ECOLOGICAL MODELLING

Ecological Modelling 177 (2004) 179–192

Effects of environmental fluctuations on the occurrence of malignant malaria—a model based study

J. Chattopadhyay[a,*], R.R. Sarkar[a], S. Chaki[a], S. Bhattacharya[b]

[a] Embryology Research Unit, Indian Statistical Institute, 203, B.T. Road, Kolkata 700108, India
[b] Department of Zoology, Asutosh College, 92 S.P. Mukherjee Road, Kolkata 700026, India

Received 10 January 2003; received in revised form 15 January 2004; accepted 10 March 2004

Abstract

This paper aims to reveal the hidden structure of the disease curve involved in the available time series data on the number of *Plasmodium falciparum* (malignant malaria) cases. At first we try to find out the important factors which impart greater influence on the occurrence of malignant malaria and then try to fit a suitable regression model. We have taken into consideration the environmental factors like temperature, humidity and rainfall along with the total number of malaria cases as the independent variables in our first regression model. Next we have incorporated a social factor viz. monthly expenditure on Malaria Control Programme by Kolkata Municipal Corporation (KMC), West Bengal, India into our model. A basic mathematical model on malaria with environmental fluctuations is considered to compare the qualitative nature of the proposed regression model. It is observed that Macdonald's stability index takes higher value in the presence of stochastic fluctuations. Moreover, it is concluded that social factors may be used for programme implementation in the case of disease outbreak.

Keywords: Malaria; *P. falciparum*; Environmental fluctuations; Regression model; Basic model; Macdonald's stability index; Control

1. Introduction

The historical and epidemiological literatures provide informations on the infectious diseases of human communities and their effects on population abundance, social organization as well as on the unfolding patterns of historical events. The application of mathematics to the study of infectious disease appear to have been initiated by Berniulli, 1760. In recent years, analysis of mathematical models and comparisons with incidence data have uncovered fundamental mechanisms that control the dynamics and persistence of parasite infections (Bartlett, 1957; Anderson and May, 1991; Rohani et al., 1999). Using a theoretical approach, it has been possible to explore the relative benefits of different potential immunization strategies. Mathematical models have been useful for estimating a critical vaccination level that will eradicate an infection (Anderson and May, 1991; Agur et al., 1993). In recent works, there has been an emphasis on the application of control theory to epidemic models (Wickwire, 1977), the study of the spatial spread of diseases (Mollison, 1977; Cliff et al., 1983; Kallen et al., 1985), the investigation of the mechanisms underlying recurrent epidemic behaviour (Hethcote et al., 1981; Aron and Schwartz, 1984), the importance of heterogeneity in transmission (Anderson and May, 1986), and the extension of the threshold theory to encompass more complex deterministic and stochastic models (Whittle, 1955; Becker, 1978; Anderson and May, 1978, 1979; May and Anderson, 1979; Ball, 1983). Surprisingly, however, despite the current sophistication of the literature, the insights gained from theoretical work have, in general, had little impact on empirical approaches to epidemiological study and the design of public health policy. Therefore, a much greater emphasis must be placed on data oriented studies though the theoretical work play a role in the solution of practical problems in disease control and in the interpretation of observed trends.

The disease malaria has a global significance as a cause of human mortality and its role in the early beginnings of epidemic and endemic theory is prominent. Malaria in human is due to infection by one of the four protozoan species belonging to the genus *Plasmodium* namely *P. falciparum*, *P. vivax*, *P. malariae* and *P. ovale*. The most pathogenic species is *P. falciparum* which is a major cause of child mortality in many areas of the developing world. Among the several other epidemics, malaria is one of the major epidemics observed in India. For this, the major transitions are between regular cycles to irregular patterns (possibly chaotic epidemics) and also between regionally synchronized oscillations to complex epidemics (may be spatially incoherent). The recent epidemiological status of malaria in Kolkata Municipal Corporation (KMC) Area, West Bengal, India has been studied by Mukhopadhyay et al. (1997). Their study revealed that both mortality and morbidity due to malaria in the area of KMC showed increasing trend. Davis and Martin (1997) characterized post-treatment clearance of young forms of *P. falciparum* from the blood with three differential equation models, a linear decline, a linear and then logarithmic decline and the Michaelis–Menten kinetic equation, which were fitted to log-transformed parasite counts from 30 semi-immune patients with synchronous parasitaemias allocated to one of six antimalarial drug regimens. The earliest attempts to provide a quantitative understanding of the dynamics of malaria transmission were that of Ross, 1911, 1915. Macdonald (1957) added a layer of biological realism to these early models by his careful attention to interpretation and estimation of parameters. The value of mathematical studies to the design of malaria control programmes and the interpretation of observed epidemiological trends has been a topic of considerable controversy (Martini, 1921; Moshkovskii, 1950; Macdonald, 1957; Bruce-Chwatt and Glanville, 1973). So far our knowledge is concerned, no study has been done on the dynamics of malaria considering the real available data on the malignant malaria disease over time. But in reality a simple malaria model can explain both kinds of transition, prevalence and distribution of parasites. Aron and May (1982) observed that the basic models demonstrate the seasonal pattern in mosquito population density, if the total mosquito population varies seasonally with an amplitude that fluctuates randomly from year to year. In order to access the actual situation of malaria in the city of Kolkata, India, time series analysis of the morbidity, mortality etc. of malaria for this place from the available epidemiological data is very much necessary. Moreover, environmental stochasticity in terms of variability in different environmental factors (temperature, humidity, rainfall etc.) and social factors (expenditure on malaria control programme etc.), for this disease cannot be ignored. Although the basic model gives a good overview of the dynamics of malarial infection, particularly the basic factors that underlie "stable" and "unstable" malaria, many of its predictions are strikingly different from reality. Macdonald (1957) estimated the stability index to explain the epidemic outbreaks of malarial infection and Aron and May (1982) found rough estimates of Macdonald's stability index for several regions where malaria is indigenous. Still, there is need to find out proper estimation of Macdonald's stability index in a more realistic sense by considering environmental fluctuations in the system. Therefore, suitable model, and forecasting on the future occurrences of the disease based on the environmental and social factors and proper estimation of Macdonald's stability index are likely to be needed from the recurrent problems of malaria epidemic.

