STATISTICS IN TRANSITION-new series, Summer 2012
Vol. 13, No. 2, pp. 405—418

TIME SERIES MODEL FOR PREDICTING THE MEAN DEATH RATE OF A DISEASE

D. K. SHANGODOYIN[1], **J. F. OJO**[2], **J. O. OLAOMI**[1], **A. O. ADEBILE**[3]

ABSTRACT

This study develops a time series model to estimate the mean death rate of either an emerging disease or re-emerging disease with a bilinear induced model. The estimated death rate converges rapidly to the true parameter value for a given mean death at time t. The derived model could be used in predicting the m-step future death rate value of a given disease. We illustrated the new concept with real life data.

Keywords: Mean death, bilinear model, Death cases, emerging and re-emerging diseases.

1. Introduction

A measure of death rate gives an indicator to quality of health services available within a locality as well as the standard of preparedness of the health provider to curb the menace of the disease; the most important to health care provider is on how best to estimate the death rate of a given contagious disease. Globally, World Health Organization (WHO) is saddled with the responsibility of how best to estimate the death rate arising from a global pandemic. The World Health Organization (WHO) convectional formula for computation of death rate is simply the ratio of the number of known deaths to the total number of confirmed cases (Mathers and Loncar (2006), WHO (2005) Chronic disease report), however, this formula is likely to underestimate the true death rate because the outcomes of many cases might be unknown or uncertain at the time these figures are compiled. Some other methods for estimating death rate have varying degrees of challenges as mentioned in Shangodoyin (2010), in particular, the generalized mixed effect model of estimating death rate discussed by Tong

[1] Department of Statistics, University of Botswana, Botswana. Corresponding author: shangodoyink@mopipi.ub.bw.
[2] Department of Statistics, University of Ibadan, Ibadan, Nigeria.
[3] Department of Statistics, Federal Polytechnic, Ede, Nigeria.

and Chang (2006) although converges quickly to the death rate computed from the complete data, the linear model specified leads to a singular precision matrix for the unknown parameters, also the choice of the singular value decomposition presented may restrain this approach for practical use and in their paper, it was concluded that further research on how to carry out the estimation of the conditional mean death rate model with constraint in which the estimated death rate should be greater than or equal to zero.

Linear models are widely used because of their unrivalled simplicity, but they cannot be applied for data that have a turning-or rate-change-point, even if the data show good linearity sufficiently far from this point. To describe such bilinear-type data so as to make possible smooth and fully parameterizable transitions between two linear segments while still maintaining a clear connection with the linear models we require a bilinear type of model. Buchwald, P (2007) utilized a completely generalized version of a linearized biexponential model (LinBiExp) to various time profiles of biological and medical interest including growth profiles, such as those of human stature, agricultural crops and fruits, multicellular tumor spheroids, single fission yeast cells, or even labour productivity, and decline profiles, such as age-effects on cognition in patients who develop dementia and lactation yields in dairy cattle.

In this study, an attempt is made to develop a bilinear time series model to estimate the overall death rate of an emerging or re-emerging disease with the inclusion of bilinear parameters on the conditional mean death model.

2. Basic theory of bilnear models

There are situations when it is felt that linear time series models may not be adequate in explaining the underlying random mechanisms as in for instance sunspot data, Canadian lynx data which cannot be adequately described by linear time series models and the test proposed by Rao and Gabr (1981) does confirm that the above series cannot be described by linear Gaussian models. However, one may ask if there exist other models, which can provide a better fit. One is led then to consider non-linear models. Jones (1978), Granger and Andersen (1978), Haggan and Ozaki (1980), Priestly (1980), Tong and Lim (1980) and Rao(1981) have considered particular types of nonlinear time series models. The nonlinear models considered by Granger and Andersen (1978) and Rao (1981) are known as bilinear time series models. This class of time series has been found to provide a better fit as well useful in many areas for example biological sciences, ecology and engineering (see Mohler (1973), Bruni, Dupillo and Koch (1974)). From what we have above, one could say that the time series models for which it may be possible to obtain optimal forecasts for several steps ahead and which can perform better than a linear model are the bilinear time series models.

