A Modified Statistical Estimation Algorithm for Partially Observed Diffusion Epidemic Models

A. H. Aliu[1]; A. A. Abiodun[2]; R. A. Ipinyomi[3]

[1]Department Mathematics and Statistics,
Rufus Giwa Polytechnic, Owo,
Ondo State Nigeria.
e-mail: ahaliu@ymail.com

[2,3]Department of Statistics,
University of Ilorin,
Ilorin, Nigeria
e-mail: alfredabiodun1@gmail.com[2]; ipinyomira@yahoo.co.uk[3]

***Abstract*—Diffusion processes governed by Stochastic Diffusion Equations (SDEs) are a well known tool for modeling continuous-time data. Consequently, there is widely interest in efficiently estimating diffusion parameters from discretely observed data. Likelihood based inference can be problematic, as the transition densities are rarely available in closed form. One widely used solution proposed by Pedersen, (1995) involved the introduction of latent data points between every pair of observations to allow an Euler-Maruyama approximation of the true transition densities to become accurate. We applied Markov Chain Monte Carlo methods to sample the conditional posterior distribution of the latent data and model parameters on discretely observed data. In this case, we modified algorithm that would explore efficient MCMC schemes that are affected with degeneracy problem. In our approached the situation where the scheme becomes degenerate does not occur. This method capable of sampling efficient estimate of diffusion parameters from discrete observed epidemic data either with or without measurement error.**

Keywords: *Diffusion process, stochastic differential equation, Bayesian inference, Numerical solution, partially observed data, Diffusion Bridge, MCMC, SEIR epidemic model.*

I. Introduction

Most epidemic data are discretely observed and undergo stochastic transition rate. Stochastic epidemic models allow more realistic description of the transmission of disease as compared to deterministic epidemic models [3], [2]. However, parameter estimation is challenging for discretely observed data for stochastic models [19], [13]. Several methods of frequentist procedures to infer on the parameters are been considered in the literatures. Most techniques struggle when inter-observation times are large.

Here, we employ an efficient Bayesian estimation approach under stochastic differential equation (SDE) technique. Stochastic differential equation (SDE) models play a prominent role in a range of application areas, including biology, chemistry, epidemiology, mechanics, microelectronics, economics, and finance [5], [15], [8], [4], [10], [11] & [6]. A complete understanding of SDE theory requires familiarity with advanced probability and stochastic processes. These processes are often referred to as a diffusion process.

Diffusion processes are a promising instrument to realistically model the time-continuous evolution of natural phenomena. Diffusion process have an advantage over some of the other stochastic formulations, in that, they can be easily derived directly from the deterministic system of ordinary differential equations and have a relatively simple form [16].Inferring the parameters of such models possess challenging in the field of study.

In this paper, we reviewed some of the empirical solution to parameter estimation problems. We adopted Bayesian imputation approach to estimate the parameter of interest. We replaced intractability transition density problems with Euler-Maruyama approximation. We also adopted data augmentation scheme so as to limit the discretization error incurred by the approximation.

II. Research Methodology

We restrict attention to estimation within the Bayesian imputation approach. The essential idea of the Bayesian imputation approach is to augment low frequency data by introducing intermediate time-points between observation times. An Euler-Maruyama scheme is then applied by approximating the transition densities over the induced

discretization.To deal with such data, we define Observation say D as:

$$D_T = \{X_{t0}^{(i)}, X_{t1}^{(i)}, \cdots, X_{tn}^{(i)}\}' \quad (1)$$

$D_n^{(1)}$ as discretely observed and $D_n^{(2)}$ as unobserved part. where, $X^{(1)}_t$ represent dimension $d_1 > 0$ and $X^{(2)}_t$ dimension $d_2 \geq 0$. With $d_1+d_2 = d$. If $d_2 = 0$, implies fully observed.

We consider a parameterized family of d-dimensional diffusion process $\{X_t,\ t \geq 0\}$ satisfied by a Stochastic Differential Equation of the form:

$$dX_t = \alpha(X_t, \theta)dt + \sqrt{\beta(X_t, \theta)}dW_t \quad (2)$$

$$, \mathbf{X0 = x_0}$$

$\mathbf{X_t}$ is the value of the process at time t, $\boldsymbol{\theta}$ is the parameter vector of length p, $\boldsymbol{\alpha(X_t, \theta)}$ is the drift functions, $\boldsymbol{\beta(X_t, \theta)}$ is the diffusion coefficient, and $\mathbf{W_t}$is standard Brownian motion (*d-vector* Wiener Process). The $\mathbf{X_0 = x_0}$is the vector of initial conditions for this process. We seek a numerical solution via the Euler-Maruyama approximation. The idea is to discretize Equation (2) by Euler Scheme as [1].