In the present study, we try to fit a suitable regression model by considering important environmental factors along with a social factor based on the available data. We compare the qualitative nature of the model with the basic malaria model under environmental stochasticity. Further, we propose a proper estimation of Macdonald's stability index for the system under environmental fluctuation, with the help of the technique developed by Sarkar et al. (2001) which is also used by Chattopadhyay et al. (2001). We also compare the value of Macdonald's stability index estimated by this technique with the value obtained by Aron and May (1982) for several regions where malaria is indigenous. Lastly, we propose a suitable value of Macdonald's stability index for Kolkata, India to reveal the realistic feature of the disease outbreaks and to chalk out a suitable eradication strategy.

2. Data analysis and observations

To observe the effect of environmental factors on the occurrence of the disease, at our disposal we have monthly data on the following variables for three consecutive years viz. 1999, 2000 and 2001.

Var 1. Total number of malaria cases
Var 2. Total number of *P. falciparum* cases (Pf. cases)
Var 3. Maximum temperature (°C)
Var 4. Minimum temperature (°C)
Var 5. Maximum humidity
Var 6. Minimum humidity
Var 7. Total rainfall (mm)

Also, to observe the effect of social factors, we have monthly expenditure data on the Malaria Control Programme by KMC for the years 2000 and 2001.

Now we are going to perform the regression analysis on this data set using the flowchart shown in the Appendix A. To build a suitable regression model, the total number of Pf. cases will be considered as the dependent variable and the other factors as the explanatory variables.

We first try to have an insight of the data and see whether any meaningful ideas about the interrelationships between the variables are available without going into more detailed analysis. For this purpose two-dimensional scatter diagrams are drawn. These are shown in the Figs. 1–6, where Var 2 represents the total number of Pf. cases.

From the scatter diagrams 1–6, we are not getting any linear relationship between the dependent and independent variables. Hence, after selecting the influential variables, we will try to refine our model by variable transformations.

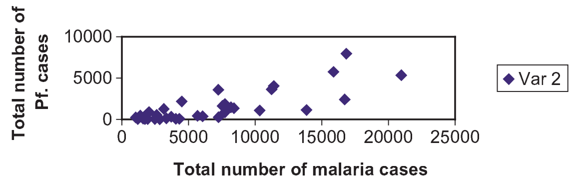


Fig. 1. Scatter diagram of Var 2 vs. Var 1.

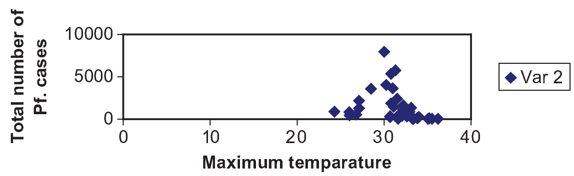


Fig. 2. Scatter diagram of Var 2 vs. Var 3.

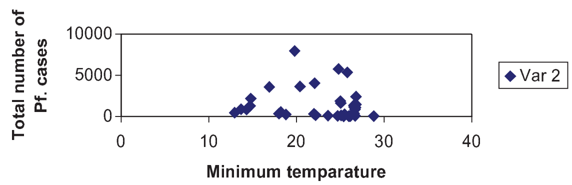


Fig. 3. Scatter diagram of Var 2 vs. Var 4.

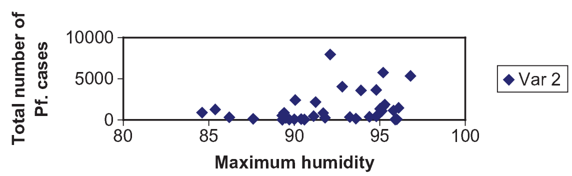


Fig. 4. Scatter diagram of Var 2 vs. Var 5.

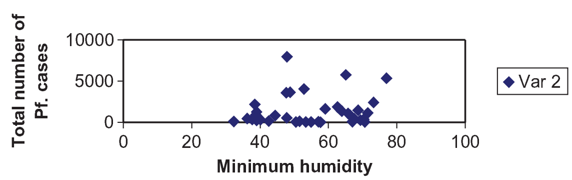


Fig. 5. Scatter diagram of Var 2 vs. Var 6.

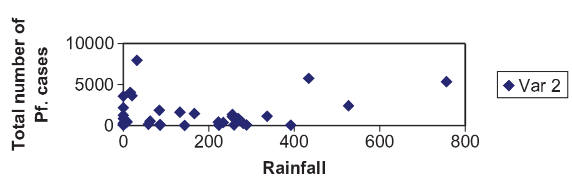


Fig. 6. Scatter diagram of Var 2 vs. Var 7.

3. Selection of influential variables and refinement of model

3.1. Variable selection

Let us consider a regression model in the form of

$$Y_i = \beta_0 + X_{1i} + \beta_2 X_{2i} + \cdots\cdots + \beta_{P-1} X_{P-1,i} + \varepsilon_i$$
$$\forall i = 1(1)n$$

it is suspected that not all X_i's are significant for the model. To find out the more useful explanatory variables we are considering three methods of variable selection namely,

1. Forward selection procedure
2. Backward elimination procedure
3. Stepwise regression procedure

3.1.1. Forward selection procedure

Regress Y on each $X_{j,}$ respectively. Choose j such that either r^2 is largest or β_j has the lowest P-value or the sum of squares due to error is the lowest. r^2, The coefficient of multiple determination is given by

$$R^2 = 1 - \frac{\text{SSE}}{\text{TSS}}$$

where SSE and TSS imply sum of squares error and total sum of squares, respectively.

Suppose X_j^* is selected. Then consider $P-2$ models

$$Y_i = \beta_0 + \beta_j^* \text{X}_{ji}^* + \beta_k X_{ki} \quad \forall k \neq j^*$$

for each of these models calculate , the partial F-statistics, $\{b_j/s(b_j)^2\}$, where b_j is the estimate of β_j.

If this value exceeds a pre-determined level (F-to enter), include $X_{j.}$ Next follow the similar procedure by considering X_j and X_j^* already present in the model.

3.1.2. Backward elimination procedure

Initially consider all the variables in the model. Choose j such that r^2 is smallest. Eliminate X_j from the model if the corresponding partial F-statistics are smaller than a pre-determined level (F-to delete).