The most general form of the bilinear model, as defined in Granger and Andersen (1978a) which is denoted by BL (p, q, r, s) is given as:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + e_t - \sum_{j=1}^{q} \theta_j e_{t-j} + \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_{ij} X_{t-i} e_{t-j}$$

where the noise e_t consists of independent, identically distributed random variables with zero means and variance σ^2. We assume also that the model is invertible. Also e_t is assumed to be independent of X_s. The generality of the above model makes it too complicated to be analyzed successfully (Tuan and Lanh,1981); as a result these authors considered the first order bilinear model BL (1, 0, 1, 1) given as $X_t = \phi_1 X_{t-1} + e_t + b_{11} X_{t-1} e_{t-1}$ also, Rao (1980) considered the bilinear model BL (p, 0, r, s) given as

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_{ij} X_{t-i} e_{t-j} + e_t$$

2.1. The vector form

It is well known that the linear autoregressive moving average models can be written in the form of a first-order vector difference equation. Anderson, 1971; Priestley, 1978, 1980 and this vector form is known as the state space form. It is convenient to study the properties of the process when the model is in the state space form because of the Markovian nature of the model (Akaike, 1974).

Consider the model BL (p, d, 0, r, 1) given as

$$X_t = \psi_1 X_{t-1} + \ldots\ldots\ldots + \psi_{p+d} X_{t-p-d} + \left(\sum_{k=1}^{r} b_{k1} X_{t-k} \right) e_{t-1} + e_t \qquad (3.2.1)$$

Let us define the matrices

$$\Psi = \begin{bmatrix} -\psi_1 & -\psi_2 & \psi_3 & & -\psi_{p+d-1} & \psi_{p+d} \\ 1 & 0 & 0 & \ldots.. & 0 & 0 \\ 0 & 1 & 0 & \ldots.. & 0 & 0 \\ 0 & 0 & 0 & \ldots\ldots & 1 & 0 \end{bmatrix},$$

$$B = \begin{pmatrix} b_{11} & b_{21} & b_{31} \ldots\ldots & b_{r1} \\ 0 & 0 & 0 \ldots\ldots\ldots & 0 \\ 0 & 0 & 0 \ldots\ldots\ldots & 0 \end{pmatrix}$$

and vectors C^T=(1, 0, 0,…..0), H^T=(1, 0, 0,…..0), and let $\mathbf{X}_t^T = (X_t, X_{t-1}, ..., X_{t-p+d})$,
t = …..-1, 0, 1,…. With this notation, we can write the model in sub section 2.1 in the vector form as:

$$\mathbf{X}_t = \Psi \mathbf{X}_{t-1} + \mathbf{B}\mathbf{X}_{t-1}e_{t-1} + \mathbf{C}e_t \qquad \textbf{and} \qquad \mathbf{X}_t = \mathbf{H}^T \mathbf{X}_t$$

It is pertinent to ask ourselves that under what conditions on the given matrices C, ψ, B and on the given sequence $\{e_t$, t = …-1, 0, 1, …} of independent identically distributed random variables with common mean zero and common variance $\sigma^2 < \infty$, does there exist a vector-valued strictly stationary process {X_t, t= …-1, 0, 1, …} satisfying $\mathbf{X}_t = \Psi \mathbf{X}_{t-1} + \mathbf{B}\mathbf{X}_{t-1}e_{t-1} + \mathbf{C}e_t$ for every t = …-1, 0, 1,… The answer to the above is considered in sub section 2.2

2.2. Stationarity and Convergence of BL (p, d, 0, r, 1)

In this section, we give a sufficient condition for the existence of strictly stationary process and convergence conforming to the bilinear model defined in section 2.1, in the case when s =1. This we do through the following theorem.

