Journal of Applied Analysis and Computation
Volume 16, Number 6, December 2026, 3071–3097
Website:http://www.jaac-online.com
DOI:10.11948/20240033

DYNAMIC BEHAVIOR AND PARAMETER ESTIMATION OF A STOCHASTIC SIS DISEASE TRANSMISSION MODEL WITH LOGISTIC POPULATION INPUT AND DEGENERATE DIFFUSION TERM*

Xuexin Li[1], Yanan Zhao[1,†] and Tong Wu[1]

Abstract To account for population growth constrained by environmental capacity, we formulate a stochastic SIS model with logistic population input and standard incidence to describe disease transmission.Due to the degenerate structure of the stochastic differential equation, we employ the Lyapunov function method and Markov semigroup theory to establish a threshold theorem, including the extinction and ergodicity of the stochastic system.In addition, since the parameters in stochastic models are usually unknown in practice, we study parameter estimation for the proposed model under discrete observations by comparing least squares estimation, pseudo-maximum likelihood estimation (pseudo-MLE), and Bayesian posterior mean estimation. We further investigate how the observation step size, time horizon, and noise intensity affect the widths of confidence intervals and the accuracy of point estimates. Numerical simulations are provided to illustrate the theoretical results and to compare the performance of the three estimation methods.The results complement the dynamical analysis by providing a comparative study of parameter identification for this stochastic SIS model.

Keywords Stochastic SIS model, logistic population input, degenerate diffusion, threshold theorem, parameter estimation.

MSC(2010) 37A30, 47D07, 60H10, 92D30.

1. Introduction

In recent years, with the continuous outbreak of infectious diseases, the study of epidemic transmission dynamics using mathematical models has become an active research topic. Infectious diseases such as seasonal influenza, bacterial dysentery, and gonorrhoea can often be described by the classical SIS model, since recovered individuals may become susceptible again. Consider

$$\begin{cases} dS(t) = (rN(t) - \dfrac{\beta S(t)I(t)}{N(t)} - \mu S(t) + \delta I(t))dt, \\ dI(t) = (\dfrac{\beta S(t)I(t)}{N(t)} - (\mu + \delta + \varepsilon)I(t))dt. \end{cases} \tag{1.1}$$

Here $S(t)$, $I(t)$, and $N(t)$ denote the numbers of susceptible individuals, infected individuals, and the total population, respectively. The parameters r, β, μ, ε, and δ represent the intrinsic

[†]The corresponding author.
[1]School of Science, Changchun University, Changchun 130021, China
[*]The work was supported by Science and Technology Development Project of Jilin Province (YDZJ202401359ZYTS).
Email: 18834855595@163.com(X. Li), zhaoyn111@163.com(Y. Zhao), tutu123aaa@163.com(T. Wu)

growth rate, contact rate, natural death rate, disease-related death rate, and recovery rate. The standardized incidence rate in (1.1) is given by $\beta SI/N$, which has been regarded as more realistic than the bilinear incidence rate when the population is large [2,13]. In model (1.1), the term rN corresponds to exponential population growth.

In reality, however, population growth is constrained by limited environmental resources. To capture this effect, logistic population growth is introduced, so that the growth rate decreases as the population size approaches the environmental carrying capacity. This provides a more realistic description than models assuming exponential growth. Brauer [6] pointed out that environmental capacity may limit population density growth, and Gao and Hethcote [8] introduced SIRS and SIS epidemic models with density-dependent growth, showing that the persistence of infectious diseases and mortality may lead to new equilibrium population sizes below the carrying capacity, or even to population extinction. More recently, stochastic epidemic models with logistic growth have also attracted attention, and their extinction, persistence, and stationary distribution have been investigated under different incidence rates and noise structures [29,31,32]. These studies indicate that logistic population input is not only a biologically meaningful modification of the classical SIS model, but also an important factor affecting long-term population dynamics.

To account for environmental fluctuations, we consider the following stochastic SIS model:

$$\begin{cases} dS(t) = \left(rN(t)(1-\dfrac{N(t)}{K}) - \dfrac{\beta S(t)I(t)}{N(t)} - \mu S(t) + \delta I(t)\right)dt - \dfrac{\sigma S(t)I(t)}{N(t)}dB(t), \\ dI(t) = \left(\dfrac{\beta S(t)I(t)}{N(t)} - (\mu+\delta+\varepsilon)I(t)\right)dt + \dfrac{\sigma S(t)I(t)}{N(t)}dB(t), \end{cases} \tag{1.2}$$

where $B(t)$ is a standard Brownian motion with $B(0) = 0$ and $\sigma^2 > 0$ is the noise intensity. The total population satisfies

$$\frac{dN}{dt} = rN\left(1-\frac{N}{K}\right) - \mu N - \varepsilon I.$$

Since the contact rate directly determines the intensity of new infections, we introduce stochastic perturbation through this parameter. As a result, the diffusion coefficient is proportional to SI/N, which vanishes on the boundary of the positive region and therefore makes the diffusion term degenerate. Thus, the proposed model reflects both environmental constraints on population growth and random fluctuations in disease transmission. For stochastic epidemic models with non-degenerate diffusion, many results on persistence, extinction, and stationary distribution have been established by Khasminskii-type arguments and Lyapunov methods [14,17,18]. However, these approaches are not directly applicable to the present model. Recent studies have also shown that degenerate diffusion often requires different analytical tools, especially when stationary distribution and extinction are concerned [11,30]. Therefore, the analysis of threshold dynamics for the present SIS model with nonlinear logistic population input and degenerate diffusion is mathematically more delicate, and it motivates the use of Lyapunov function techniques together with integral Markov semigroup theory.

Besides long-term dynamical analysis, it is also important to estimate model parameters from discrete observations in practical applications, since quantities such as growth rate, mortality rate, and noise intensity are usually unknown. Parameter estimation for stochastic differential equations has been widely studied in the literature [4,15,28]. For stochastic epidemic models, Gray et al. [22] used pseudo-MLE and least squares estimation for a stochastic SIS model, Tang et al. [10] applied maximum likelihood estimation to a stochastic SIS model with environmental effects, and more recent Bayesian and simulation-based inference methods have also been considered [37]. Nevertheless, logistic population input, degenerate stochastic perturbation, and parameter estimation are often studied in separate settings, while their combined role in a stochastic SIS model has received relatively limited attention. Motivated by this, the present paper studies threshold dynamics and parameter estimation within the same stochastic SIS framework.

The main contributions are as follows. First, we formulate a stochastic SIS model with logistic population input and standard incidence, where perturbation of the contact rate leads to a degenerate diffusion term. Second, we establish a threshold theorem for the stochastic system, including disease extinction and ergodicity of the stationary distribution. Third, we investigate parameter estimation for the proposed model using least squares estimation, pseudo-MLE, and Bayesian posterior mean estimation, and compare the effects of observation step size, time horizon, and noise intensity on estimation accuracy.

The rest of the paper is organized as follows. Section 2 presents the existence and uniqueness result and the threshold theorem, together with numerical illustrations. Section 3 studies parameter estimation and discusses the main factors affecting estimation accuracy. Section 4 concludes the paper.

2. Main results

In this paper, (Ω, Σ, m) is a σ-finite measurable space, $E = \{f \in L^1(\Omega, \Sigma, m) : f \geq 0, \|f\| = 1\}$ contains all the densities and $\mathcal{P}(t, x, y, A)$ is the transition probability function of the diffusion process $(S(t), I(t))$ of system (1.2) for any $A \in \Sigma$.

Theorem 2.1. *For any given initial value $(S_0, I_0) \in G$, system (1.2) has a unique global positive solution $(S(t), I(t)) \in G$ for all $t \geq 0$ with probability one, where $G = \{(S(t), I(t)) \in \mathbb{R}^2_+ : 0 < S(t) + I(t) < K\}$ is a positive invariant set of system (1.2).*

Following the classical proof method, Theorem 2.1 can be proved by defining an appropriate function $V(S, I) = (S - 1 - \ln S) + (I - 1 - \ln I)$ as the routine in [36]. On this basis of the existence and uniqueness of positive solution, we show the threshold theorem between the ergodicity and extinction.

Remark 2.1. The diffusion coefficient in system (1.2) is proportional to $\frac{S(t)I(t)}{N(t)}$, which vanishes on the boundary of the positive domain. Hence, if $S_0 > 0$ and $I_0 > 0$, the solution remains in the biologically feasible region $S(t) > 0$, $I(t) > 0$ almost surely for all $t \geq 0$.

Theorem 2.2. *Let $(S(t), I(t))$ be the solution of system (1.2) with initial value $(S_0, I_0) \in G$.*

(i) Ergodicity: Let the distribution of $(S(t), I(t))$ have a density $p(t, x, y)$ that satisfies the Fokker–Planck equation associated with system (1.2). If the threshold $R_0^s := \frac{\beta}{\mu+\varepsilon+\delta+\frac{\sigma^2}{2}} > 1$, then there exists a unique density $p_*(x, y)$ satisfying

$$\lim_{t\to\infty} \iint_U |p(t, x, y) - p_*(x, y)| dxdy = 0,$$

where $p_*(x, y)$ is a stationary solution of the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2 \left[\frac{\partial^2(x^2y^2p)}{\partial x^2} - 2\frac{\partial^2(x^2y^2p)}{\partial x\partial y} + \frac{\partial^2(x^2y^2p)}{\partial y^2}\right] - \frac{\partial(f_1(x, y)p)}{\partial x} - \frac{\partial(f_2(x, y)p)}{\partial y}, \tag{2.1}$$

where $f_1(S,I) = rN(1-\frac{N}{K}) - \frac{\beta SI}{N} - \mu S + \delta I, \quad f_2(S,I) = \frac{\beta SI}{N} - (\mu+\delta+\varepsilon)I$, *and its support set is*

$$U = \operatorname{supp} p_* = \{(S(t), I(t)) \in G, \frac{K}{r}(r-\mu-\varepsilon) < S(t)+I(t) < \frac{K}{r}(r-\mu)\}. \tag{2.2}$$

(ii) Extinction: If $R_0^s < 1$ *and* $\sigma^2 \le \beta$, *or* $\sigma^2 > \max\{\frac{\beta^2}{2(\mu+\delta+\varepsilon)}, \beta\}$, *then the solution to system* (1.2) *satisfies:*

$$\limsup_{t\to\infty} \frac{\log I(t)}{t} \le -C < 0, \tag{2.3}$$

where C *is a positive constant.*

Proof. First of all, we will prove the ergodicity part in four steps by the method in [25].