3.1.3. Stepwise regression procedure

This is probably the most widely used variable selection method. It was developed to economize on computational efforts, as compared with the all-possible-regression approach, while arriving at a reasonably "good" subset of independent variables. Essentially, this search method develops a sequence of regression models, at each addition or deletion of X variable. The criterion for adding or deleting an X variable can be stated equivalently in terms of error sum of squares reduction, coefficient of partial correlation, or F-statistics.

3.1.4. Search algorithm

Neter et al. (1985) describes the stepwise regression search algorithm in terms of the F-statistic for the partial F-test.

1. The stepwise regression routine first fits a simple regression model for each of the $P-1$ potential X variables. For each simple regression model the F-statistic

$$F = \frac{\text{MSR}}{\text{MSE}}$$

where MSR is the regression mean square and MSE is the mean-square error for testing whether or not the slope is zero is obtained.

$$F_k = \frac{\text{MSR}(X_k)}{\text{MSE}(X_k)}$$

recall that $\text{MSR}(X_k) = \text{SSR}(X_k)$ measures the reduction in the total variation of Y associated with the use of the variable X_k. The X variable with the largest F-value is the candidate for first addition. If this F-value exceeds a pre-determined level, the X variable is added. Otherwise, the programme terminates with no X variable considered sufficiently helpful to enter the regression model.

2. Assume X_7 is the variable entered at step 1. The stepwise regression routine now fits all regression models with two X variables. Where X_7 is one of the pair. For each such regression model, the partial F-test statistic

$$F_k = \frac{\text{MSR}(X_k|X_7)}{\text{MSE}(X_7, X_k)} = \left\{\frac{b_k}{s(b_k)}\right\}^2$$

is obtained. This is the statistic for testing whether or not $\beta_k = 0$ when X_7 and X_k are the variables in the model. The X variable with the largest F-value is the candidate or addition at the second stage.

If this F-value exceeds a pre-determined level, the second X variable is added. Otherwise the programme terminates.

3. Suppose X_3 is added at the second stage. Now the stepwise regression routine examines whether any of the other X variables already in the model should be dropped. For our illustration, there is at this stage only one other X variable in the model, X_7, so that only one partial F-test statistic is obtained

$$F_7 = \frac{\text{MSR}(X_7|X_3)}{\text{MSE}(X_3, X_7)}$$

tat later stages, there would be a number of these F-statistics, for each of the variables in the model besides the one last added. The variable for which this F-value is smallest is the candidate for deletion. If this F-value falls below a pre-determined limit, the variable is dropped from the model; otherwise, it is retained.

4. Suppose X_7 is retained so that both X_3 and X_7 are now in the model. The stepwise regression routine now examines which X variable is the next candidate for addition, then examines whether any of the variables already in the model should now be dropped, and so on until no further X variables can either be added or deleted, at which point the search terminates.

It should be noted that the stepwise regression algorithm allows an X variable, brought into the model at an earlier stage, to be dropped subsequently if it is no longer helpful in conjunction with variables added at later stages. We have extensively used the SAS Software (http://support.sas.com) to perform the regression analysis.

Results of the forward selection procedure:
Step 1. Variable 1 (total number of malaria cases) entered
Model r^2 = 0.6515 and $C(P)$ = 28.1410

Step 2. Variable 6 (minimum humidity) entered
Model r^2 = 0.8327 and $C(P)$ = 5.1055

No other variable met the 0.05 significance level for entry into the model.

Results of the backward elimination procedure:
Step 0. All variables entered
Model r^2 = 0.8769 and $C(P)$ = 7.0

Step 1. Variable 3 (maximum temparature) removed
Model r^2 = 0.8769 and $C(P)$ = 5.006

Step 2. Variable 5 (maximum humidity) removed
Model r^2 = 0.8669 and $C(P)$ = 4.3803

Step 3. Variable 6 (minimum humidity) removed
Model r^2 = 0.8587 and $C(P)$ = 3.5191

All variables left (total number of malaria cases and rainfall) in the model are significant at the level 0.05.

Results of the stepwise regression procedure:
Step 1. Variable 1 (total number of malaria cases) entered
Model r^2 = 0.6515 and $C(P)$ = 28.14

Step 2. Variable 6 (minimum humidity) entered
Model r^2 = 0.8327 and $C(P)$ = 5.1055

All variables left in the model are significant at the 0.05 level. No other variable met the 0.05 significance level for entry into the model.

Deleting some independent variables usually biases the estimates of the parameters left in the model. Suppose out of k, we retain first $P-1$ independent variables. So $b_0, b_1, \ldots, b_{p-1}$ are biased implying that the predicted values are also biased. One measure of this bias is called Mallows' $C(P)$. The key property for it's application is that, if a new model does not lead to much bias in the predicted values, then

$$E[C(P)] \approx P$$

thus, when the $C(P)$ values for all possible regression models are plotted against P, those models with little bias will tend to fall near the line $C(P) = P$. Models with substantial bias will tend to fall considerably above this line.

$$C(P) = \frac{\text{SSEP}_P}{\text{MSE}(X_1, \ldots, X_{P-1})} - (n - 2P)$$

where SSE_P is the error sum of squares for the fitted subset regression model with $P-1$ explanatory variables.

Thus, by using the above three variable selection procedures we have obtained the following variables to be incorporated into our model

1. Total number of malaria cases
2. Minimum humidity
3. Rainfall

It is to be noted that all the analyses done so far are based on the available data for the years 1999 and 2000. We have considered total number of Pf. cases as the dependent variable. In a separate analysis, taking total number of malaria cases as the dependent variable it is found that there is a fair amount of positive correlation (0.30428) between minimum temparature and total number of malaria cases. Also, the variable selection procedure supports this fact. Hence it is not unreasonable to incorporate minimum temperature into our present model for total number as affecting the total number of malaria cases in the area indirectly affects the total number of Pf. cases.

So we are now in a position to build a regression model for the total number of Pf. cases by taking the following explanatory variables

1. Total number of malaria cases
2. Minimum temparature
3. Minimum humidity
4. Rainfall

3.2. Model refinement

Simple Linear regression may not provide a good fit. So we try to investigate curvature and interrelations between the dependent and selected explanatory variables more fully.

3.2.1. Transforming the response with one dependent variable

A good transformation should make residuals smaller. We now consider different transformation of one or both of the original variables before carrying out the regression analysis. Simple transformations of either the dependent variable Y or the independent variable X, or of both, are often sufficient to make the simple linear regression model appropriate for the transformed data. Let us now present a graphical method for getting appropriate transformations.