Theorem: Let $\{e_t, t \in Z\}$ be a sequence of independent identically distributed random variables defined on a probability space (Ω, IR, P) such that E e_t = 0 and $Ee_t^2 = \sigma^2 < \infty$. Let Ψ and B be two matrices of order p x p such that $\rho(\Psi \otimes \Psi + \sigma^2 B \otimes B) = \lambda < 1$. Let C be any column vector with components c_1, c_2, …, c_p. Then the series of random vectors

$$\sum_{r\geq 1} \prod_{i=1}^{r} (\Psi + Be_{t-j}) Ce_{t-r}$$

converges absolutely almost surely as well as in the mean for every fixed t in Z. Further, if

$$X_t = Ce_t + \sum_{r\geq 1} \prod_{i=1}^{r} (\Psi + Be_{t-j}) Ce_{t-r}, \quad t \in Z,$$

then $\{X_t, t \in Z\}$ is a strictly stationary process conforming to the bilinear model $\mathbf{X}_t = \Psi \mathbf{X}_{t-1} + \mathbf{B}\mathbf{X}_{t-1}e_{t-1} + \mathbf{C}e_t$, for every t in Z. Conversely, if $\{X_t, t \in Z\}$ is a strictly stationary process satisfying $X_t = Ce_t + \Psi X_{t-1} + BX_{t-1}e_{t-1}$, for every t in Z for some sequence $\{e_t, t \in Z\}$ of independent identically distributed random variables with $Ee_t = 0$ and $Ee_t^2 = \sigma^2 < \infty$ and some matrices Ψ , B, C of orders

p x p, r x r, p x 1 respectively with $\rho\left(\Psi \otimes \Psi + \sigma^2 B \otimes B\right) = \lambda < 1$, then $X_t = Ce_t + \sum_{r\geq 1} \prod_{i=1}^{r} \left(\Psi + Be_{t-j}\right) Ce_{t-r}$, for every t in Z.

The proof of this theorem was considered in Shangodoyin, Ojo and Marcin (2010).

3. Mean Death rate model specification

Assume that p_j is the probability that a confirmed case ends up in death on the j-th day after the confirmation of the disease; also q_j is the corresponding probability of being recovered and discharged from the hospital on the j-th day. Then $P = \sum_{j=0}^{\infty} p_j$ is the death rate of the disease and $Q = 1 - P = \sum_{j=0}^{\infty} q_j$ is the recovery rate; all cases are assumed to be independent of each other and have identical death rate and probability distribution of time to death or discharge. Let C_t be *the* number of confirmed cases at time t, and D_t be the *number* of deaths at time t. The death on the t-th time must be from the previously known cases at time t-j, then the probability that it is from time t-j equals, p_j, say, and there are C_{t-j} at time t-j (Tong and Chan (2006)). The conditional mean of D_t given C_{t-j}; for j=0,1..... as

$$E\left(D_t / C_{t-j}\right) = \mu_t = \sum_{-\infty}^{\infty} p_j C_{t-j} \tag{1}$$

For a finite general linear process, the expression in (1) is written as

$$\mu_t = \sum_{a_1}^{a_2} p_j C_{t-j} \tag{2}$$

Where $a_1, a_2 \geq 0$ are lower and upper bounds of the time to death.

The expected value of death at time t is a linear function of the number of confirmed cases in time t-j, with assumption that all cases are independent of each other and have identical death and probability distribution of time to death may tempt analyst to adopt linear estimation techniques to determine p_j; this could pose some problems because if $\langle C_t \rangle = 0 \ \forall$ t<0, then theoretically D_t are serially dependent since they come from the earlier confirmed cases. Serial

correlation among the D_t may be accounted for by including a latent process on the right side of (2).