Step 1. We prove that $\mathcal{P}(t, x_0, y_0, A)$ has a continuous density $\rho(t, x, y; x_0, y_0)$ satisfying $T(t)f(x,y) = \iint_G \rho(t,x,y;s,t)f(s,t)dsdt, f \in E$, then the semigroup $\{T(t)\}_{t\ge 0}$ is an integral Markov semigroup.

Let

$$\mathbf{m}[S,I] = (rN(1-\frac{N}{K}) - \frac{\beta SI}{N} - \mu S + \delta I,\ \frac{\beta SI}{N} - (\mu+\delta+\varepsilon)I)^T,\ \mathbf{n}[S,I] = (-\frac{\sigma SI}{N}, \frac{\sigma SI}{N})^T,$$

then the Lie bracket $[\mathbf{m}, \mathbf{n}]$ is defined as

$$[\mathbf{m}, \mathbf{n}]_j(x) = \sum_{k=1}^{d} \left(m_k \frac{\partial n_j}{\partial x_k}(x) - n_k \frac{\partial m_j}{\partial x_k}(x) \right), j = 1, 2, \ldots, d.$$

To calculate

$$|[\mathbf{m}, \mathbf{n}], \mathbf{n}| = \begin{vmatrix} m_1\frac{\partial n_1}{\partial S} - n_1\frac{\partial m_1}{\partial S} + m_2\frac{\partial n_1}{\partial I} - n_2\frac{\partial m_1}{\partial I} & n_1 \\ m_1\frac{\partial n_2}{\partial S} - n_1\frac{\partial m_2}{\partial S} + m_2\frac{\partial n_2}{\partial I} - n_2\frac{\partial m_2}{\partial I} & n_2 \end{vmatrix} = |(\frac{\sigma SI}{N})^2\delta| \ne 0.$$

It can be seen $\mathbf{n}(S,I)$, $[\mathbf{m},\mathbf{n}](S,I)$ are linearly independent for $(S,I) \in G$ and span $\mathbb{R}^2$. Hence the vector fields $\mathbf{m}$ and $\mathbf{n}$ satisfy the Hörmander condition [3] in the interior of the domain, which ensures the hypoellipticity of the process and supports the ergodicity analysis of the system.

Step 2. We prove there exists $T > 0$, satisfying $\rho(T, x, y; x_0, y_0) > 0$ for each $(x_0, y_0) \in U$ and $(x, y) \in U$, then $\int_0^\infty T(t)f dt > 0$ *a.e.* for each $f \in E$.

For $\xi \in L^2([0,T];\mathbb{R})$ and $(x_0, y_0) \in G$, we consider the integral equations of Stratonovitch type of system (1.2)

$$\begin{cases} S_\xi(t) = S_0 + \displaystyle\int_0^t [f_1(S_\xi(s), I_\xi(s)) - \frac{\sigma S_\xi(s) I_\xi(s)}{S_\xi(s)+I_\xi(s)}\xi(s)]ds, \\ I_\xi(t) = I_0 + \displaystyle\int_0^t [f_2(S_\xi(s), I_\xi(s)) + \frac{\sigma S_\xi(s) I_\xi(s)}{S_\xi(s)+I_\xi(s)}\xi(s)]ds, \end{cases} \tag{2.4}$$

where

$$f_1 = rN(1-\frac{N}{K}) - \frac{\beta SI}{N} - \mu S + \delta I + \frac{\sigma^2(S-I)}{2(S+I)^2}SI,\ f_2 = \frac{\beta SI}{N} - \frac{\sigma^2(S-I)}{2(S+I)^2}SI - (\mu+\delta+\varepsilon)I.$$

Firstly we define the Frechét derivative $C_{x_0,y_0;\xi}$: $l \mapsto \mathbf{z}_{\xi+l}(T) = (x_{\xi+l}(T), y_{\xi+l}(T))^T$, from $L^2([0,T];\mathbb{R})$ to $\mathbb{R}^2$. For $x = x_\xi(T)$, $y = y_\xi(T)$, in order to prove $\rho(T, x, y; x_0, y_0) > 0$, we need to prove $C_{x_0,y_0;\xi}$ has rank 2 for some ξ.

Let $l(t) = \frac{x_\xi(t)+y_\xi(t)}{x_\xi(t)y_\xi(t)}\mathbf{1}_{[T-\epsilon,T]}$, $t \in [0,T]$, where $\mathbf{1}_{[T-\epsilon,T]}$ is the characteristic function, $\epsilon \in (0,T)$, then $C_{x_0,y_0;\xi}l = \epsilon\mathbf{r} + \frac{1}{2}\epsilon^2\Lambda(T)\mathbf{r} + o(\epsilon^2)$, $\quad \mathbf{r} = (-\sigma, \sigma)^T$, and

$$\Lambda(T)\mathbf{r} = \begin{pmatrix} \sigma(-f'_{1S} + f'_{1I}) - \sigma^3(\frac{\partial b}{\partial S} - \frac{\partial b}{\partial I}) + \sigma^2\xi(\frac{\partial a}{\partial S} - \frac{\partial a}{\partial I}) \\ \sigma(-f'_{2S} + f'_{2I}) + \sigma^3(\frac{\partial b}{\partial S} - \frac{\partial b}{\partial I}) - \sigma^2\xi(\frac{\partial a}{\partial S} - \frac{\partial a}{\partial I}) \end{pmatrix},$$

where $a = SI/(S+I)$, $b = (S-I)SI/2(S+I)^2$. Hence, we get $C_{x_0,y_0;\xi}$ has rank 2 according to $\mathbf{r}$ and $\Lambda(T)\mathbf{r}$ are linearly independent.

Secondly we need to prove that for any $(x_0, y_0), (x, y) \in U$ and $T > 0$, there exists a control function ξ satisfying $(x_\xi(0), y_\xi(0)) = (x_0, y_0)$ and $(x_\xi(T), y_\xi(T)) = (x, y)$. To facilitate the statement, we replace system (2.4) with the following differential equations.

$$\begin{cases} x'_\xi = f_1(x_\xi, y_\xi) - \sigma\xi\dfrac{x_\xi y_\xi}{x_\xi + y_\xi}, \\ y'_\xi = f_2(x_\xi, y_\xi) + \sigma\xi\dfrac{x_\xi y_\xi}{x_\xi + y_\xi}. \end{cases}$$

Let $z_\xi = x_\xi + y_\xi$, then

$$\begin{cases} x'_\xi = g_1(x_\xi, z_\xi) - \sigma\xi\dfrac{x_\xi(z_\xi - x_\xi)}{z_\xi}, \\ z'_\xi = g_2(x_\xi, z_\xi), \end{cases}$$

where $g_1(x,z) = f_1(x, z-x)$, $\quad g_2(x,z) = rz(1 - \frac{z}{K}) - \mu z - \varepsilon(z-x)$. we obtain

$$x_\xi = z_\xi + \frac{1}{\alpha}[z'_\xi - rz_\xi(1 - \frac{z_\xi}{K}) + \mu z_\xi], \ t \in [0,T], \tag{2.5}$$

and

$$rz_\xi(1 - \frac{z_\xi}{K}) - (\mu+\varepsilon)z_\xi < z'_\xi < rz_\xi(1 - \frac{z_\xi}{K}) - \mu z_\xi, \ t \in [0,T]. \tag{2.6}$$

Let $U_0 = \{(x,z) \in R^2_+ : 0 < x < K, \frac{K}{r}(r - \varepsilon - \mu) < z < \frac{K}{r}(r-\mu) \ \text{ and } \ x < z\}$. Based on the variable transformation, we just need to prove there exists a control function ξ satisfying $(x_\xi(0), z_\xi(0)) = (x_0, z_0)$ and $(x_\xi(T), z_\xi(T)) = (x, z)$ in U_0.

We divide z_ξ on $[0, \varsigma_1]$, $(\varsigma_1, T - \varsigma_2)$, $[T - \varsigma_2, T]$. Construct a C^2-function

$$z_\xi : [0, \varsigma_1] \to (\lambda - K\frac{\varepsilon}{r} + \theta, \lambda - \theta) \tag{2.7}$$

where $\lambda = \frac{K}{r}(r-\mu)$, $\theta = \frac{1}{2}\min\{z_0 - \lambda + K\frac{\varepsilon}{r}, z_1 - \lambda + K\frac{\varepsilon}{r}, \lambda - z_0, \lambda - z_1\}$. Then we have

$$-\frac{r\theta^2}{K} + (r - \varepsilon - \mu)\theta < z'_\xi < -\frac{r\theta^2}{K} + (r-\mu)\theta \tag{2.8}$$

where $-\frac{r\theta^2}{K} + (r - \mu - \varepsilon)\theta < 0$ and $-\frac{r\theta^2}{K} + (r-\mu)\theta > 0$. So we can construct z_ξ such that $z_\xi(0) = z_0$, $\ z'_\xi(0) = g_2(x_0, z_0)$, $\ z'_\xi(\varsigma_1) = 0$, $t \in [0, \varsigma_1]$ and satisfies inequality (2.6). Similarly, we apply the method to the interval $[T - \varsigma_2, T]$. Let T be sufficiently large, then z_ξ can be extended from $[0,\varsigma_1]\cup[T-\varsigma_2,T]$ to the entire interval $[0,T]$ satisfying $z_\xi(0)=z_0$, $z'_\xi(0)=g_2(x_0,z_0)$, $z_\xi(T)=z$, $z'_\xi(T)=g_2(x,z)$ and (2.6). Thus, according to (2.5), we can find a ξ and T satisfying $(x_\xi(0),y_\xi(0))=(x_0,y_0)$ and $(x_\xi(T),y_\xi(T))=(x,y)$.