Let us divide the range of the independent variable into three portions, making a good compromise between getting an equal number of data points in each portion and making the three portions roughly equal. For each of the three sets of data points thus created,

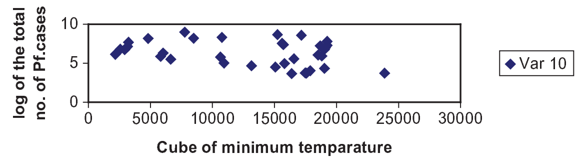


Fig. 7. Scatter diagram for transformed variables (log of total no. of Pf. cases vs. cube of minimum temperature).

find a point (which may or may not be one of the data points) which is a good representative of the set. For each set a good choice is the point whose co-ordinates are the medians of the x and y values for the points in the set. We now find the slope of the line joining the first two points (going from left to right) and the slope of the line joining the last two points. If these two are equal, then the data points should describe a straight line. If not, the middle of the three points will be below (the convex case) or above (the concave case) the line joining the other two. Then we follow the flowchart shown below.

$-1/y^2$	$\Uparrow$	x^5
$-1/\sqrt{y}$	if convex up the ladder	x^4
$\log y$		x^3
$\sqrt{y}$		x^2
y	$\Leftarrow$ We are here $\Rightarrow$	x
y^2		$\sqrt{x}$
y^3	if concave down the ladder	$\log x$
y^4	$\Downarrow$	$-1/\sqrt{x}$

starting from the linear relationships between the unchanged variables we have found that a fair amount of linear relationships exist between the logarithm of the total number of Pf. cases and the cubes of minimum temperature, minimum humidity and rainfall. These relationships are clearly visible from the Figs. 7–9, which show the scatter diagram for the transformed variables. After comparing these figures with the figures (Figs. 3, 5 and 6) for unchanged variables, we

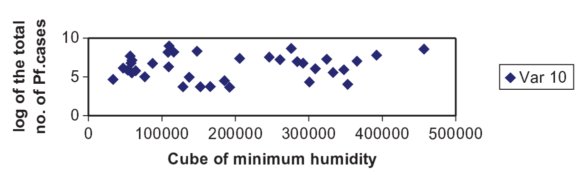


Fig. 8. Scatter diagram for transformed variables (log of total no. of Pf. cases vs. cube of minimum humidity).

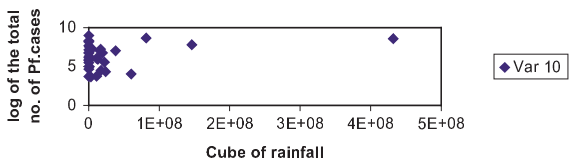


Fig. 9. Scatter diagram for transformed variables (log of total no. of Pf. cases vs. cube of rainfall).

decided to incorporate the transformed variables into our model.

In Figs. 7–9, Var 10 represents the logarithm of the total number of Pf. cases.

4. The proposed regression model

So far our analysis has been based on the monthly figures for the years 1999 and 2000.We have not considered the social factor in our model. The result of the regression procedure is given in Table 1.

Now we incorporate the social factor into our regression model. Since there is a high positive correlation between the total number of malaria cases and total number of Pf. cases, in order to restore this interrelationship we transform the former to it is logarithmic form and consider it in the model. Again it is quite reasonable to assume that there is a strong negative correlation between the total number of Pf. cases and the monthly expenditure on the Malaria Control Programme. In our former model we have transformed the original dependent variable into its logarithmic form. The Figs. 10 and 11 reveal that the linear relationship is stronger for the log-transformed monthly expenditure than for the monthly expenditure with the log of the total number of Pf. cases.

Thus our final regression model will be based on the following transformed variables

Dependent variable:	Logarithm of the total number of Pf. cases
Independent variables:	1. Logarithm of the total number of malaria cases 2. Cube of minimum temperature 3. Cube of minimum humidity 4. Cube of rainfall 5. Logarithm of the monthly expenditure

5. Results and comparison with basic mathematical model

The results of the regression procedure of our final model is given in Table 2.

We observe that the model r^2 has been increased from 0.8436 in the former model to 0.9265 in the final model which implies that the present model is better. Fig. 12 shows the residuals against the time points.

Fig. 13 shows the total number of Pf. cases for different months during the period 1999–2001. From the available monthly expenditure data on the Malaria

Table 1
The regression procedure

Source	DF	SS	MS	F-value	Pr > F
Analysis of variance					
Model	4	48.83468	12.20867	25.61	<0.0001
Error	19	9.05712	0.47669		
Corrected total	23	57.89180			
Variable	DF	Parameter estimate	Standard error	t-value	Pr > |t|
Parameter estimates					
Intercept	1	−5.49890	2.02682	−2.71	0.0138
Tmcases_log	1	1.62940	0.25200	6.47	<0.0001
Mintempcube	1	−2.09559E-8	0.00000268	−0.01	0.9938
Minhumcube	1	−0.00015395	0.00003769	−4.08	0.0006
Raincube	1	1.38684E-9	2.104503E-9	0.66	0.5178

Dependent variable: logarithm of the total number of Pf. cases. r^2 = 0.8436. Adjusted r^{2*} = 0.8106. *Adjusted r^2 takes into account the number of parameters in the model through degrees of freedom, which r^2 does not. Adjusted $r^2 = 1 - (\text{MSE}(\text{TSS}/n - 1))$.

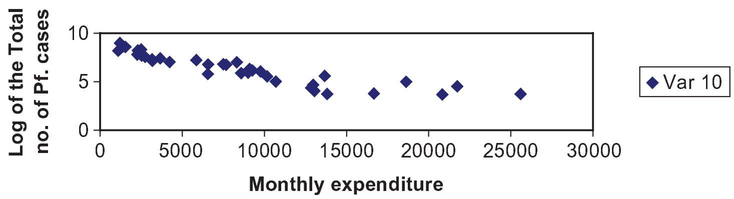


Fig. 10. Scatter diagram for transformed variables (log of total no. of Pf. cases vs. monthly expenditure).