$$\Delta X_t \equiv X_{t+\Delta t} - X_t = \alpha(X_t, \theta)\Delta t + \beta(X_t, \theta)^{\frac{1}{2}}\Delta W_t \quad (3)$$

Where$\Delta W_t \sim N_d(0, I\Delta t)$. Since, most diffusion process undergo Markov chain, we assume equidistant observation times with the likelihood function of the observation given parameters is of the form:

$$L(\theta \mid D_T) = \prod_{k=0}^{n-1} \pi(x_{k+1} \mid x_k, \theta) \quad (4)$$

Where, $\pi(x_{tk+1}|x_{tk}, \theta)$ denotes the transition density from $X_{tk} = x_{tk}$ to $X_{tk+1} = x_{tk+1}$.

This likelihood function is very rarely available in closed form. The maximum likelihood estimation would be intractable. We therefore considered Bayesian method of estimation.

BAYESIAN INFERENCE

Our modification based on Bayesian inference approach. In statistics, Bayesian inference is a method of inference in which Baye's rule is used to update the probability estimate for a hypothesis as additional evidence is required. The idea behind Bayesian inference is that the likelihood and prior are combined using Bayes' theorem to compute the posterior distribution.

The posterior density from (4) is given thus:

$$\pi(\theta \mid D_T) \propto \pi(\theta) \times \prod_{k=0}^{n-1} \pi^{Euler}(x_{k+1} \mid x_k, \theta) \quad (5)$$

Where $\pi(\theta)$ is the prior density, the Euler- Maruyama approximation might not be accurate if interval $[t_{k+1}, t_k]$ is too large. We therefore adopted a *data augmentation* approach.

In data augmentation we inserting m-1 additional time points in between $[t_{k+1}, t_k]$.

$$t_k = \tau_{km} < \tau_{km+1} < \cdots < \tau_{(k+1)m} = t_{k+1} \quad , k = 0, \ldots, K \quad (6)$$

Where, $\Delta\tau = \tau_{km+1} - \tau_{km} = \frac{t_{k+1} - t_k}{m}$

Therefore, the joint posterior for parameters and imputed data as

$$\pi(\theta, x \mid D_T^{(i)}) = \pi(\theta) \times \prod_{k=0}^{nm-1} \pi^{Euler}(x_{\tau k+1}^{(i)} \mid x_{\tau k}^{(i)}, \theta) \quad (7)$$

where Euler density $\pi^{Euler}(x_{\tau k+1}^{(i)} \mid x_{\tau k}^{(i)}, \theta) = N_d(x_{\tau k+1}^{(i)}; x_{\tau k}^{(i)} + \alpha(x_{\tau k}^{(i)}, \theta)\Delta\tau, \beta(x_{\tau k}^{(i)}, \theta)\Delta\tau)$

$N_d(.;\ \mu,\ \Sigma)$ denotes the multivariate Gaussian density with mean μ and variance-covariance Σ

Sampling Procedure

The posterior distribution is typically analytically intractable, we therefore sample via Markov Chain Mote Carlo (MCMC) scheme.

(i) for path update, we sample $x \mid x_0, x_T, \vartheta$

(ii) For parameter update, we sample $\vartheta \mid x_0, x_T, x$

In path updating, various diffusion bridges proposal mechanism for sample the skeleton path had been proposed in the literature, such as Diffusion bridge by [18], Modified diffusion bridge by [9], Regularized sampler by [14] among others,

Here we adopted Modified Diffusion Bridge proposed by [9].

Assuming the starting point ($x_0 = x_{\tau k}$) and the end point ($x_T = x_{\tau m}$) are observed, the path update proposal would now be our aim, to get this we defined a proposal distribution: $q(x_{\tau k+1} \mid x_{\tau k}, x_{\tau m}, \theta)$and find out the $\mu_{\tau k}$, $\Sigma_{\tau k}$.