Step 3. We prove that if $R_0^s>1$, then $\lim_{t\to\infty}\iint_U T(t)f(x,y)dxdy=1$. Considering system (1.2), we obtain

$$\frac{dN(t)}{dt}=rN(t)(1-\frac{N(t)}{K})-(\mu+\varepsilon)N(t)+\varepsilon S(t), \tag{2.9}$$

and

$$rN(1-\frac{N}{K})-(\mu+\varepsilon)N<N'(t)<rN(1-\frac{N}{K})-\mu N,\quad (S,I)\in U. \tag{2.10}$$

Now we just need to prove U is the support set, namely (2.2). That is, there exists $t_0=t_0(\omega)$ for almost every $\omega\in\Omega$, such that

$$\lambda-K\frac{\varepsilon}{r}<N_t(\omega)<\lambda,\quad \text{for } t>t_0. \tag{2.11}$$

When $N_0\in(\lambda-K\frac{\varepsilon}{r},\lambda)$, by (2.10) we can obtain the desired conclusion.

When $N_0\in(0,\lambda-K\frac{\varepsilon}{r})$, if (2.11) is not valid, according to (2.10), there exists $\Omega_1\subset\Omega$ and $\text{Prob}(\Omega_1)>0$, then $N_t(\omega)\in(0,\lambda-K\frac{\varepsilon}{r}),\omega\in\Omega_1$. It can be known $N_t(\omega)$ is strictly increasing and $\lim_{t\to\infty}N_t(\omega)=\lambda-K\frac{\varepsilon}{r}$. So $\lim_{t\to\infty}S_t(\omega)=0$ and $\lim_{t\to\infty}I_t(\omega)=\lambda-K\frac{\varepsilon}{r}$, $\omega\in\Omega_1$ on the basis of $g_2(x,z)$. Then it yields from system (1.2)

$$\frac{\log I-\log I_0}{t}=\frac{\beta}{t}\int_0^t\frac{S}{N}dr-(\mu+\delta+\varepsilon)-\frac{\sigma^2}{2t}\int_0^t\frac{S^2}{N^2}dr+\frac{\sigma}{t}\int_0^t\frac{S}{N}dB(r), \tag{2.12}$$

then

$$\lim_{t\to\infty}\left[\frac{\beta}{t}\int_0^t\frac{S}{N}dr-(\mu+\delta+\varepsilon)-\frac{\sigma^2}{2t}\int_0^t\frac{S^2}{N^2}dr+\frac{\sigma}{t}\int_0^t\frac{S}{N}dB(r)\right]=-(\mu+\delta+\varepsilon),\ \text{a.s.}$$

On the other hand, by (2.12), we have $\lim_{t\to\infty}\frac{\log I}{t}=0$ on Ω_1, which is a contradiction. Similarly, when $N_0\in(\lambda-K\frac{\varepsilon}{r},\lambda)$, (2.11) can be proved. Therefore, our claim (2.11) holds.

Step 4. We prove that if $R_0^s>1$, then $\{T(t)\}_{t\geq 0}$ relative to compact sets is asymptotically stable. Construct a Lyapunov function $V:U\to\bar{\mathbb{R}}_+$ by

$$V(S(t),I(t))=-\log S(t)-\log(\lambda-S(t)-I(t))-l\log I(t)-\log(S(t)+I(t)-\lambda+K\frac{\varepsilon}{r})-V(\theta,r\theta). \tag{2.13}$$

Fixed $l=(\beta+\mu+2r+2+\frac{\sigma^2}{2})/(\beta-(\mu+\varepsilon+\frac{\sigma^2}{2}))>0$, such that

$$\beta+\mu+2r-l[\beta-(\mu+\varepsilon+\frac{\sigma^2}{2})]+\frac{\sigma^2}{2}=-2, \tag{2.14}$$

and $V(\theta,r\theta)$, $\theta=K(r-\mu)/r(l+1)$ guarantees the positivity of the function $V(S,I)$. We get

$$\begin{aligned}\mathscr{L}^*V\leq&[-\frac{rN(1-\frac{N}{K})}{S}+\beta+\mu+\frac{\sigma^2}{2}]+[-\beta+(\mu+\varepsilon+\frac{\sigma^2}{2})+\frac{\beta}{\lambda-K\frac{\varepsilon}{r}}I]\\&+[\frac{rN}{K}-\frac{dI}{\lambda-N}]+[\frac{r}{K}N-\frac{\varepsilon S}{N-\lambda+K\frac{\varepsilon}{r}}].\end{aligned}$$

For $(S, I) \in U$, define

$$U_{\kappa,\epsilon} = \{\kappa < S < \lambda - \kappa, \kappa < I < \lambda - \kappa, \lambda - K\frac{\varepsilon}{r} + \epsilon < S + I < \lambda - \epsilon\}.$$

Here κ is a sufficient small positive number and $\epsilon = \kappa^2$, such that (i)$-\frac{f_{\min}}{\kappa} + 2\beta + 2\mu + 2r + \varepsilon + \sigma^2 < -1$, f_{min} is the minimum value for $f(N) = N(1 - \frac{N}{K}), (S, I) \in U$ (ii)$\frac{l\beta\kappa}{\lambda - K\frac{\varepsilon}{r}} < -1$, and (iii)$-\frac{\varepsilon}{\kappa} + 2\beta + 2\mu + 2r + \varepsilon + \sigma^2 < -1$.

Let $A^1_{\kappa,\epsilon} = \{(S, I) \in U : 0 < S < \kappa\}$, and $A^2_{\kappa,\epsilon} = \{(S, I) \in U : 0 < I < \kappa\}$,

$$A^3_{\kappa,\epsilon} = \{(S, I) \in U : \kappa \le S < \lambda, \kappa \le I < \lambda, \lambda - \kappa^2 < S + I < \lambda\},$$
$$A^4_{\kappa,\epsilon} = \{(S, I) \in U : \kappa \le S < \lambda, \kappa \le I < \lambda, \lambda - K\frac{\varepsilon}{r} < S + I < \lambda - K\frac{\varepsilon}{r} + \kappa^2\}.$$

Then $U \setminus U_{\kappa,\epsilon} = A^1_{\kappa,\epsilon} \bigcup A^2_{\kappa,\epsilon} \bigcup A^3_{\kappa,\epsilon} \bigcup A^4_{\kappa,\epsilon}$. So we consider the four kinds of situations:

$$\begin{aligned}
&\mathscr{L}^*V \le -\frac{f_{\min}}{\kappa} + \beta + \mu + \frac{\sigma^2}{2} + \mu + \varepsilon + \frac{\sigma^2}{2} + \beta + r + r < -1, (S, I) \in A^1_{\kappa,\epsilon},\\
&\mathscr{L}^*V \le \beta + \mu + \frac{\sigma^2}{2} - l[\beta - (\mu + \varepsilon + \frac{\sigma^2}{2})] + \frac{l\beta\kappa r}{\lambda r - K\varepsilon} + r + r < -1, (S, I) \in A^2_{\kappa,\epsilon},\\
&\mathscr{L}^*V \le \beta + \mu + \frac{\sigma^2}{2} + \mu + \varepsilon + \frac{\sigma^2}{2} + \beta + (-\frac{\varepsilon}{\kappa} + r) + r < -1, (S, I) \in A^3_{\kappa,\epsilon},\\
&\mathscr{L}^*V \le \beta + \mu + \frac{\sigma^2}{2} + \mu + \varepsilon + \frac{\sigma^2}{2} + \beta + r + (-\frac{\varepsilon}{\kappa} + r) < -1, (S, I) \in A^4_{\kappa,\epsilon}.
\end{aligned}$$

To sum up, we get $\sup_{x \in U \setminus U_{\kappa,\epsilon}} \mathscr{L}^*V(x) < 0$. With the discussion in the literature [23], there exists the Khasminskii function $V(S, I)$. Therefore $\{T(t)\}_{t \ge 0}$ relative to compact sets is asymptotically stable, and there's only a stationary solution on U.

Finally, we will prove the extinction part.

Case 1. If $R^s_0 < 1$, $\sigma^2 \le \beta$, then

$$\begin{aligned}
\frac{\log I(t) - \log I_0}{t} &= \frac{1}{t}\int_0^t [\frac{\beta S}{N} - (\mu + \delta + \varepsilon) - \frac{\sigma^2 S^2}{2N^2}]ds + \frac{1}{t}\int_0^t \frac{\sigma S}{N} dB(s)\\
&\le \frac{1}{t}\int_0^t [(R^s_0 - 1)(\mu + \delta + \varepsilon + \frac{\sigma^2}{2})]ds + \frac{1}{t}\int_0^t \frac{\sigma S}{N} dB(s).
\end{aligned} \tag{2.15}$$

Case 2. If $\sigma^2 > \max\{\frac{\beta^2}{2(\mu+\delta+\varepsilon)}, \beta\}$, considering the following inequality

$$\begin{aligned}
\frac{\log I(t) - \log I_0}{t} &\le \frac{1}{t}\int_0^t [-\frac{\sigma^2}{2}(\frac{S}{N} - \frac{\beta}{\sigma^2})^2 + \frac{\beta^2}{2\sigma^2} - (\mu + \delta + \varepsilon)]ds + \frac{1}{t}\int_0^t \frac{\sigma S}{N} dB(s)\\
&\le \frac{1}{t}\int_0^t [\frac{\beta^2}{2\sigma^2} - (\mu + \delta + \varepsilon)]ds + \frac{1}{t}\int_0^t \frac{\sigma S}{N} dB(s).
\end{aligned} \tag{2.16}$$

Thus, applying the limit superior of (2.15) and (2.16), and according to the first equation of the stochastic model (1.2), we get

$$\limsup_{t\to\infty} \frac{\log I(t)}{t} \le -a < 0 \quad \text{and} \quad \lim_{t\to\infty} \frac{1}{t}\int S(t)dt = \frac{(r - \mu)K}{r} \quad \text{a.s.} \tag{2.17}$$

a is positive constant. This completes the proof. □

The following examples illustrate our results. From the theory introduced in [9], the EM approximate solution converges in probability to the true solutions of models (1.2) and its corresponding deterministic model.