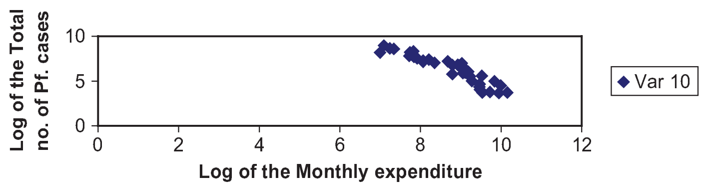


Fig. 11. Scatter diagram for transformed variables (log of total no. of Pf. cases vs. log of monthly expenditure).

Table 2
The regression procedure

Source	DF	SS	MS	F-value	Pr > i
Analysis of variance					
Model	4	80.65554	16.13111	75.65	<0.0001
Error	19	6.39660	0.21322		
Corrected total	23	87.05214			
Variable	DF	Parameter estimate	Standard error	t-value	Pr > $\|t\|$
Parameter estimates					
Intercept	1	8.04166	3.18760	2.52	0.0172
Tmcases_log	1	0.86017	0.22224	3.87	0.0005
Mintempcube	1	−0.00009258	0.00002382	−3.89	0.0005
Minhumcube	1	8.084269E-7	0.00000155	0.52	0.6057
Raincube	1	−6.6386E-10	1.39868E-9	−0.47	0.6385
Monthlyexp_log	1	−0.90353	0.19238	−4.70	<0.0001

Dependent variable logarithm of the total number of Pf. cases. r^2 = 0.9265. Adjusted r^2 = 0.9143.

Control Programme by KMC, we try to get hold of two representative figures for the year 2000 and 2001. They (Var 8) are shown on Fig. 13 by two spots. It is clearly visible that as soon as the social factor, under consideration comes into play, the occurrence of malignant malaria in the Kolkata Metropoliton Corporation Area had been considerably reduced. Also it is to be noted that the number of Pf. cases are small corresponding to those months in which more money had been spent on the control programme. So it is recommended to keep a steady balance in spending money throughout the months of a year in order to

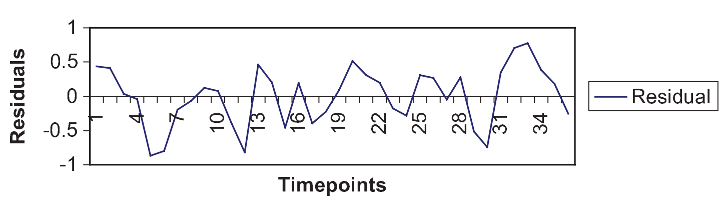


Fig. 12. Residual plot.

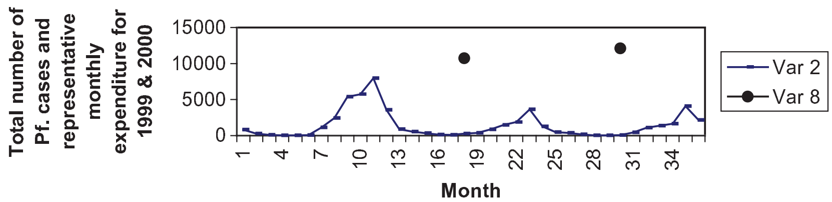


Fig. 13. Total number of Pf. cases for different months during the period 1999–2001.

reduce the brisk and slack movement in the curve of the total number of Pf. cases.

We have reached the above conclusion by performing statistical analysis on the available data. Fig. 14 shows the predicted number of Pf. cases for different months during the period 1999–2001. We have divided the graph into two parts. The first part (for the year 1999, i.e. first 12 time points) shows the figure without considering the social factor, while the second part (for the years 2000 and 2001, i.e. the remaining 24 time points) incorporates the social factor.

We have already mentioned in the introduction that fluctuations in mosquito population change the pattern of outbreak of the disease. In our study, so far we have proposed a suitable regression model based on the available epidemiological data and considered the variations in environmental as well as social factors to demonstrate a more realistic scenario behind the disease outbreak.

5.1. Comparison with a mathematical model

Now we are in a position to observe the similarity between the proposed model and the known basic mathematical models with environmental fluctuations. The earliest attempt to provide a quantitative understanding of the dynamics of malaria transmission was that of Ross (1911, 1915). The basic deterministic model which incorporates the interaction between the infected human hosts and the mosquito vector population may be written in the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma y(1 - x) - rx \tag{1}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = acx(1 - y) - \mu y \tag{2}$$

where x and y are the proportions of the human and female mosquito populations that are infected, a is the bite rate of a single mosquito, c is the proportion of bites by susceptible mosquitoes on infected people that produce a patent infection, r is the individual recovery rate per human, μ is the individual death rate for mosquitoes and $\sigma = abM/N$, where N and M are the (constant) sizes of the human and female mosquito populations respectively and b is the proportion of infected bites that produce an infection. In this case, small changes in mosquito density or the biting rate are more likely to result in substantial changes in the proportion of humans infected. This is the essence of Macdonald (1957) conclusion that ($ac/\mu = m$, say) is an index of stability; in areas where mosquito vectors bite humans relatively often and have relatively long life spans, this index is high and malaria tends to be endemic (Macdonald's stable malaria); conversely, where mosquitoes bite on humans less often and have shorter life spans, the index is low and malaria tends to be subject to epidemic outbreaks (Macdonald's unstable malaria). But the situ-

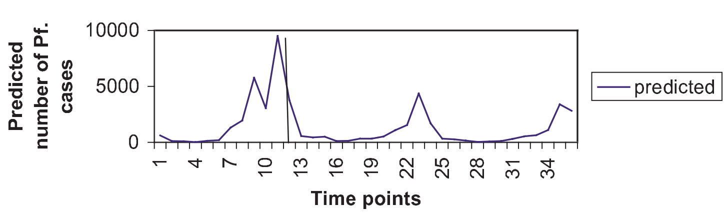


Fig. 14. Predicted number of Pf. cases.

ation depicted in this study was purely deterministic and the effects of environmental fluctuations have not been considered. To observe the environmental variability in the abundance of mosquito populations, we therefore modify the basic model by incorporating a random fluctuation in the form of colour noise, which follow the Ornstein–Uhlenbeck process (Uhlenbeck and Ornstein, 1954), in the second equation (which describes changes in the proportion of mosquitoes infected) of the basic model. The modified model can be written in the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma y(1-x) - rx \tag{3}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = acx(1-y) - \mu y + \eta(t)y \tag{4}$$

where the perturbation coefficient $\eta(t)$ is the Ornstein–Uhlenbeck process. The mathematical expectation and correlation function of the process $\eta(t)$ are given by

$$\langle \eta(t) \rangle = 0,\ \langle \eta(t_1)\,\eta(t_2) \rangle = \varepsilon\,\delta_0 \exp(-\delta_0|t_1 - t_2|) \tag{5}$$

where ε, $\delta_0 > 0$ are respectively the intensity and the correlation time of the noise and $\langle . \rangle$ represents averages over the ensemble of the stochastic process. The correlation function η(t) is the solution of the stochastic differential equation