Modified Diffusion Bridge for univariate model is of the form:(8)

$$X_{\tau k+1}^{(i)} \mid x_{\tau k}^{(i)}, x_{\tau m}^{(i)} \sim N_d(\mu_{\tau k}, \Sigma_{\tau k}) \quad ; k = 0, \ldots, m-1$$

where

$$\mu_{\tau k} = x_{\tau k}^{(i)} + \frac{x_{\tau m}^{(i)} - x_{\tau k}^{(i)}}{\tau_m - \tau_k}\Delta\tau \qquad \Sigma_{\tau k} = \frac{\tau_m - \tau_{k+1}}{\tau_m - \tau_k}\beta(x_{\tau k}^{(i)}, \theta)\Delta\tau$$

The marginal posterior density for the imputed data $\pi(x \mid x^{(i)}_{\tau k-1}, x^{(i)}_{\tau m}, \theta)$ has acceptance Probability of the form:

$$\alpha(X_{\tau k}^{*(i)}, X_{\tau k}^{(i)}) = \min\left\{1, \frac{\pi^{Euler}(x_{\tau k+1}^{*(i)} \mid x_{\tau k}^{(i)}, x_{\tau m}^{(i)}, \theta) \times q(x_{\tau k+1}^{(i)} \mid x_{\tau k}^{*(i)}, x_{\tau m}^{(i)}, \theta)}{\pi^{Euler}(x_{\tau k+1}^{(i)} \mid x_{\tau k}^{(i)}, x_{\tau m}^{(i)}, \theta) \times q(x_{\tau k+1}^{*(i)} \mid x_{\tau k}^{(i)}, x_{\tau m}^{(i)}, \theta)}\right\}$$

(9)

Under this update scheme, the proposal mechanism of the MCMC becomes degenerate as $m \to \infty$, meaning that, there is dependence between the parameters and the imputed values, likewise there is dependence between values of the imputed latent process itself. This was first highlighted as a

problem by [18]. To overcome this, we consider innovation scheme earlier proposed by [12], though not applicable to discrete observation.

CONTRIBUTION

Our contribution is on Modified Innovation scheme, that is, the MCMC sampling strategy to be considered was the innovation scheme, first introduced by [7]. In diffusion there is one-to-one relationship between ΔX_t and ΔW_t.

$$\Delta X_t^{(i)} = \alpha(X_t^{(i)},\theta)\Delta t + \sqrt{\beta(X_t^{(i)},\theta)}\Delta W_t^{(i)} \quad (10)$$

$$\Delta W_t^{(i)} = \beta(X_t^{(i)},\theta)^{-\frac{1}{2}}\{\Delta X_t^{(i)} - \alpha(X_t^{(i)},\theta)\Delta t\}$$

which implies:

$$= \beta(X_t^{(i)},\theta)^{-\frac{1}{2}}\left\{\alpha(X_t^{(i)},\theta) - \frac{x_{tj+1}^{(i)} - X_t^{(i)}}{t_{j+1}-t}\right\}dt + dW_t^{(i)} \quad (11)$$

Rather than sample from the distribution of conditional on the missing imputed data, the innovation scheme uses a subtle reparameterisation, by sampling conditional on the driving Brownian motion, and the latent path ($x_{\tau k}$) is obtained deterministically and consistent with the parameters of the model, therefore, this overcoming the dependence problem.

Here, we sampled the parameters of interest (θ), given the Brownian driving ($w_{\tau k}$) and observation (D_T) thus:

$$\pi(\theta \mid w, D_T) \propto \pi(\theta) \times \pi\{f(w,\theta) \mid D_T\} J(\theta) \quad (12)$$

where J(θ) is the Jacobian for one increment is

$$J(\theta) = \left|\frac{\partial \Delta W_t}{\partial X_t}\right| = \left|\beta^*(X_t,\theta)\right|^{-1/2}$$

The target distribution therefore becomes

$$\pi(\theta \mid w, D_T^{(i)}) \propto \pi(\theta) \prod_{k=0}^{nm-1} \pi^{Euler}(X_{\tau k+1} \mid X_{\tau k},\theta) \prod_{k=0}^{n-1} \prod_{j=0}^{m-2} \left|\beta^*(X_{\tau k+j},\theta)\right|^{-1/2} \quad (13)$$

Having set this update scheme, the new acceptance probability now becomes

$$\alpha(\theta^*,\theta) = \min\left\{1, \frac{\pi(\theta^*)\pi\{f(w,\theta^*) \mid D_T^{(i)}\} J(\theta^*)}{\pi(\theta)\pi\{f(w,\theta) \mid D_T^{(i)}\} J(\theta)}\right\} \quad (14)$$