Example 2.1. We use the Euler-Maruyama (EM) method to perform 10,000 time steps for model (1.2) with the initial value of $(S(0), I(0)) = (0.8,\ 0.4)^T$. The step size is $\Delta t = 0.001$ and the parameters are given by

$$K = 1,\ r = 0.2,\ \mu = 0.1,\ \beta = 0.5,\ \delta = 0.2,\ \varepsilon = 0.3,\ \sigma = 0.4. \tag{2.18}$$

This set of parameters satisfies the condition $R_0^s < 1$ and $\sigma^2 \leq \beta$ of Theorem 2.2 (ii). Hence, the conditions of Theorem 2.2(ii) are satisfied. Therefore, the disease becomes extinct with probability 1. In the sense of time mean, the limit of the number of infected persons approaches a constant. As shown in Figure 1.

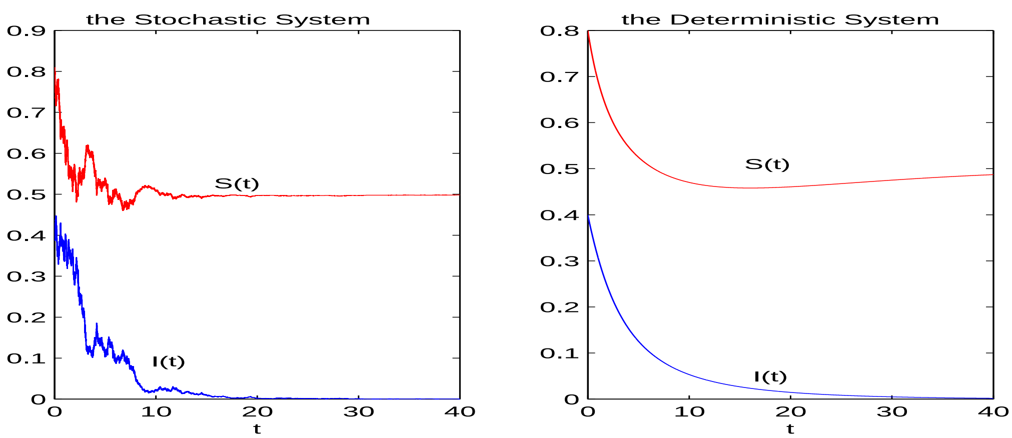

Figure 1. Numerical simulations of the paths $S(t)$ and $I(t)$ of SDE model (1.2) using the EM scheme with stepsize $\Delta t = 0.001$ and an initial value $(S(0), I(0)) = (0.8,\ 0.4)^T$ with the system parameters provided by (2.18).

Example 2.2. We assume that the parameters of system (1.2) are given by

$$K = 1,\ r = 0.2,\ \mu = 0.1,\ \beta = 0.8,\ \delta = 0.2,\ \varepsilon = 0.3,\ \sigma = 0.4. \tag{2.19}$$

Hence, the condition of Theorem 2.2(i) is satisfied. Therefore, the stochastic system admits a unique stationary distribution, and the disease persists in the long run. Figure 2 illustrates this persistent behavior.

Example 2.3. We make the system parameters of model (1.2) the same as in Example 2.2, but $\sigma = 0.9$ is given. This set of parameters satisfies the condition $\sigma^2 > \max\{\frac{\beta^2}{2(\mu+\delta+\varepsilon)}, \beta\}$ of Theorem 2.2 (ii). It shows that the disease tends to extinction almost surely. However, for the corresponding deterministic model we obtain $R_0 > 1$, indicating that local equilibria in Γ are globally asymptotically stable. We provide simulations, shown in Figure 3, to support the extinction conclusion in Theorem 2.2 that the large white noise can suppress infectious disease outbreaks.

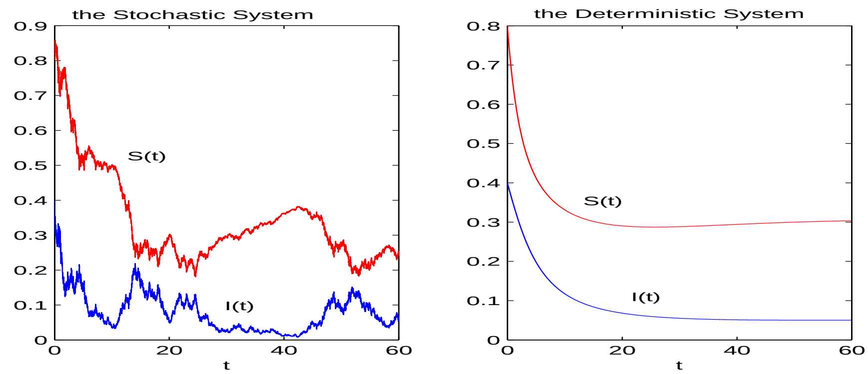

Figure 2. Numerical simulations of the paths $S(t)$ and $I(t)$ of SDE model (1.2) using the EM scheme with stepsize $\Delta t = 0.001$ and an initial value $(S(0), I(0)) = (0.8,\ 0.4)^T$ with the system parameters provided by (2.19).

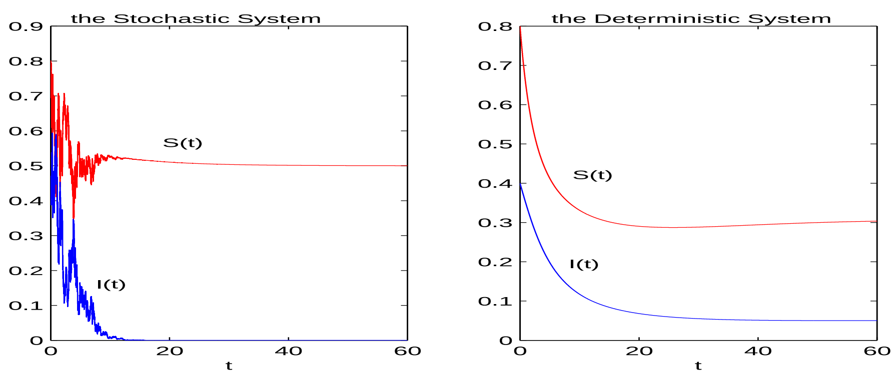

Figure 3. Numerical simulations of the paths $S(t)$ and $I(t)$ of SDE model (1.2) using the EM scheme with stepsize $\Delta t = 0.001$ and an initial value $(S(0), I(0)) = (0.8,\ 0.4)^T$ with the system parameters provided by (2.19), but $\sigma = 0.9$, such that $R_0 > 1$.

3. Parameter estimation

3.1. Least squares estimation

In this section, we first approximate the model using the Euler-Maruyama (EM) scheme [19,20], which transforms the stochastic differential equation system into a form suitable for regression analysis. Since the discretized equations share the same Brownian increment, the resulting regression framework should be interpreted as an approximation under discrete sampling. We then apply regression methods to estimate the model parameters and obtain both point estimates and $100(1-\alpha)\%$ confidence intervals. We also investigate factors that influence the length of these confidence intervals. Finally, numerical simulations are performed to illustrate the practical implications of the proposed theoretical results.

3.1.1. Regression model

Assuming that $(\{S_k\}_{k=0}^n, \{I_k\}_{k=0}^n)^T$ represents the observed data from the process described by equation (1.2), and given a step size of Δt and initial settings $(S_0 = S(0), I_0 = I(0))^T$, we apply the EM scheme to discretize the interval $[k\Delta t, (k+1)\Delta t]$. The resulting discretization is as follows

$$\begin{cases} S_{k+1} - S_k = (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t - \frac{\sigma S_k I_k}{N_k}\Delta W_k, \\ I_{k+1} - I_k = (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t + \frac{\sigma S_k I_k}{N_k}\Delta W_k, \end{cases} \tag{3.1}$$

where $\Delta W_k = W_{k+1} - W_k$.

Equation (3.1) can be rewritten as

$$\begin{cases} {y^1}_{k+1} = r{u^1}_{k+1} + \mu {u^2}_{k+1} + \delta {u^3}_{k+1} - c - \sigma Z_{k+1}, \\ {y^2}_{k+1} = \eta v_{k+1} + c + \sigma Z_{k+1}, \end{cases} \tag{3.2}$$

where ${y^1}_{k+1} = \frac{(S_{k+1}-S_k)N_k}{S_k I_k \sqrt{\Delta t}}$, ${u^1}_{k+1} = \frac{{N_k}^2(1-\frac{N_k}{K})\sqrt{\Delta t}}{S_k I_k}$, ${u^2}_{k+1} = -\frac{N_k\sqrt{\Delta t}}{I_k}$, ${u^3}_{k+1} = \frac{N_k\sqrt{\Delta t}}{S_k}$, ${y^2}_{k+1} = \frac{(I_{k+1}-I_k)N_k}{S_k I_k \sqrt{\Delta t}}$, $\eta = \mu + \delta + \varepsilon$, $v_{k+1} = -\frac{N_k\sqrt{\Delta t}}{S_k}$, $c = \sqrt{\Delta t}\beta$, $Z_{k+1} \sim N(0,1)$.