Also for finite uncensored aggregate data on re-emerging and/or emerging disease the analysis of the expected deaths taking into consideration the distribution of its past realizations and coupled with the large number of lags of C_{t-j} and $\mu_{\text{t-j}}$ can easily introduce multicollinearity in the model. This problem may not be adequately captured by linear model presented in equation (2), a problem identified by Tong and Chan (2006); and it could be corrected by incorporating smoothness assumption in the estimation of the unknown parameters using some parsimonious parametric class of models. A class of parsimonious models that have been found to provide a better fit some biological, ecological and medical data is the bilinear time series models (Mohler (1973) and Bruni etal (1974)). Martins (1999) and Gonclaves (2000) have shown that this class of models provide an optimal forecast for several steps ahead and can perform better than a linear model. In order to account for serial correlation in μ_t and to ensure that ε_j and p_j vary smoothly with lag j, we will adopt a modified bilinear model (MBL) (1, 0, 1, 1) to the left hand side of equation (2); for details see Rao (1981), Chen (1992), Shangodoyin, Ojo and Kozak (2010).
Assuming that u is obtained using the order selection criterion AIC with starting lower bound j=1, then the derived model is:

$$\mu_t - \alpha\mu_{t-1} - \beta\mu_{t-1}e_{t-1} = \sum_{u_1}^{u_2} p_j C_{t-j} + e_t \tag{3}$$

We then write the bilinear model for estimating the mean of the deaths at time t as:

$$\mu_t = \alpha\mu_{t-1} + \beta\mu_{t-1}e_{t-1} + \sum_{1}^{u} p_j C_{t-j} + e_t \tag{4}$$

4. Model estimation

Suppose that μ_t are generated by equation (4), the sequence of random deviates $\{e_t\}$ could be determined from the relation

$$e_t = \mu_t - \alpha\mu_{t-1} - \beta\mu_{t-1}e_{t-1} - \sum_{1}^{u} p_j C_{t-j} \tag{5}$$

To estimate the unknown parameters in equation (5), we make the following assumptions:

(i) The errors $\{e_t\}$ are independent and identically distributed with mean zero and variance σ^2 with finite kurtosis.

(ii) The root of the $p(B) = \sum p_j B^j$ polynomial lies outside the unit circle, this ensures stationary of $\{C_{t-j}\}$.

(iii) The values of $|\alpha|$<1 and $|\beta| < 1$, ensure that invertibility condition required of the bilinear process is satisfied.

Following Walker (1964), we shall refer to the parameters $\boldsymbol{\lambda}' = [\alpha, \beta, \mathbf{p}]$ as the structural parameters, the complete set of r=1+1+u+1 parameters will be denoted by $\boldsymbol{\Omega} = [\boldsymbol{\lambda}, \sigma^2]$. The task here is to get the estimator of $\boldsymbol{\lambda}$ using information on all realizations of the entire history of the system generating μ_t. The effect of transients can be minimized if the difference equation (5) is started off from a value of time t for which all the previous values of both C_t and μ_t are known. Let us specify the starting values of μ_t as $\dot{\mu}_t = \frac{\sum_1^t D_t}{t}$; ∀t=1,2.....

Thus, with the errors in equation (5) and appropriate initial values, say, $e_{t,0}$

The least squares estimator of $\boldsymbol{\lambda}' = [\alpha, \beta, \mathbf{p}]$ are obtained by minimizing the function $Q(\lambda) = \sum_1^n e_t^2$ with respect to initial values $\boldsymbol{\lambda}_0$. These estimators are derived by non-linear least squares estimation method discussed as follows:

Let $\boldsymbol{\lambda}_0' = [\alpha, \beta, p_1, p_2, ...p_u]$ for explicitness, we shall write $\lambda_{1,0} = \alpha$, $\lambda_{2,0} = \beta$, $\lambda_{3,0} = p_1$, $\lambda_{4,0} = p_2, ..., \lambda_{r-1,0} = p_{u-1}$ and $\lambda_{r,0} = p_u$ for the r-structural parameters. The first and second order partial derivatives of $Q(\boldsymbol{\lambda})$ with respect to the parameters $\boldsymbol{\lambda}_0'$ are respectively

$$G_i = \frac{\partial\left(Q(\boldsymbol{\lambda})\right)}{\partial\dot{\lambda}_i} = 2\sum_1^n e_t\left(\frac{\partial e_t}{\partial\dot{\lambda}_i}\right), \forall i = 1, 2, ..., r$$

and $$h_{ij} = \frac{\partial^2\left(Q(\boldsymbol{\lambda})\right)}{\partial\dot{\lambda}_i\partial\dot{\lambda}_j} = 2\sum_1^n\left(\frac{\partial e_t}{\partial\dot{\lambda}_i}\right)\left(\frac{\partial e_t}{\partial\dot{\lambda}_j}\right) + 2\sum_1^n e_t\frac{\partial^2 e_t}{\partial\dot{\lambda}_i\partial\dot{\lambda}_j}$$