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -\delta_0\,\eta + \delta_0\sqrt{(2\varepsilon)}\frac{\mathrm{d}\omega}{\mathrm{d}t} \tag{6}$$

where $\xi(t) = \mathrm{d}\omega/\mathrm{d}t$ denotes the standard zero mean Gaussian white noise characterized by

$$\langle \xi(t) \rangle = 0,\ \langle \xi(t_1)\,\xi(t_2) \rangle = \delta(t_1 - t_2) \tag{7}$$

with $\delta(t)$ as the Dirac delta function. Recently, Sarkar et al. (2001) developed a method to estimate the optimal values of the parameters and the safe region. They have solved the stochastic differential equations involved in the system, applied the idea of Tchebycheff's inequality to obtain a tolerance interval and minimized the deviations of the populations about the equilibrium level for an eco-epidemiological model of Tilapia–Pelican populations proposed by Chattopadhyay and Bairagi (2001). Chattopadhyay et al. (2001) successfully used this technique to estimate the inaccessible parameters in a plant-herbivore-parasitoid system under environmental stochasticity. In this study, we used their technique to obtain the critical value of the Macdonald's stability index (m) and observed that under the effect of environmental fluctuation, the Macdonald's stability index has the following realistic value

$$m = \min\left(\frac{r + \sqrt{r + 4\sigma^2}}{2\sigma}, \frac{6\sigma(r+\sigma) + 2\sigma - r}{3\sigma}\right) \tag{8}$$

To validate our analytical result, we consider the values of the parameters used by Aron and May (1982) in simulation of the basic model with a large reproductive rate and variable mosquito density. We observed that for $a = 20$ per year, $\mu = 50$ per year, $r = 4$ per year, $N = 20$, $M = 50$, $b = 1$ and $c = 0.85$, the Macdonald's stability index (m) is 0.34, when there is no environmental fluctuation. But the estimation of Macdonald's stability index under environmental fluctuation by our method is $m = 1.0402$, which is much higher than the deterministic situation. This depicts the fact that malaria becomes endemic (Macdonald's stable malaria) though previously estimated as epidemic outbreaks (Macdonald's unstable malaria). It is interesting to note that this technique of estimation will help us to provide sufficient information regarding a good estimate of Macdonald's stability index for several regions where malaria is indigenous (see, Table 14.6, p. 398, Anderson and May, 1991).

Now, to reveal the realistic feature of the outbreaks of malaria disease in Kolkata, India, we have compared the qualitative nature of the disease curve for our proposed regression model (obtained from the available data) with the simulated diagram (see, Fig. 15) obtained from the basic model under environmental fluctuation. For this simulation we consider the default values of the system parameters as $a = 2$ per year, $\mu = 4$ per year, $r = 3$ per year, $N = 3$, $M = 5$, $b = 2$ and $c = 0.95$. We observe that when there is no environmental stochasticity, then the Macdonald's stability index (m) is 2.25 and under the effect of environmental fluctuation the value of m is 1.633, which is much lower than the deterministic situation. This reveals the fact that malaria is epidemic in this area while considering the realistic scenario, whereas the deterministic estimation shows that malaria is endemic. Thus, there is a need of proper programme implementation for the eradication of malarial outbreaks.

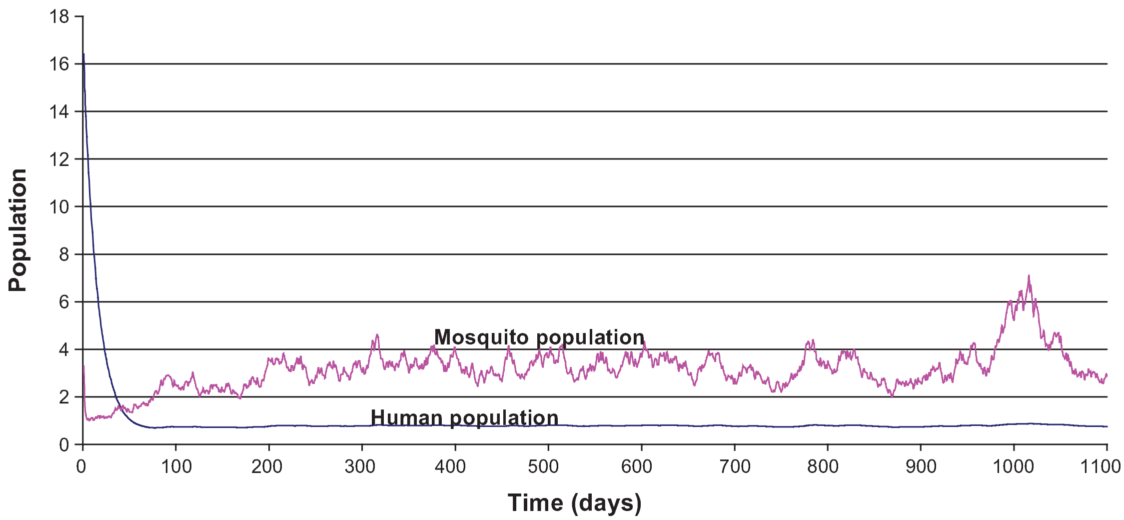


Fig. 15. Numerical simulation of the basic model under environmental fluctuation (see text for the default values of the parameters).