This appears to be a simple linear regression model where y^1 and y^2 are random variables. Regression theory can still be used to estimate the parameters β, r, μ, δ and η, as the estimates are obtained by the method of least squares, which involves minimising criterion function

$$\sum_{k=0}^{n-1} ({y^1}_{k+1} - r{u^1}_{k+1} - \mu {u^2}_{k+1} - \delta {u^3}_{k+1} + c)^2$$

and

$$\sum_{k=0}^{n-1} ({y^2}_{k+1} - \eta v_{k+1} - c)^2.$$

Rawlings [24] provided a detailed discussion of multivariate linear regression problems using the general matrix form

$$Y = X\theta + \varepsilon, \tag{3.3}$$

where

$$Y = \begin{pmatrix} Y^1 \\ Y^2 \end{pmatrix}, X = \begin{pmatrix} {X^1}_{(n*4)} & \mathrm{O}_{(n*2)} \\ \mathrm{O}_{(n*4)} & {X^2}_{(n*2)} \end{pmatrix}, \theta = \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix}, \varepsilon = (\varepsilon_1, \cdots, \varepsilon_n, \varepsilon_1, \cdots, \varepsilon_n)^T.$$

The submatrices in the Y, X and θ are

$$Y^1 = \begin{pmatrix} {y^1}_{(1)} \\ {y^1}_{(2)} \\ \vdots \\ {y^1}_{(n)} \end{pmatrix}, {X^1}_{(n*4)} = \begin{pmatrix} -\sqrt{\Delta t} & {u^1}_{(1)} & {u^2}_{(1)} & {u^3}_{(1)} \\ -\sqrt{\Delta t} & {u^1}_{(2)} & {u^2}_{(2)} & {u^3}_{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ -\sqrt{\Delta t} & {u^1}_{(n)} & {u^2}_{(n)} & {u^3}_{(n)} \end{pmatrix}, \theta^1 = \begin{pmatrix} \beta \\ r \\ \mu \\ \delta \end{pmatrix},$$

$$Y^2 = \begin{pmatrix} {y^2}_{(1)} \\ {y^2}_{(2)} \\ \vdots \\ {y^2}_{(n)} \end{pmatrix}, {X^2}_{(n*4)} = \begin{pmatrix} \sqrt{\Delta t} & v_{(1)} \\ \sqrt{\Delta t} & v_{(2)} \\ \vdots & \vdots \\ \sqrt{\Delta t} & v_{(n)} \end{pmatrix}, \theta^2 = \begin{pmatrix} \beta \\ \eta \end{pmatrix},$$

$$\mathrm{O}_{(n*2)} = \begin{pmatrix} 0\,0 \\ 0\,0 \\ \vdots\,\vdots \\ 0\,0 \end{pmatrix}, \mathrm{O}_{(n*4)} = \begin{pmatrix} 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ \vdots\,\vdots\,\vdots\,\vdots\,\vdots \\ 0\,0\,0\,0 \end{pmatrix}.$$

3.1.2. Point estimators

We can utilize the formulae from multiple linear regression theory to derive the estimators for θ^1 and θ^2 as

$$\begin{aligned}\hat{\theta} &= \begin{pmatrix} \hat{\theta}^1 \\ \hat{\theta}^2 \end{pmatrix} \\ &= (X^T X)^{-1}(X^T Y) \\ &= \left[\begin{pmatrix} (X^1)^T & \mathrm{O}_{(n*2)} \\ \mathrm{O}_{(n*4)} & (X^2)^T \end{pmatrix} \begin{pmatrix} X^1 & \mathrm{O}_{(n*2)} \\ \mathrm{O}_{(n*4)} & X^2 \end{pmatrix}\right]^{-1} \begin{bmatrix} (X^1)^T Y^1 \\ (X^2)^T Y^2 \end{bmatrix} \\ &= \begin{bmatrix} \left((X^1)^T X^1\right)^{-1} \left((X^1)^T Y^1\right) \\ \left((X^2)^T X^2\right)^{-1} \left((X^2)^T Y^2\right) \end{bmatrix}.\end{aligned} \tag{3.4}$$

The point estimators for θ^1 and θ^2 are calculated as

$$\begin{aligned}\hat{\theta}^1 &= \begin{pmatrix} \hat{\beta} \\ \hat{r} \\ \hat{\mu} \\ \hat{\delta} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{A_{11}}{A}\sqrt{\Delta t}\sum {y^1}_{(k)} + \frac{A_{21}}{A}\sum {u^1}_{(k)}{y^1}_{(k)} + \frac{A_{31}}{A}\sum {u^2}_{(k)}{y^1}_{(k)} + \frac{A_{41}}{A}\sum {u^3}_{(k)}{y^1}_{(k)} \\ -\frac{A_{12}}{A}\sqrt{\Delta t}\sum {y^1}_{(k)} + \frac{A_{22}}{A}\sum {u^1}_{(k)}{y^1}_{(k)} + \frac{A_{32}}{A}\sum {u^2}_{(k)}{y^1}_{(k)} + \frac{A_{42}}{A}\sum {u^3}_{(k)}{y^1}_{(k)} \\ -\frac{A_{13}}{A}\sqrt{\Delta t}\sum {y^1}_{(k)} + \frac{A_{23}}{A}\sum {u^1}_{(k)}{y^1}_{(k)} + \frac{A_{33}}{A}\sum {u^2}_{(k)}{y^1}_{(k)} + \frac{A_{43}}{A}\sum {u^3}_{(k)}{y^1}_{(k)} \\ -\frac{A_{14}}{A}\sqrt{\Delta t}\sum {y^1}_{(k)} + \frac{A_{24}}{A}\sum {u^1}_{(k)}{y^1}_{(k)} + \frac{A_{34}}{A}\sum {u^2}_{(k)}{y^1}_{(k)} + \frac{A_{44}}{A}\sum {u^3}_{(k)}{y^1}_{(k)} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{A_{11}}{A}\sum \frac{a_k N_k}{S_k I_k} + \frac{A_{21}}{A}\sum \frac{(N_k)^3 a_k h_k}{(S_k)^2 (I_k)^2} - \frac{A_{31}}{A}\sum \frac{(N_k)^2 a_k}{S_k (I_k)^2} + \frac{A_{41}}{A}\sum \frac{(N_k)^2 a_k}{(S_k)^2 I_k} \\ -\frac{A_{12}}{A}\sum \frac{a_k N_k}{S_k I_k} + \frac{A_{22}}{A}\sum \frac{(N_k)^3 a_k h_k}{(S_k)^2 (I_k)^2} - \frac{A_{32}}{A}\sum \frac{(N_k)^2 a_k)}{S_k (I_k)^2} + \frac{A_{42}}{A}\sum \frac{(N_k)^2 a_k}{(S_k)^2 I_k} \\ -\frac{A_{13}}{A}\sum \frac{a_k N_k}{S_k I_k} + \frac{A_{23}}{A}\sum \frac{(N_k)^3 a_k h_k}{(S_k)^2 (I_k)^2} - \frac{A_{33}}{A}\sum \frac{(N_k)^2 a_k}{S_k (I_k)^2} + \frac{A_{43}}{A}\sum \frac{(N_k)^2 a_k}{(S_k)^2 I_k} \\ -\frac{A_{14}}{A}\sum \frac{a_k N_k}{S_k I_k} + \frac{A_{24}}{A}\sum \frac{(N_k)^3 a_k h_k}{(S_k)^2 (I_k)^2} - \frac{A_{34}}{A}\sum \frac{(N_k)^2 a_k}{S_k (I_k)^2} + \frac{A_{44}}{A}\sum \frac{(N_k)^2 a_k}{(S_k)^2 I_k} \end{pmatrix}\end{aligned}$$

and

$$\hat{\theta}^2 = \begin{pmatrix} \hat{\beta} \\ \hat{\eta} \end{pmatrix} = \begin{pmatrix} \frac{B_{11}}{B}\sqrt{\Delta t}\sum y^2{}_{(k)} + \frac{B_{21}}{B}\sum v_k y^2{}_{(k)} \\ \frac{B_{12}}{B}\sqrt{\Delta t}\sum y^2{}_{(k)} + \frac{B_{22}}{B}\sum v_k y^2{}_{(k)} \end{pmatrix} = \begin{pmatrix} \frac{B_{11}}{B}\sum \frac{b_k N_k}{S_k I_k} - \frac{B_{21}}{B}\sum \frac{(N_k)^2 b_k}{(S_k)^2 I_k} \\ \frac{B_{12}}{B}\sum \frac{b_k N_k}{S_k I_k} - \frac{B_{22}}{B}\sum \frac{(N_k)^2 b_k}{(S_k)^2 I_k} \end{pmatrix},$$

where

$$A = \left|(X^1)^T X^1\right| = \begin{vmatrix} n\sqrt{\Delta t} & -\sqrt{\Delta t}\sum u^1{}_{(k)} & -\sqrt{\Delta t}\sum u^2{}_{(k)} & -\sqrt{\Delta t}\sum u^3{}_{(k)} \\ -\sqrt{\Delta t}\sum u^1{}_{(k)} & \sum (u^1{}_{(k)})^2 & \sum u^1{}_{(k)} u^2{}_{(k)} & \sum u^1{}_{(k)} u^3{}_{(k)} \\ -\sqrt{\Delta t}\sum u^2{}_{(k)} & \sum u^2{}_{(k)} u^1{}_{(k)} & \sum (u^2{}_{(k)})^2 & \sum u^2{}_{(k)} u^3{}_{(k)} \\ -\sqrt{\Delta t}\sum u^3{}_{(k)} & \sum u^3{}_{(k)} u^1{}_{(k)} & \sum u^3{}_{(k)} u^2{}_{(k)} & \sum (u^3{}_{(k)})^2 \end{vmatrix}$$

$$= \begin{vmatrix} n\sqrt{\Delta t} & -\Delta t\sum \frac{(N_k)^2 h_k}{S_k I_k} & \Delta t \sum \frac{N_k}{I_k} & -\Delta t\sum \frac{N_k}{S_k} \\ -\Delta t\sum \frac{(N_k)^2 h_k}{S_k I_k} & \Delta t\sum \left(\frac{(N_k)^2 h_k}{S_k I_k}\right)^2 & -\Delta t\sum \frac{(N_k)^3 h_k}{S_k (I_k)^2} & \Delta t\sum \frac{(N_k)^3 h_k}{(S_k)^2 I_k} \\ \Delta t\sum \frac{N_k}{I_k} & -\Delta t\sum \frac{(N_k)^3 h_k}{S_k (I_k)^2} & \Delta t\sum \left(\frac{N_k}{I_k}\right)^2 & -\Delta t\sum \frac{(N_k)^2}{S_k I_k} \\ -\Delta t\sum \frac{N_k}{S_k} & \Delta t\sum \frac{(N_k)^3 h_k}{(S_k)^2 I_k} & -\Delta t\sum \frac{(N_k)^2}{S_k I_k} & \Delta t\sum \left(\frac{N_k}{S_k}\right)^2 \end{vmatrix}$$

and

$$B = \left|(X^2)^T X^2\right| = \begin{vmatrix} n\Delta t & \sqrt{\Delta t}\sum v_k \\ \sqrt{\Delta t}\sum v_k & \sum (v_k)^2 \end{vmatrix} = \begin{vmatrix} n\Delta t & -\Delta t\sum \frac{N_k}{S_k} \\ -\Delta t\sum \frac{N_k}{S_k} & \Delta t\sum (\frac{N_k}{S_k})^2 \end{vmatrix},$$

$\{A_{ij}, i, j = 1, 2, 3, 4\}$ are the cofactor of the determinant A, $\{B_{ij}, i, j = 1, 2, 3, 4\}$, are the cofactor of the determinant B, $a_k = S_{k+1} - S_k, b_k = I_{k+1} - I_k,\ h_k = 1 - \frac{N_k}{K}$, Σ represent $\sum_{k=0}^{n-1}$.