The elements in G_i are derived from:

$$\frac{\partial e_t}{\partial \alpha} = \dot{\mu}_{t-1}; \frac{\partial e_t}{\partial \beta} = \dot{\mu}_{t-1} e_{t-1}; \frac{\partial e_t}{\partial \varepsilon_j} = C_{t-j}, \forall j = 1, 2, ..., u$$

To estimate the structural parameters, let $\dfrac{\partial^2 e_t}{\partial \dot{\lambda}_i \partial \dot{\lambda}_j} = 0$

Such that for any given values of C_{t-j} and $\dot{\mu}_t$, we evaluate the recursive equations $\sum_1^n e_t \left(\dfrac{\partial e_t}{\partial \dot{\lambda}_i} \right) = 0$ and $\sum_1^n \left(\dfrac{\partial e_t}{\partial \dot{\lambda}_i} \right) \left(\dfrac{\partial e_t}{\partial \dot{\lambda}_j} \right) = 0$

let $\mathbf{V}'(\boldsymbol{\lambda}) = \left[\dfrac{\partial Q(\boldsymbol{\lambda})}{\partial \dot{\lambda}_1}, \dfrac{\partial Q(\boldsymbol{\lambda})}{\partial \dot{\lambda}_2}, ..., \dfrac{\partial Q(\boldsymbol{\lambda})}{\partial \dot{\lambda}_r} \right]$ an $\mathbf{H}(\boldsymbol{\lambda}) = \left[\dfrac{\partial Q(\boldsymbol{\lambda})}{\partial \dot{\lambda}_i \partial \dot{\lambda}_j} \right]$

be the matrices of the first and second partial derivatives. Expanding $V(\boldsymbol{\lambda})$ near $\boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}$ in Taylor series, we obtain:

$$\left[\mathbf{V}(\hat{\boldsymbol{\lambda}}) \right]_{\lambda = \hat{\lambda}} = 0 = \mathbf{V}(\boldsymbol{\lambda}) + \mathbf{H}(\boldsymbol{\lambda})(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda})$$

This gives the estimators of the structural parameters as $\hat{\boldsymbol{\lambda}} = \dot{\boldsymbol{\lambda}} - \mathbf{H}^{-1}(\dot{\boldsymbol{\lambda}}) \mathbf{V}(\dot{\boldsymbol{\lambda}})$

Now, suppose that we restrict to u=1, in this case the lag to death is a unit of the measuring period, say, one (year or month of week or day). The simultaneous equations resulting from substituting all the necessary elements of G_i and $\mathrm{h_{ij}}$ are :

$$\sum_1^n e_t \left(\frac{\partial e_t}{\partial \dot{\alpha}} \right) = \sum_1^n \dot{\mu}_t \dot{\mu}_{t-1} - \dot{\alpha} \sum_1^n \dot{\mu}_{t-1}^2 - \dot{\beta} \sum_1^n \dot{\mu}_{t-1}^2 e_{t-1} - \dot{p}_1 \sum_1^n \dot{\mu}_{t-1} C_{t-1} = 0 \tag{6}$$

$$\sum_1^n e_t \left(\frac{\partial e_t}{\partial \dot{\beta}} \right) = \sum_1^n \dot{\mu}_t \dot{\mu}_{t-1} e_{t-1} - \dot{\alpha} \sum_1^n \dot{\mu}_{t-1}^2 e_{t-1} - \dot{\beta} \sum_1^n \dot{\mu}_{t-1}^2 e^2{}_{t-1} - \dot{\varepsilon}_1 \sum_1^n \dot{\mu}_{t-1} C_{t-1} e_{t-1} = 0 \tag{7}$$