6. Conclusion

Epidemiological research on micro-parasitic infection is largely based on two distinct measures of parasite abundance within communities of people. The first of these is the incidence of infection or disease. The second measure is the prevalence of infection or disease. The measurement of incidence or prevalence is often based on the stratification of the population under study with respect to a variety of factors such as age, sex, social factors, environmental variability etc. Malaria, particularly falciparum malaria, is a major cause of morbidity and mortality throughout many regions of the world. The quantitative understanding of the dynamics of malaria transmission under the effect of environmental and social variations is therefore an utmost need in this context. In this paper we have tried to investigate the realistic features behind the disease outbreaks based on available data. We have (1) proposed a suitable regression model considering the important environmental and social factors. The results establish the significant dependence of total number of malaria cases on only minimum humidity and rainfall. We have obtained this result based on the data available from KMC, Kolkata, India. It may not be true in the world scenario, thus rigorous studies in this particular aspect are urgently needed. But we believe the techniques adopted here may be applicable in global perspective. We have also tried (2) to compare the qualitative nature of our proposed model with the basic mathematical model under environmental stochasticity. Further, we have proposed (3) a suitable measure for the estimation of Macdonald's stability index of the basic model under environmental fluctuations and compared this value with the value obtained by Aron and May (1982). It has been observed that Macdonald's stability index for Kolkata Municipal Corporation (West Bengal, India) is much lower due to environmental and social factors, which depicts the epidemic outbreak (Macdonald's unstable malaria). This prediction helps us to control the disease outbreak through proper implementation of a vaccination or the immunization programme. Therefore, quantitative understanding for the estimation of parameters is very important in this context to design proper malarial control programmes. Our study on estimation of parameters for the population abundance under environmental fluctuation gives a much more realistic interpretation in this direction.

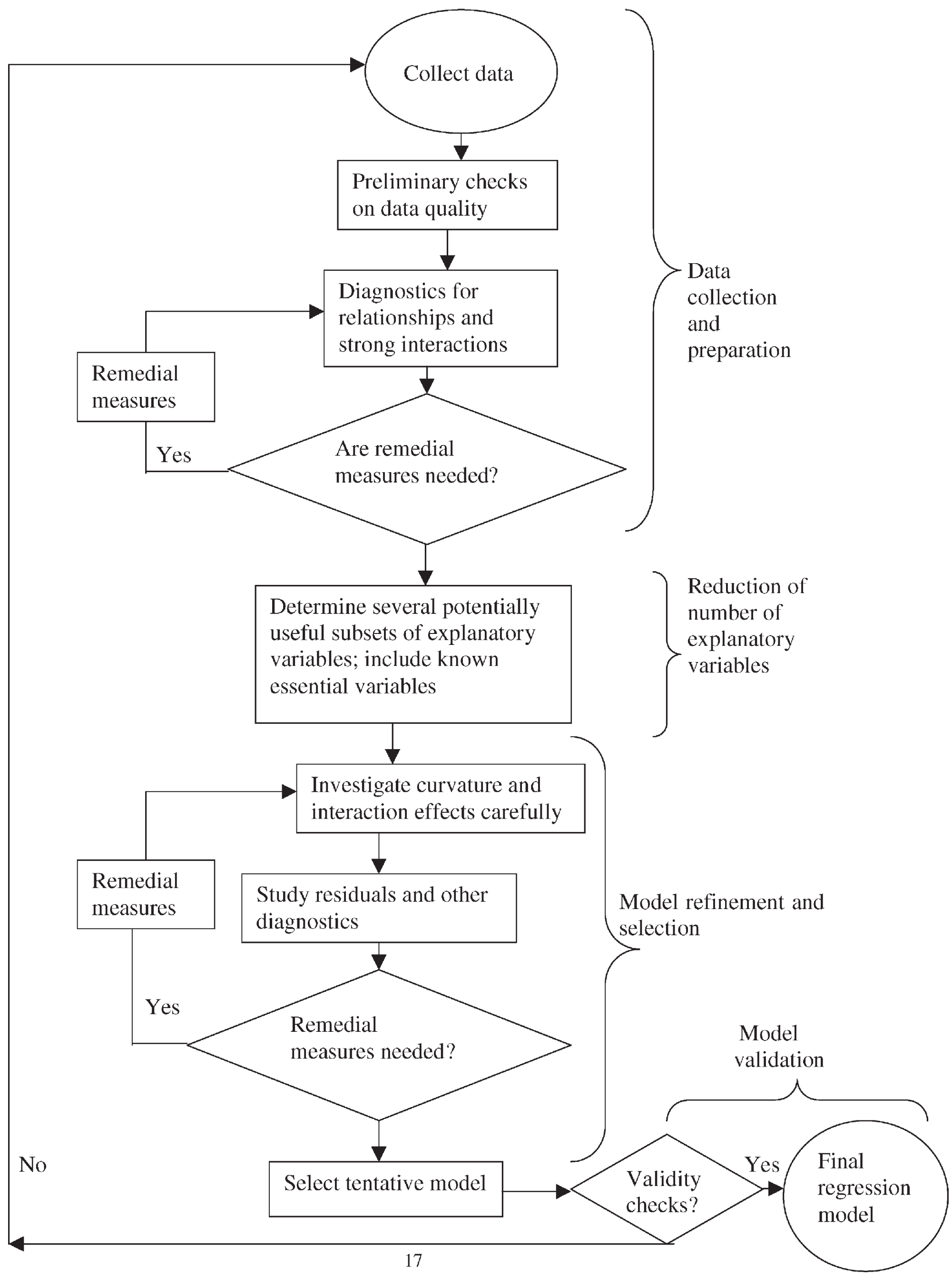
Before ending this article, we like to compare the advantages and disadvantages of our malaria model with other models of disease spreading. The development of mathematical models has been very useful in the study of dynamics of infectious diseases. In recent years, there has been significant change in modelling approaches and methods on terrestrial ecology, system ecology and epidemiology (see, Jørgensen, 1990, 1997, 2000). It is well observed that analysis of mathematical models and its comparison with incidence data have uncovered fundamental mechanisms that control the disease dynamics. So far as literature is concerned, several authors have tried to describe the dynamics of malaria and its control through different approaches, for example: basic differential equation models by Ross (1911, 1915); interpretation of parameters by adding a biological realisms to the early differential equation models by Macdonald (1957); observing the epidemiological trends to design the malarial control programme by Bruce-Chwatt and Glanville (1973), introducing seasonal variations in the mosquito population to the basic models by Aron and May (1982) and taking into account the spatial heterogeneity of environment by Torres-Sorando and Rodriguez (1997). However, there is a scarcity of models which take into account the environmental and social factors based on the available data.