Example 3.1. Assume the initial value is $(S(0), I(0)) = (0.8, 0.4)^T$ and the parameters of model (1.2) are

$$T = 10,\ \beta = 0.5,\ \mu = 0.1,\ K = 1,\ r = 0.2,\ \varepsilon = 0.3,\ \delta = 0.2,\ \sigma = 0.02.$$

We employ the EM method to simulate the $(S(t), I(t))^T$ process using the aforementioned parameters. The time step used is very small, $\Delta t = 0.001$, and we save these $(S(t), I(t))$ values as our true data set. Subsequently, we sample every 10th data point from this data set to obtain a sample for parameter estimation, resulting in n=1000 observations with $\Delta t = 0.01$. Using this sample, we obtain point estimates for the parameters, $\hat{\theta}^1 = (0.60, 0.23, 0.12, 0.30)^T$ and $\hat{\theta}^2 = (0.46, 0.57)^T$.

The true value of the parameter β is 0.5, and the numerical simulation results indicate that it is more efficient to estimate the parameter β using Y^2 and X^2 as variables. Therefore, in the least squares estimation, we first use the second equation of (3.2) to estimate the parameter β, and then substitute the estimated value of β into the first equation of (3.2) to estimate the parameter $\hat{\theta}^{1'} = (r, \mu, \delta)^T$. Then we have point estimators as

$$\hat{\theta}^{1'} = \begin{pmatrix} \hat{r} \\ \hat{\mu} \\ \hat{\delta} \end{pmatrix} \tag{3.5}$$

$$= ((X^{1'})^T X^{1'})^{-1}((X^{1'})^T Y^{1'})$$

$$= \begin{pmatrix} \frac{C_{11}}{C}\sum u^1{}_{(k)} y^{1'}{}_{(k)} + \frac{C_{21}}{C}\sum u^2{}_{(k)} y^{1'}{}_{(k)} + \frac{C_{31}}{C}\sum u^3{}_{(k)} y^{1'}{}_{(k)} \\ \frac{C_{12}}{C}\sum u^1{}_{(k)} y^{1'}{}_{(k)} + \frac{C_{22}}{C}\sum u^2{}_{(k)} y^{1'}{}_{(k)} + \frac{C_{32}}{C}\sum u^3{}_{(k)} y^{1'}{}_{(k)} \\ \frac{C_{13}}{C}\sum u^1{}_{(k)} y^{1'}{}_{(k)} + \frac{C_{23}}{C}\sum u^2{}_{(k)} y^{1'}{}_{(k)} + \frac{C_{33}}{C}\sum u^3{}_{(k)} y^{1'}{}_{(k)} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{C_{11}}{C}\sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - \frac{C_{21}}{C}\sum \frac{N_k g_k}{S_k (I_k)^2} + \frac{C_{31}}{C}\sum \frac{N_k g_k}{(S_k)^2 I_k} \\ \frac{C_{12}}{C}\sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - \frac{C_{22}}{C}\sum \frac{N_k g_k}{S_k (I_k)^2} + \frac{C_{32}}{C}\sum \frac{N_k g_k}{(S_k)^2 I_k} \\ \frac{C_{13}}{C}\sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - \frac{C_{23}}{C}\sum \frac{N_k g_k}{S_k (I_k)^2} + \frac{C_{33}}{C}\sum \frac{N_k g_k}{(S_k)^2 I_k} \end{pmatrix}, \tag{3.6}$$

where

$$X^{1'} = \begin{pmatrix} u^1{}_{(1)} & u^2{}_{(1)} & u^3{}_{(1)} \\ u^1{}_{(2)} & u^2{}_{(2)} & u^3{}_{(2)} \\ \vdots & \vdots & \vdots \\ u^1{}_{(n)} & u^2{}_{(n)} & u^3{}_{(n)} \end{pmatrix}, Y^{1'} = \begin{pmatrix} y^{1'}{}_{(1)} \\ y^{1'}{}_{(2)} \\ \vdots \\ y^{1'}{}_{(n)} \end{pmatrix} = \begin{pmatrix} y^1{}_{(1)} + \sqrt{\Delta t}\hat{\beta} \\ y^1{}_{(2)} + \sqrt{\Delta t}\hat{\beta} \\ \vdots \\ y^1{}_{(n)} + \sqrt{\Delta t}\hat{\beta} \end{pmatrix}$$

and

$$C = \left|(X^{1'})^T (X^{1'})\right|$$

$$= \begin{vmatrix} \sum (u^1{}_{(k)})^2 & \sum u^1{}_{(k)} u^2{}_{(k)} & \sum u^1{}_{(k)} u^3{}_{(k)} \\ \sum u^2{}_{(k)} u^1{}_{(k)} & \sum (u^2{}_{(k)})^2 & \sum u^2{}_{(k)} u^3{}_{(k)} \\ \sum u^3{}_{(k)} u^1{}_{(k)} & \sum u^3{}_{(k)} u^2{}_{(k)} & \sum (u^3{}_{(k)})^2 \end{vmatrix}$$

$$= \begin{vmatrix} \Delta t \sum \left(\frac{(N_k)^2 h_k}{S_k I_k} \right)^2 & -\Delta t \sum \frac{(N_k)^3 h_k}{S_k (I_k)^2} & \Delta t \sum \frac{(N_k)^3 h_k}{(S_k)^2 I_k} \\ -\Delta t \sum \frac{(N_k)^3 h_k}{S_k (I_k)^2} & \Delta t \sum \left(\frac{N_k}{I_k} \right)^2 & -\Delta t \sum \frac{(N_k)^2}{S_k I_k} \\ \Delta t \sum \frac{(N_k)^3 h_k}{(S_k)^2 I_k} & -\Delta t \sum \frac{(N_k)^2}{S_k I_k} & \Delta t \sum \left(\frac{N_k}{S_k} \right)^2 \end{vmatrix}.$$

The point estimate for parameters r, μ, δ, η and β in the model are

$$\hat{\beta} = \frac{B_{11}}{B} \sum \frac{b_k N_k}{S_k I_k} - \frac{B_{21}}{B} \sum \frac{(N_k)^2 b_k}{(S_k)^2 I_k},$$

$$\hat{r} = \frac{C_{11}}{C} \sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - \frac{C_{21}}{C} \sum \frac{N_k g_k}{S_k (I_k)^2} + \frac{C_{31}}{C} \sum \frac{N_k g_k}{(S_k)^2 I_k},$$

$$\hat{\mu} = \frac{C_{12}}{C} \sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - \frac{C_{22}}{C} \sum \frac{N_k g_k}{S_k (I_k)^2} + \frac{C_{32}}{C} \sum \frac{N_k g_k}{(S_k)^2 I_k},$$

$$\hat{\delta} = \frac{C_{13}}{C} \sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - \frac{C_{23}}{C} \sum \frac{N_k g_k}{S_k (I_k)^2} + \frac{C_{33}}{C} \sum \frac{N_k g_k}{(S_k)^2 I_k},$$

$$\hat{\eta} = \frac{B_{12}}{B} \sum \frac{b_k N_k}{S_k I_k} - \frac{B_{22}}{B} \sum \frac{(N_k)^2 b_k}{(S_k)^2 I_k}$$

where $\{B_{ij}, i, j = 1, 2, 3, 4\}$ are the cofactors of the determinant B, $\{C_{ij}, i, j = 1, 2, 3, 4\}$ are the cofactors of the determinant C, $a_k = S_{k+1} - S_k$, $b_k = I_{k+1} - I_k$, $h_k = 1 - \frac{N_k}{K}$, $g_k = a_k N_k + \hat{\beta} S_k I_k \Delta t$, Σ represent $\sum_{k=0}^{n-1}$.

3.1.3. Variance of estimated parameters

Interval estimation allows the range of values of the estimate to be judged with some degree of probability, thus recognising the degree of clustering and dispersion of the sample series. It also provides more information than simple point estimation. In order to obtain confidence intervals for the parameters, the variance of parameters $\theta^{1'}$ and θ^2 needs to be estimated, and the variance estimation formula is

$$Var(\hat{\theta}^{1'}) = \left[(X^{1'})^T X^{1'} \right]^{-1} \sigma^2, Var(\hat{\theta}^2) = \left[(X^2)^T X^2 \right]^{-1} \sigma^2. \tag{3.7}$$

The parameter is estimated from the residual mean square formula

$$\hat{\sigma}^2 = \frac{(Y - X\hat{\theta})^T (Y - X\hat{\theta})}{n - p}, \tag{3.8}$$

where p is the number of parameters being estimated. Both formulas in (3.2) can estimate σ^2, by means of numerical simulation, we find that there is no obvious difference in the results of their estimates of σ^2. Then we will estimate σ^2 by the second formula in (3.2). From equation (3.4),(3.8) we can obtain

$$\hat{\sigma}^2 = \frac{(Y^2)^T Y^2 - (Y^2)^T X^2 \hat{\theta}^2}{n-2},\ \hat{\theta}^2 = ((X^2)^T X^2)^{-1}((X^2)^T Y^2). \tag{3.9}$$

Then (3.9) can be written as

$$\hat{\sigma}^2 = \frac{1}{n-2}(\sum (y^2{}_k)^2 - (\sqrt{\Delta t}\sum y^2{}_k)\hat{\beta} - (\sum y^2{}_k v_k)\hat{\eta}). \tag{3.10}$$

Theorem 3.1. *The estimator $\hat{\sigma}^2$ in* (3.2) *is an asymptotically unbiased estimator for σ^2, i.e., when $n \to \infty$*

$$\hat{\sigma}^2 \to \sigma^2.$$

The proof follows the same argument as in [22].