$$\sum_1^n e_t \left(\frac{\partial e_t}{\partial \dot{p}_1} \right) = \sum_1^n \dot{\mu}_t C_{t-1} - \dot{\alpha} \sum_1^n \dot{\mu}_{t-1} C_{t-1} - \dot{\beta} \sum_1^n \dot{\mu}_{t-1} C_{t-1} e_{t-1} - \dot{p}_1 \sum_1^n C_{t-1}^2 = 0 \tag{8}$$

The equations derived from the second partial derivatives are:

$$\sum_1^n \left(\frac{\partial e_t}{\partial \dot{\alpha}} \right) \left(\frac{\partial e_t}{\partial \dot{\beta}} \right) = \sum_1^n \dot{\mu}_{t-1}^2 e_{t-1} = 0 \tag{9}$$

$$\sum_1^n \left(\frac{\partial e_t}{\partial \dot{\alpha}} \right) \left(\frac{\partial e_t}{\partial \dot{p}_1} \right) = \sum_1^n \dot{\mu}_{t-1} C_{t-1} = 0 \tag{10}$$

$$\sum_1^n \left(\frac{\partial e_t}{\partial \dot{\beta}} \right) \left(\frac{\partial e_t}{\partial \dot{p}_1} \right) = \sum_1^n \dot{\mu}_{t-1} C_{t-1} e_{t-1} = 0 \tag{11}$$

Substituting equations (9) and (10) in (6) gives:

$$\dot{\alpha} = \frac{\sum_{1}^{n} \dot{\mu}_t \dot{\mu}_{t-1}}{\sum_{1}^{n} \dot{\mu}_{t-1}^2} \tag{12}$$

Also, using equations (9) and (11) in (7) gives:

$$\dot{\beta} = \frac{\sum_{1}^{n} \dot{\mu}_t \dot{\mu}_{t-1} e_{t-1}}{\sum_{1}^{n} (\dot{\mu}_{t-1} e_{t-1})^2} \tag{13}$$

The estimator for $\dot{p}_1$ is obtained using equations (10) and (11) in (8) gives:

$$\dot{p}_1 = \frac{\sum_{1}^{n} \dot{\mu}_t C_{t-1}}{\sum_{1}^{n} C_{t-1}^2} \tag{14}$$

To determine the mean and variance of the estimators defined in equations (12), (13) and (14) we start by re-writing equation (12) as

$$\hat{\alpha} = \frac{\sum_{1}^{n} \hat{\mu}_t \hat{\mu}_{t-1}}{\sum_{1}^{n} \hat{\mu}_{t-1}^2} = f(\hat{\mu}_o, ..., \hat{\mu}_n)$$

Assuming Taylor expansion is valid, we can approximate

$$\hat{\alpha} = f(\hat{\mu}_o, ..., \hat{\mu}_n) \cong f(\mu_o, ..., \mu_n) + \sum_{i=0}^{n} (\hat{\mu}_i - \mu_i) \frac{\partial f}{\partial \hat{\mu}_i} |_{\hat{\mu}_i = \mu_i} + ... \tag{15}$$

The equation (15) yields

$$E(\hat{\alpha}) \cong f(\mu_o, ..., \mu_n) \tag{16}$$

and

$$V(\hat{\alpha}) \cong \sum_{i=0}^{n} \{(\frac{\partial f}{\partial \hat{\mu}_i})^2 |_{\hat{\mu}_i = \mu_i}\} V(\hat{\mu}_i) + \sum_{i \neq}^{n} \sum_{j=1}^{n} \{(\frac{\partial f}{\partial \hat{\mu}_i} \frac{\partial f}{\partial \hat{\mu}_j}) |_{\hat{\mu}_i = \mu_i, \hat{\mu}_j = \mu_j}\} Cov(\hat{\mu}_i, \hat{\mu}_j) \tag{17}$$

The expression of variance (17) is extremely complex and hence no analytical form of $V(\hat{\alpha})$ is available. However, one can compute the $V(\hat{\alpha})$ numerically by using any of the variance estimation methods such as Jackknife method, Bootstrap method and Random group method.