III. ANALYSIS

We demonstrate the performance of aforementioned methods described above by applying it to synthesis simulated epidemic system of diffusion model. We considered stochastic infection model (SEIR Model) which undergo diffusion system of model:

$$dXt = \begin{pmatrix} \frac{-\beta X1X3}{N} \\ \frac{\beta X1X3}{N} - \gamma X2 \\ \gamma X2 - \alpha X3 \end{pmatrix} dt + \begin{pmatrix} \frac{\beta X1X3}{N} & \frac{-\beta X1X3}{N} & 0 \\ \frac{-\beta X1X3}{N} & \frac{\beta X1X3}{N} + \gamma X2 & -\gamma X2 \\ 0 & -\gamma X2 & \gamma X2 + \alpha X3 \end{pmatrix}^{1/2} dWt \quad (15)$$

Here, the state variable $X^{(i)}_t = (x_1, x_2, x_3)^T$, where, x_1 denotes Susceptible individuals, x_2 represent Exposed, and x_3 Infectious individuals with their initial condition for the state variables are (500000, 1000,10) respectively. The parameter of interest denoted by $\theta = (\beta, \gamma, \alpha)^T$. We initialized the sampler with $0 < \beta < 1$, $0 < \gamma < 0.7$ and $0.1 < k < 1$ that represent transmission rate, exposed rate and infection rate respectively. We performed iteration for 10^4times with three different number of imputed time points (m = 5, 15 and 50). In parameter proposal, we used independent sampler of the form$N_d(0, \psi_j^2)$ distribution for the proposal of parameter of interest, where ψ_j^2 is the turned variance of {0.009, 0.009, 0.001}for the parameter respectively.

To show that the proposed method does not degenerate when increasing the number of imputed time points, we applied modified innovation scheme. We set the starting time point at $t_0 = 0$ and end-time at T = 30, with equidistant time interval $\Delta\tau = 0.001$.

We choose an uninformative prior for each of the parameter, and apply the MCMC scheme to infer the posterior values of the model.

We compared the empirical method (Naïve) with our new method.The path and parameters update and the results were depicted below.

Implementation was done with the aid of R-software programming.

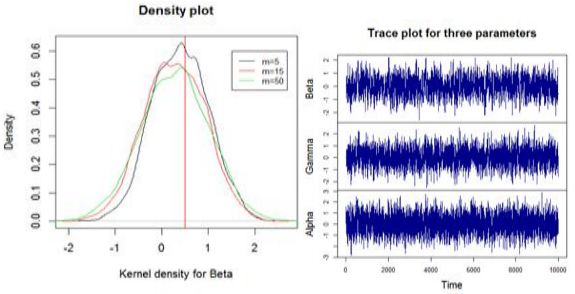

Figure 1(a) Figure 1(b)

Figure-1(a) shows the density plot for the innovation scheme for three different imputed values, the three imputed were very closed. And 1(b) shows the trace plot for the three parameters, the trace plot mixing very well.

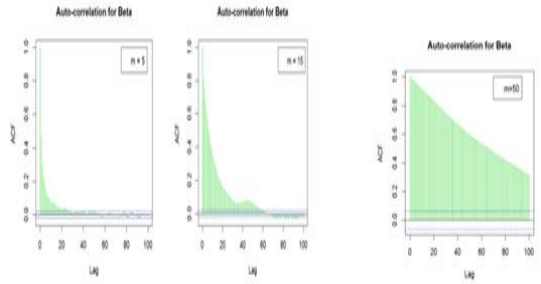

Figure 2(a). Auto-correlation for the Naive method scheme

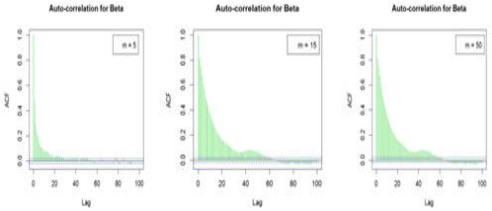

Figure 2(b). Auto-correlation for the Modified innovation scheme. *Source: Simulated SEIR synthetic data.*

Figure-2(a) & (b) shows auto-correlation forboth traditional naive method for parameter beta and modified innovation scheme.

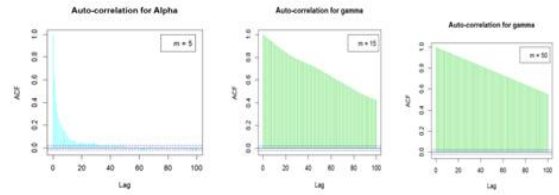

Figure 3(a) Auto-correlation for the Naive method

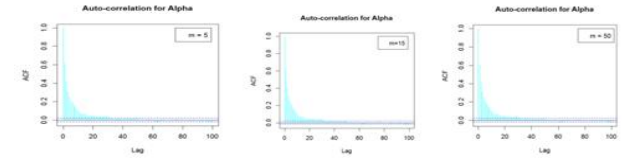

Figure 3(b). Auto-correlation for the Modified innovation scheme. *Source: Simulated SEIR synthetic data.*

Figure-3(a) & (b) shows the naive method for the parameter alpha and modified Innovation scheme.