Using $\hat{\sigma}^2$ to estimate σ^2, obtain the variance of the parameter being estimated

$$Var(\hat{\theta}^{1'}) = \begin{pmatrix} \frac{C_{11}}{C} & \frac{C_{21}}{C} & \frac{C_{31}}{C} \\ \frac{C_{12}}{C} & \frac{C_{22}}{C} & \frac{C_{32}}{C} \\ \frac{C_{13}}{C} & \frac{C_{23}}{C} & \frac{C_{33}}{C} \end{pmatrix} \hat{\sigma}^2, Var(\hat{\theta}^2) = \begin{pmatrix} \frac{B_{11}}{B} & \frac{B_{21}}{B} \\ \frac{B_{12}}{B} & \frac{B_{22}}{B} \end{pmatrix} \hat{\sigma}^2.$$

3.1.4. Interval estimates

According to the theory of least squares regression [24], if σ^2 is known, the components of the parameter estimates $\hat{\theta}^{1'}$, $\hat{\theta}^2$ satisfy a normal distribution. When the number of observations n is sufficiently large, $\hat{\sigma}^2$ can be substituted for σ^2, so the confidence intervals (CIs) of β, r, μ, δ and ε at the $(1-\alpha)$ confidence level are

$$\hat{\beta} \pm Z_{\alpha/2}\sqrt{Var(\hat{\beta})} = \frac{1}{B}(B_{11}\sum \frac{b_k N_k}{S_k I_k} - B_{21}\sum \frac{(N_k)^2 b_k}{(S_k)^2 I_k}) \pm Z_{\alpha/2}\sqrt{\frac{B_{11}}{B}\hat{\sigma}^2},$$

$$\hat{r} \pm Z_{\alpha/2}\sqrt{Var(\hat{r})} = \frac{1}{C}(C_{11}\sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - C_{21}\sum \frac{N_k g_k}{S_k (I_k)^2} + C_{31}\sum \frac{N_k g_k}{(S_k)^2 I_k}) \pm Z_{\alpha/2}\sqrt{\frac{C_{11}}{C}\hat{\sigma}^2},$$

$$\hat{\mu} \pm Z_{\alpha/2}\sqrt{Var(\hat{\mu})} = \frac{1}{C}(C_{12}\sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - C_{22}\sum \frac{N_k g_k}{S_k (I_k)^2} + C_{32}\sum \frac{N_k g_k}{(S_k)^2 I_k}) \pm Z_{\alpha/2}\sqrt{\frac{C_{22}}{C}\hat{\sigma}^2},$$

$$\hat{\delta} \pm Z_{\alpha/2}\sqrt{Var(\hat{\delta})} = \frac{1}{C}(C_{13}\sum \frac{(N_k)^2 h_k g_k}{(S_k)^2 (I_k)^2} - C_{23}\sum \frac{N_k g_k}{S_k (I_k)^2} + C_{33}\sum \frac{N_k g_k}{(S_k)^2 I_k}) \pm Z_{\alpha/2}\sqrt{\frac{C_{33}}{C}\hat{\sigma}^2},$$

and

$$\hat{\eta} \pm Z_{\alpha/2}\sqrt{Var(\hat{\eta})} = \frac{1}{B}(B_{12}\sum \frac{b_k N_k}{S_k I_k} - B_{22}\sum \frac{(N_k)^2 b_k}{(S_k)^2 I_k}) \pm Z_{\alpha/2}\sqrt{\frac{B_{22}}{B}\hat{\sigma}^2},$$

where $Z_{\alpha/2}$ is the quantile of a standard normal distribution, $Z_{0.025} = 1.96$ is the 95% confidence interval. The asymptotic widths of the CIs for θ, which is

$$2 \times Z_{\alpha/2}\sqrt{Var(\hat{\theta})}. \tag{3.11}$$

Table 1. CIs for Example 3.2.

		Sample A ($T=20, \Delta t=0.01$)		Sample B ($T=20, \Delta t=0.0025$)		Sample C ($T=80, \Delta t=0.01$)	
	True value	CI	Width of CI	CI	Width of CI	CI	Width of CI
$\hat{\beta}$	0.5	(0.39,0.54)	0.15	(0.44,0.59)	0.15	(0.43,0.52)	0.09
$\hat{r}$	0.2	(0.18,0.22)	0.03	(0.17,0.21)	0.03	(0.20,0.20)	0.00
$\hat{\mu}$	0.1	(0.09,0.11)	0.02	(0.08,0.11)	0.02	(0.10,0.10)	0.00
$\hat{\delta}$	0.2	(0.16,0.18)	0.03	(0.19,0.22)	0.03	(0.17,0.18)	0.01
$\hat{\eta}$	0.6	(0.50,0.63)	0.13	(0.54,0.67)	0.13	(0.53,0.62)	0.09

Example 3.2. We simulate the model (1.2) using the Euler-Maruyama (EM) method with initial values of $(S(0), I(0)) = (0.8, 0.4)^T$. We assume that the parameters of system (1.2) are

$$\beta = 0.5,\ \mu = 0.1,\ K = 1,\ r = 0.2,\ \varepsilon = 0.3,\ \delta = 0.2,\ \sigma = 0.02.$$

In order to compare the effect of Δt and T on the parameter interval estimation, we first set up $\Delta t = 0.01, n = 2000, T = 20$ as the control group, and then carry out experiments using the following two methods. Method 1: Reduce Δt by a factor of four, keeping T unchanged. Method 2: Increase T by a factor of four, keeping n and Δt unchanged.

We set $T = 20$ for the first 2 datasets and $T = 80$ for the third dataset in order to obtain sample data. Set $\Delta t = 0.001, n = 20000, T = 20$ for the numerical simulation and sample the simulated data at 10 data point intervals to obtain a control group (Sample A), so $\Delta t = 0.01$ and $n = 2000$ in this case. Using the same sampling method as Sample A, we generate Sample B with $\Delta t = 0.0025$ and $n = 8000$, as well as Sample C with $\Delta t = 0.01$ and $n = 8000$.

We can see from Table 1 that Method 1 (Sample B) reduces Δt, making the observations more intensive, but does not narrow the CIs for the parameters. Method 2 (Sample C) uses a sample of longer observations at one site and gives a narrower CIs. Thus, from this example, we can conclude that the length of the parameter confidence interval in the model is significantly influenced by the total time length T, but it is independent of the time interval Δt.

Example 3.3. We simulate the model (1.2) using the Euler-Maruyama (EM) method with an initial value of $(S(0), I(0)) = (0.8, 0.4)^T$. We assume that the parameters of system (1.2) are given by

$$\beta = 0.5, \mu = 0.1, K = 1, r = 0.2, \varepsilon = 0.3, \delta = 0.2, \sigma = 0.02.$$

In order to observe the effect of the different values of T on the 95% confidence interval of the parameters, we set $\Delta t = 0.01$. We consider interval lengths T of $T = 20,\ 25,\ 30,\ 35,\ 40$ and 45. For each T value, we will use the same method as in Example 3.1 to simulate and sample the corresponding data set. To observe the effect of Δt on the length of the confidence interval, we set $T = 20$. We now vary the value of the step size Δt to $\Delta t = 0.001,\ 0.002,\ 0.004,\ 0.005,\ 0.008$ and 0.01.

As depicted in Figure a, it is evident that with an increase in T, there is a gradual decrease in the length of the 95% confidence interval, displaying a distinct declining trend. However, as shown in Figure b, there is no discernible trend in the length of the 95% confidence interval as Δt varies. This is consistent with the result obtained by Gray [22], which we verify here by numerical simulations.

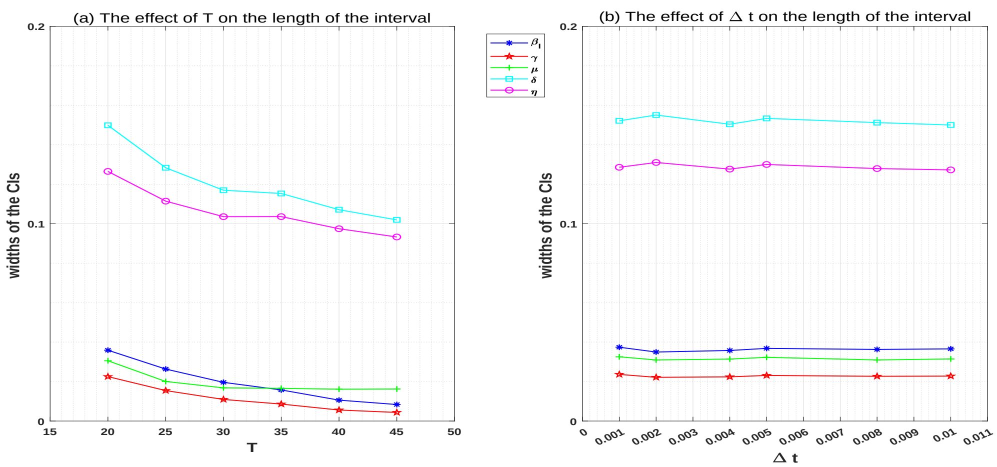

Figure 4. The parameter values in Example 3.3 generate samples for parameter estimation, a shows the parameters 95% confidence region length change over time length T. b shows the change of the parameters 95% confidence region length with step size Δt.

3.2. Pseudo-maximum likelihood estimation

In our model, it is not possible to obtain an explicit expression for the maximum likelihood function of r, β, μ, δ, ε and σ^2, mainly because it is difficult to find the corresponding likelihood function. We are unable to employ exact maximum likelihood methods in this case. Therefore, an approximation method is needed to obtain estimates for these parameters. A common approach is the pseudo-likelihood method, which can be used to estimate r, β, μ, δ, ε and σ^2. However, it is important to note that this is still an approximation method and the estimation results may not be as accurate as those obtained from the exact maximum likelihood method.