5. Data Analysis and Discussion

This new methodology shall be considered on two real life data; data was collected on Tuberculosis incidence recorded monthly at Julia Molefe Clinic, Gaborone, Botswana and Malaria incidence data was recorded on annual basis at the University College hospital, Ibadan, Nigeria. We utilized equation (14) to compute the death rate; this formula relies both on the moving death average and the number of cases observed for each respective month or year over time.

In table 1, we presented the computed statistics for the death rate of Tuberculosis data using equation (14) and have 0.07358 about 7.3% with estimator variance 0.0016; using the WHO formula we obtained 0.076 about 7.6% the variance of estimator is 0.0684.

In table 2, we presented some computed statistics based on estimator derived in equation (14) and WHO convectional estimator for death rate using Malaria data; the estimator in equation (14) is 0.004 about 0.4% with estimator variance 0.0000005, whereas the WHO estimator for this data is 0.006 about 0.6% with estimator variance 0.000385. Since the ratio of estimators $\frac{Var(estimator \text{ given in equation 14})}{\text{Var(WHO estimator)}} < 1$ in both real life data considered, we affirmed that the estimator presented in equation (14) is more efficient than the WHO estimator; also, WHO formula is not likely to give the true estimate of death rate because the outcomes of some cases might not be known at the time of compilation besides equation (14) may track the monthly (or periodic) number of deaths relatively well.

To predict the mean death rate for the Tuberculosis data we used equations (12), (13) and (14) to compute the estimates of $\dot{\alpha}$, $\dot{\beta}$ and $\dot{p}_1$ as 0.857, -1.21 and 0.081 respectively, thus the prediction model for this data is $\mu_t = \underset{(0.002)}{0.857}\,\mu_{t-1} - \underset{(0.078)}{1.21}\,\mu_{t-1}e_{t-1} + \underset{(0.0006)}{0.081}C_{t-1}$ assuming a one-step time to death for this disease. Similarly for the Malaria data, the computed estimates of

$\dot{\alpha}$, $\dot{\beta}$ and $\dot{p}_1$ are 1.10, -0.1161 and 0.004 respectively; also the prediction model for this data is $\mu_t = \underset{(0.056)}{1.10}\,\mu_{t-1} - \underset{(0.145)}{0.12}\,\mu_{t-1}e_{t-1} + \underset{(0.000001)}{0.004}\,C_{t-1}$ assuming a one-step time to death for this disease.

6. Conclusion

The assumption that the time to death is at most one (day or month or year depending on the nature of the disease); then the overall death rate can be computed using $\dot{p}_1$ provided $C_{t-j} > 0$, then the quantity in (14) will be less than unity, this is another view to the further research advocated by Chan and Tong (2006) approach. Also we can use equation (4) to predict the death rate of a given disease over a given period. It is interesting to note that the estimator in equation (14) performs better than the WHO conventional estimator in the two real-life cases considered.

Appendix

Table 1. TB DEATH STATISTICS (January, 2010- July, 2010)

TIME (t) in Month	C_t	D_t	μ_t	C_{t-1}	e_t
1	16	1	1	14	-0.134
2	9	1	1	16	-1.173
3	11	0	1	9	-0.606
4	12	2	1	11	-0.768
5	5	2	1	12	-0.849
6	13	0	1	5	-0.282
7	12	0	1	13	-0.930
Total	78	6			

Appendix

Table 2. ANNUAL MALARIA DEATH STATISTICS

TIME (t) in Month	C_t	D_t	μ_t	C_{t-1}	e_t
2000	3664	10	10	1670	-5.15
2001	2483	5	7.5	3664	-18.17
2002	2507	13	9.3	2483	-8.88
2003	2113	9	9.3	2507	-9.38
2004	2004	11	9.6	2113	-9.08
2005	2712	37	14.2	2004	-4.38
2006	3072	25	15.7	2712	-10.77
Total	18555	110			

REFERENCES

AKAIKE, H. (1973). Maximum Likelihood Identification of Gaussian Auto-regresssive Moving Average Models. Biometrika 60, 255-265.