In our model approach, we develop a regression model which describes the pattern of the malaria curve under important environmental and social influences. Next, we compare this model with the basic malaria model by introducing environmental stochasticity. This approach help us to estimate the Macdonald's stability index for the system under environmental fluctuation and reveal the realistic feature of the disease outbreaks as well as suitable control strategy. The above descriptions clearly indicate that the model and the results obtained in this paper are different from other models of disease spreading. We hope these results will be useful for different countries as well. However, we like to mention that the statistical model is based on linear regression, but non-linear regression model may give some other interesting results. We also like to mention that before applying our results for environmental fluctuating systems, comparison of the mathematical findings with real life social data is urgently needed.

Acknowledgements

The authors are grateful to the Kolkata Municipal Corporation (KMC), West Bengal, India for providing the data on malaria. The authors are also grateful to Mr. Joy Mukhopadhyay for helping us in data collection. The authors are grateful to the anonymous referees for their valuable comments and suggestions.

Appendix A



17

References

Agur, Z., Cojocaru, L., Mazor, G., Anderson, R.M., Damon, Y.L., 1993. Pulse mass measles vaccination across age cohorts. Proc. Nat. Acad. Sci. U.S.A. 90, 11698–11702.

Anderson, R.M., May, R.M. 1991. Infectious Diseases of Humans-Dynamics and Control. Oxford University Press, UK.

Anderson, R.M., May, R.M., 1986. The invasion, persistence and spread of infectious diseases within animal and plant communities. Phil. Trans. R. Soc. B314, 533–570.

Anderson, R.M., May, R.M., 1979. Population biology of infectious diseases: part 1. Nature 280, 361–367.

Anderson, R.M., May, R.M., 1978. Regulation and stability of host-parasite population interactions. I. Regulatory processes. J. Anim. Ecol. 47, 219–247.

Aron, J.L., Schwartz, I.B., 1984. Seasonally and period-doubling bifurcation in an epidemic model. J. Theor. Biol. 100, 665–679.

Aron, J.L., May, R.M. 1982. The population dynamics of malaria. In: Anderson, R.M. (Ed.), Population Dynamics of Infectious Diseases. Chapman and Hall, London, pp. 139–179.

Ball, F., 1983. The threshold behaviour of epidemic models. J. Appl. Prob. 20, 227–241.

Bartlett, M.S., 1957. Measles periodicity and community size. J. R. Stat. Soc. A120, 48–70.

Becker, N., 1978. The use of epidemic models. Biometrics 35, 295–305.

Berniulli, D. 1760. Essai d'une nouvelle analyse de la mortalite' cause'e par la petite ve'role et des advantages de l'inoculation pour la pre'venir. Me'm. Math. Phys. Acad. Roy. Sci. Paris, pp. 1–45.

Bruce-Chwatt, L.J., Glanville, V.J. 1973. Dynamics of tropical disease. Oxford University Press, London.

Chattopadhyay, J., Sarkar, R.R., Fritzsche-Hoballah, M.E., Turlings, T.C.J., Bersier, L.-F., 2001. Parasitoids may determine plant fitness-a mathematical model based on experimental data. J. Theor. Biol. 212, 295–302.

Chattopadhyay, J., Bairagi, N., 2001. Pelican at risk in Salton sea-an ecoepidemiological study. Ecol. Model. 136, 103–112.

Cliff, A.P., Haggett, P., Ord, J.K., 1983. Forecasting epidemic pathways for measles in Iceland: the use of simultaneous equation and logic models. Ecol. Dis. 2, 377–396.

Davis, T.M.E., Martin, R.B., 1997. Clearence of young parasite forms following treatment of *falciparum* malaria in humans: comparison of three simple mathematical models. Epidemiol. Inf. 119, 61–69.

Hethcote, H.W., Stech, H.W., van der Driessche, P., 1981. Nonlinear oscillations in epidemic models. SIAM J. Appl. Math. 40, 1–9.

Jørgensen, S.E., 1990. About your journal: ecological modelling. Ecol. Model. 50, 1–4.

Jørgensen, S.E., 1997. Ecological modelling by ecological modelling. Ecol. Model. 100, 5–10.

Jørgensen, S.E., 2000. 25 years of ecological modelling by ecological modelling. Ecol. Model. 126, 95–99.

Kallen, A., Arcuri, P., Murray, J.D., 1985. A simple model for the spatial spread and control of rabies. J. Theor. Biol. 116, 377–393.

Macdonald, G. 1957. The epidemiology and control of malaria. Oxford University Press, London.

Martini, E. 1921. Berechnungen und Beobachtungen. In: Epidemiologie und bekampfung der malaria. Gente, Humburg.

May, R.M., Anderson, R.M., 1979. Population biology of infectious diseases: part II. Nature 280, 455–461.

Mollison, D., 1977. Spatial contact model for ecological and epidemic spread. J. R. Stat. Soc. B39, 283–326.

Moshkovskii, Sh.D. 1950. Basic laws of the epidemiology of malaria. AMN, Moscow.

Mukhopadhyay, A.K., Karmakar, P., Hati, A.K., Dey, P., 1997. recent epidemiological status of malaria in Calcutta Municipal Corporation Area, West Bengal, India. J. Malariol. 34, 188–196.

Neter, J., Wasserman, W., Kutner, M.H. 1985. Applied Linear Statistical Models. Irwin, Homewood, Illinois.

Rohani, P., Earn, D.J.D., Finkenstadtt, B., Grenfell, B.T., 1999. Opposite patterns of synchrony in sympatric disease metapopulations. Science 286, 968–971.

Ross, R. 1911. The Prevention of Malaria, second ed. Murray, London.

Ross, R., 1915. Some a priori pathometric equations. Br. Med. J. 1, 546–547.

Sarkar, R., Chattopadhyay, J., Bairagi, N., 2001. Effects of environmental fluctuation on an eco-epidemiological model on Salton Sea. Environmetrics 12, 289–300.

Torres-Sorando, L., Rodriguez, D.J., 1997. Models of spatio-temporal dynamics in malaria. Ecol. Model. 104, 231–240.

Uhlenbeck, G.E., Ornstein, L.S. 1954. On the theory of Brownian motion. In: Wax, N. (Ed.), Selected Papers on Noise and Stochastic Process. Dover, New York, USA.

Whittle, P., 1955. The outcome of a stochastic epidemic-a note on Bailey's paper. Biometrika 42, 154–162.

Wickwire, K., 1977. Mathematical models for the control of pests and infectious diseases: a survey. Theor. Popul. Biol. 11, 182–238.