IV. RESULTS AND DISCUSSIONS

The modified method adopted capable of sampling efficient estimate of diffusion parameters from discrete observed epidemic data for infinite number of imputed time points. See figures 1(a), 2(b) and 3(b).

The results obtained from posterior distribution in modified innovation scheme when the number of imputed points increases does not worsen the mixing of the chain, figure 2(b) and 3(b). Also, under the modified innovative scheme as number of imputed tend to infinite ($m \to \infty$), we have both parameters and path update that are consistent.

V. CONCLUSION

We consider a diffusion process approach based on a stochastic discrete-time approximation diffusion process. With the aims of estimate unobserved latent data and parameters of given epidemic system of model when the number of imputed time point is very large. We presented a naive class of estimation with our new method the modified innovation scheme (13) which are computationally and statistically efficient, and can be readily applied in situations where the discrete-observation of the process is possible.

In our approached the situation where the scheme becomes degenerate does not occur.

ACKNOWLEDGMENT

We are grateful to the members of Nigeria statistical society for helpful discussion during the oral presentation of this manuscript during the first international conference and the editor for their constructive reviews of the manuscript.

REFERENCES

[1] Allen, E. J. (2007). Modelling with Stochastic Differential Equations. Published by Springer, Dordrecht, The Netherlands.

[2] Andersson, H. and T. Britton, 2000. Stochastic epidemic models and their statistical analysis. Lecture Notes in Statistics, Vol. 151, Springer, New York.

[3] Becker, N., 1989. Analysis of infectious disease data. London: Monographs on statistics and applied probability, Chapman & Hall.

[4] Bibby, B. and M. Sørensen, 2001. Simplified estimating functions for diffusion models with a high-dimensional parameter. Scandinavian Journal of Statistics, 28(1): 99-112.

[5] Black, F. and M. Scholes, 1973. The pricing of options and corporate liabilities. Journal of Political Economy, 81(3): 637-654, The University of Chicago press.

[6] Chiarella, C., H. Hung and T.D. Tô, 2009. The volatility structure of the fixed income market under the HJM framework: A nonlinear filtering approach. Computational Statistics and Data Analysis, Elsevier 53(6): 2075-2088.

[7] Chib, S., M.K. Pitt and N. Shephard, 2006. Likelihood based inference for diffusion driven models. Economics Papers No. 2004-W20, Economics Group, Nuffield College, University of Oxford.

[8] Cox, J., J. Ingersoll and S. Ross, 1985a. An intertemporal general equilibrium model of asset prices. Econometrica, 53(2): 363-384.

[9] Durham, G.B. and A.R. Gallant, 2001. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. Journal of Business and Economic Statistics, 20(3): 297-338.

[10] Elerian, O., S. Chib and N. Shephard, 2001. Likelihood inference for discretely observed nonlinear diffusions. Econometrica, 69(4): 959–993.

[11] Eraker, B., 2001. MCMC analysis of diffusion models with application to finance. Journal of Business & Economic Statistics, 19(2): 177–191.

[12] Golightly, A. and D.J. Wilkinson, 2008. Bayesian inference for nonlinear multivariate diffusion models observed with error. Computational Statistics and Data Analysis, 52(3): 1674-1693.

[13] Jimenez, J., R. Biscay and T. Ozaki, 2005. Inference methods for discretely observed continuous-time stochastic volatility models: A commented overview. Asia-Pacific Financial Markets, 12(2): 109-141.

[14] Lindstrom, E., 2012. A regularised bridge sampler for sparsely sampled diffusions. Stat.Comput. 22(1): 615-623.

[15] Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics, 3(1976): 125-144.

[16] Øksendal, B. K, 2003. Stochastic differential equations: An introduction with applications. 6th Edn.: Springer. New York.

[17] Pedersen, A. R. (1995). Consistency and asymptotic normality of an approximate maximum likelihood estimator forDiscretely observed diffusion processes. Bernoulli 1(3): 257-279.

[18] Roberts, G.O. and O. Stramer, 2001. On inference for partially observed nonlinear diffusion models using the metropolis-hastings algorithm. Biometrika, 88(3): 603-621.

[19] Sørensen, H., 2004. Parametric inference for diffusion processes observed at discrete points in time: A survey. International Statistical Review, 72(3): 337–354.