ALTMAN, K.L. (2003). The SARS epidemic. The front line research. New York Times. May 07 Edition.

ANDERSON, T.W. (1971). The Statistical Analysis of Time Series, New York and London Wiley.

BUCHWALD, P. (2007). A general bilinear model to describe growth or decline time profiles Math Biosci. 2007 205(1):108-36.

BRUNI, C., DUPILLO, G. and KOCH, G. (1974). Bilinear Systems: An Appealing Class of Nearly Linear System in Theory and Application. IEEE Trans. Auto Control Ac-19, 334-338.

CHEN, C.W.S. (1992). Bayesian Inferences and forecasting in bilinear time series models. Communications in Statistics – Theory and Methods 21(6), 1725-1743.

DONNELLY, C.A., et. al (2003). Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. The Lancet, 361,pg 1761-1766.

GRANGER, C.W.J. and ANDERSON, A.P. (1978). Introduction to Bilinear Time Series Models. Vandenhoeck and Puprecht.

GONCLAVES, E., JACOB P. and MENDES-LOPES N. (2000). A decision procedure for bilinear time series based on the Asymptotic Separation. Statistics, 333-348.

HAGGAN, V. and OZAKI, T. (1980). Amplitude Dependent Exponential AR Model Fitting for Non-Linear Random Vibrations. Proc. International Time Series Meeting, Nottinghamm, ed. O. D. Anderson. North Holland.

JONES, D.A. (1978). Non-linear Autoregressive Processes. Proc. Royal Soc. London (A) 360, 71-95.

KINDIG, D.A., SEPLAKI, C.L. and LIBBY, D.L. (2002). Death variation in US subpopulations. Bulletin of WHO, 80(1), pg 9-15.

LI, W.K. and HUI, Y.V. (1983). Estimation of random coefficient autoregressive process: An empirical Bayes approach . Journal of Time Series Analysis, Vol. 4, No. 2. pg 89-94.

LOHMANN, M.S. (2005). Data quality control and observations error estimation. www.ucar.edu.

MARTINS, C.M. (1997). A Note on the Third-Order Moment Structure of a bilinear model with Non-Independent Shocks. Potrugaliae Mathematica vol 56, 58-89.

MATHERS, C.D. and LONCAR, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. PLos Medicine. Vol. 3. Issue 11, pp 2011-2030.

MOHLER, R.R. (1973). Bilinear Control Processes. New York: Academic Press.

PRIESTELY, M.B. (1978). Non-Linear Models in Time Series Analysis. The Statistician 27, 159-176.

SHANGODOYIN, D.K., OJO, J.F. and KOZAK, M. (2010). Subsetting and identification of optimal models in one-dimensional bilinear time series modelling. International Journal of management science and Engineering management England, UK, 5(4), pp 252-260.

SHANGODOYIN, D.K. (2010). Using time series models to determine the death rate of a given disease. To appear in International Encyclopaedia of Statistical Science (Springer, USA).

SUBBA RAO, T. (1981). On the theory of Bilinear time Series Models ,Jour. R. Statist. Soc.B. 43, 244-255.

TONG, H. AND CHAN, K. (2006). Estimating the death rate of an emerging disease by Time Series Analysis. Technical report, Department of Statistics & Actuarial Science, University of Iowa, Iowa, USA.

TUAN DINH PHAM and LANH TAT TRAN (1981). On the First Order Bilinear Time Series Model. Jour. Appli. Prob. 18, 617-627.

WALKER, A.M. (1962). Large sample estimation of parameters of autoregressive processes with moving average residuals. Biometrika, 49, 117-131.