, Hii J, Delacollette C, Bakyaita N, Aregawi MW. Estimating trends in the burden of malaria atcountry level. Am J Trop Med Hyg 2007; 77(6 Suppl): 133-137.

[42] Roberts DR, Laughlin LL, Hsheih P, Legters LJ. DDT, Global strategies, and a malaria control crisis in South America. Emerg Infect Dis 1997; 3: 295. http://dx.doi.org/10.3201/eid0303.970305

[43] Jensen TP, Bukirwa H, Njama-Meya D, Francis D, Kamya MR, Rosenthal P, Dorsey G. Use of the slide positivity rate to estimate changes in malaria incidence in a cohort of Ugandan children. Malar J 2009; 8: 213. http://dx.doi.org/10.1186/1475-2875-8-213

[44] MetzgerW, Giron A, Vivas-Martnez S, Gonzlez J, Charrasco A, Mordmller B, Magris M. A rapid malaria appraisal in the Venezuelan Amazon. Malar J 2009; 8: 291. http://dx.doi.org/10.1186/1475-2875-8-291

[45] Subbarao SK, Vasantha K, Raghavendra K, Sharma V, Sharma G. Anopheles culicifacies: siblings species composition and its relationship to malaria incidence. J Am Mosq Control Assoc 1988; 4: 29.

[46] Montanari R, Bangali A, Talukder K, Baqui A, Maheswary N, Gosh A, Rahman M, Mahmood A. Three case definitions of malaria and their effect on diagnosis, treatment and surveillance in Cox's Bazar district, Bangladesh. Bull World Health Organ 2001; 79: 648-656.

[47] Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW. The global distribution and population at risk of malaria: past, present, and future. Lancet Infect Dis 2004; 4: 327-336. http://dx.doi.org/10.1016/S1473-3099(04)01043-6

[48] Lee PW, Liu CT, Rampao H, do Rosario V, Shaio MF. Pre-elimination of malaria on the island of Prncipe. Malar J 2010; 9: 26. http://dx.doi.org/10.1186/1475-2875-9-26

[49] Roy SB, Sarkar RR, Sinha S. Theoretical investigation of malaria prevalence in two Indian cities using the response surface method. Malar J 2011; 10: 301. http://dx.doi.org/10.1186/1475-2875-10-301

[50] Joshi PL, Chandra R, Bhattacharya M, Vaish HC. Validity of using slide positivity rate (SPR) in identification of high risk malarious segments in rural areas. J Commun Dis 1997; 29: 41.

[51] China Malaria Elimination Plan [http:/ / www.moh.gov.cn/ publicfiles/ business/ htmlfiles/ mohjbyfkzj/ s3593/ 201005/ 47529.htm] website

[52] China National Malaria Office of Global Fund: Management and technique program in high malaria transmission areas in China, Shanghai 2003.

Received on 20-08-2015 Accepted on 10-11-2015 Published on 08-06-2016

http://dx.doi.org/10.6000/1929-6029.2016.05.02.6