The Euler scheme discretizes the process as (3.1), where Δt is the observation interval time and $\Delta t > 0$, k is the kth observation moment. For a given previous moment state $F_k = (S_k, I_k)^T$, there are

$$E(S_{k+1}|F_k) = S_k + (rN_k(1-\frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t,$$
$$E(I_{k+1}|F_k) = I_k + (\frac{\beta S_k I_k}{N_k} - (\mu+\delta+\varepsilon)I_k)\Delta t.$$

If two random variables X and Y obey a normal distribution, then their joint probability distribution is a two-dimensional normal distribution, the probability density function is

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp[-\frac{1}{2(1-\rho^2)}(\frac{(x-\mu_x)^2}{{\sigma_x}^2} + \frac{(y-\mu_y)^2}{{\sigma_y}^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y})]$$

of which $\rho = \frac{E(XY)-E(X)E(Y)}{\sigma_x\sigma_y}$. In the SIS (Susceptible-Infectious-Susceptible) model of infectious diseases, there is a correlation between the variables S (susceptible) and I (infectious). So (S_{k+1}, I_{k+1}) has a two-dimensional normal distribution with a conditional probability density function of

$$f((S_{k+1}, I_{k+1})\,|F_k)$$
$$= \frac{1}{2\pi\Delta t\frac{\sigma^2 {S_k}^2 {I_k}^2}{{N_k}^2}\sqrt{1-\rho^2}}\exp[-\frac{{N_k}^2}{2(1-\rho^2)\sigma^2{S_k}^2{I_k}^2\Delta t}$$

$$
\begin{aligned}
&\times (S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)^2 \\
&+ (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)^2 \\
&- 2\rho(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t) \\
&\times (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t))],
\end{aligned} \tag{3.12}
$$

where $\rho = \frac{E(S_{k+1}I_{k+1}) - E(S_{k+1})E(I_{k+1})}{\sigma_{S_{k+1}}\sigma_{I_{k+1}}}$. Therefore, for a given F_0, the joint conditional density function for $\{(S_1, I_1), (S_2, I_2), \cdots, (S_n, I_n)\}$ is

$$
\begin{aligned}
&f(((S_1, I_1), (S_2, I_2), \cdots, (S_n, I_n)) \,|F_0\,) \\
=&(\frac{1}{2\pi\Delta t\sqrt{1-\rho^2\sigma^2}})^n \prod_{k=0}^{n} \frac{{N_k}^2}{{S_k}^2 {I_k}^2} \exp[-\frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t} \\
&\times (S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)^2 \\
&+ (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)^2 \\
&- 2\rho(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t) \\
&\times (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t))].
\end{aligned} \tag{3.13}
$$

Since the constants have no effect on the parameter estimates, the constants are ignored and the natural logarithm of the joint conditional probability function is taken to obtain the log pseudo-likelihood function as

$$
\begin{aligned}
L(\theta) =& - n \ln \sigma^2 + \sum_{k=0}^{n} (2\ln N_k - 2\ln S_k - 2\ln I_k) - \sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t} \\
&\times [(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)^2 \\
&+ (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)^2 \\
&- 2\rho(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t) \\
&\times (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t))].
\end{aligned} \tag{3.14}
$$

Taking the partial derivative of the parameter $\theta = [r,\ \beta,\ \mu,\ \delta,\ \varepsilon,\ \sigma^2]$ we get

$$
\begin{aligned}
\frac{\partial L(\theta)}{\partial r} =& - \sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t} [-2N_k(1 - \frac{N_k}{K})\Delta t(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) \\
&- \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t) + 2\rho N_k(1 - \frac{N_k}{K})\Delta t(I_{k+1} - I_k
\end{aligned}
$$

$$- (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)], \tag{3.15}$$

$$\frac{\partial L(\theta)}{\partial \beta} = -\sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t}[2\frac{S_k I_k \Delta t}{N_k}(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k}$$
$$- \mu S_k + \delta I_k)\Delta t) - 2\frac{S_k I_k \Delta t}{N_k}(I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)$$
$$+ 2\rho \frac{S_k I_k \Delta t}{N_k}(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)$$
$$- 2\rho \frac{S_k I_k \Delta t}{N_k}(I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)], \tag{3.16}$$

$$\frac{\partial L(\theta)}{\partial \mu} = -\sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t}[2S_k \Delta t(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k$$
$$+ \delta I_k)\Delta t) + 2I_k \Delta t(I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t) - 2\rho S_k \Delta t(I_{k+1}$$
$$- I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t) - 2\rho I_k \Delta t(S_{k+1} - S_k$$
$$- (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)], \tag{3.17}$$

$$\frac{\partial L(\theta)}{\partial \delta} = -\sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t}[-2I_k \Delta t(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k$$
$$+ \delta I_k)\Delta t) + 2I_k \Delta t(I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t) + 2\rho I_k \Delta t(I_{k+1} - I_k$$
$$- (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t) - 2\rho I_k \Delta t(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K})$$
$$- \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t), \tag{3.18}$$

$$\frac{\partial L(\theta)}{\partial \varepsilon} = -\sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t}[2I_k \Delta t(I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)$$
$$- 2\rho I_k \Delta t(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)], \tag{3.19}$$

$$\frac{\partial L(\theta)}{\partial \sigma^2} = -\frac{n}{\sigma^2} + \frac{1}{(\sigma^2)^2}\sum_{k=0}^{n-1} \frac{{N_k}^2}{2(1-\rho^2)\sigma^2 {S_k}^2 {I_k}^2 \Delta t}[(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K})$$
$$- \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)^2 + (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)^2$$
$$- 2\rho(S_{k+1} - S_k - (rN_k(1 - \frac{N_k}{K}) - \frac{\beta S_k I_k}{N_k} - \mu S_k + \delta I_k)\Delta t)$$
$$\times (I_{k+1} - I_k - (\frac{\beta S_k I_k}{N_k} - (\mu + \delta + \varepsilon) I_k)\Delta t)]. \tag{3.20}$$

Let the partial derivative of the parameters be zero, which is equivalent to solving the system of equations

Table 2. Pseudo-maximum likelihood estimation.

	$\hat{\beta}$	$\hat{r}$	$\hat{\mu}$	$\hat{\delta}$	$\hat{\varepsilon}$
Sample A $(T = 20, \Delta t = 0.01)$	0.495	0.234	0.123	0.226	0.242
Sample B $(T = 20, \Delta t = 0.0025)$	0.699	0.260	0.353	0.353	0.277
Sample C $(T = 80, \Delta t = 0.01)$	0.494	0.198	0.202	0.202	0.292

Table 3. Pseudo-maximum likelihood estimation of absolute error.

	$\lvert\hat{\beta} - \beta\rvert$	$\lvert\hat{r} - r\rvert$	$\lvert\hat{\mu} - \mu\rvert$	$\lvert\hat{\delta} - \delta\rvert$	$\lvert\hat{\varepsilon} - \varepsilon\rvert$
Sample A $(T = 20, \Delta t = 0.01)$	0.005	0.034	0.023	0.026	0.058
Sample B $(T = 20, \Delta t = 0.0025)$	0.199	0.060	0.033	0.153	0.023
Sample C $(T = 80, \Delta t = 0.01)$	0.006	0.002	0.001	0.002	0.008

$$\sum_{k=0}^{n-1} \frac{{N_k}^3(1-\frac{N_k}{K})}{{S_k}^2{I_k}^2}[\rho(I_{k+1}-I_k)-(S_{k+1}-S_k)+rN_k(1-\frac{N_k}{K})\Delta t-\beta(1+\rho)\frac{S_kI_k}{N_k}\Delta t$$
$$+\mu(\rho I_k-S_k)\Delta t+\delta(1+\rho)I_k\Delta t+\varepsilon\rho I_k\Delta t]=0,$$
$$\sum_{k=0}^{n-1} \frac{N_k}{S_kI_k}[(S_{k+1}-S_k)-(I_{k+1}-I_k)-rN_k(1-\frac{N_k}{K})\Delta t+2\beta\frac{S_kI_k}{N_k}\Delta t+\mu(S_k-I_k)\Delta t$$
$$-2\delta I_k\Delta t-\varepsilon I_k\Delta t]=0,$$
$$\sum_{k=0}^{n-1} \frac{{N_k}^2}{{S_k}^2{I_k}^2}[(S_k-\rho I_k)(S_{k+1}-S_k)+(I_k-\rho S_k)(I_{k+1}-I_k)-r(S_k-\rho I_k)N_k(1-\frac{N_k}{K})\Delta t$$
$$+\beta((S_k-\rho I_k)-(I_k-\rho S_k))\frac{S_kI_k}{N_k}\Delta t+\mu((S_k-\rho I_k)S_k+(I_k-\rho S_k)I_k)\Delta t$$
$$+\delta((I_k-\rho S_k)-(S_k-\rho I_k))I_k\Delta t+\varepsilon(I_k-\rho S_k)I_k\Delta t]=0,$$
$$\sum_{k=0}^{n-1} \frac{{N_k}^2}{{S_k}^2I_k}[(I_{k+1}-I_k)-(S_{k+1}-S_k)+rN_k(1-\frac{N_k}{K})\Delta t-2\beta\frac{S_kI_k}{N_k}\Delta t+\mu(I_k-S_k)\Delta t$$
$$+2\delta I_k\Delta t+\varepsilon I_k\Delta t]=0,$$
$$\sum_{k=0}^{n-1} \frac{{N_k}^2}{{S_k}^2I_k}[(I_{k+1}-I_k)-\rho(S_{k+1}-S_k)+r\rho N_k(1-\frac{N_k}{K})\Delta t-\beta(1+\rho)\frac{S_kI_k}{N_k}\Delta t+\mu(I_k$$
$$-\rho S_k)\Delta t+\delta(1+\rho)I_k\Delta t+\varepsilon I_k\Delta t]=0.$$

The resulting parameter estimates $\hat{r}$, $\hat{\beta}$, $\hat{\mu}$, $\hat{\delta}$ and $\hat{\varepsilon}$ are put into (3.20) to solve for σ^2.

Example 3.4. Using the same sample data, A, B, and C, as in Example 3.2, we can perform pseudo-maximum likelihood estimation of the parameters and calculate the absolute error between the estimated values and the true values.

The numerical results in Table 2 and 3 show that the estimation performance for parameters β and δ is poor when $\Delta t = 0.0025$. In addition, increasing the time length T can improve the maximum likelihood estimation, while decreasing the step size Δt can actually worsen the estimation.

3.3. Bayesian posterior distribution mean estimation

In this section, the Bayesian estimation method is employed within the framework of linear models. The assumption of a normal prior distribution for the parameters is made. Under this assumption, the posterior distribution remains normal as well. The Bayesian estimation based on the quadratic loss function is determined by the posterior mean [5].

Consider the following multivariate linear model (3.2), where $Y \in R_{2n\times 1}$ is the observed, $X \in R_{2n\times p}$ is the known design matrix, p is the number of parameters, and ε is a $2n * 1$ matrix of random errors assumed to have a matrix normal distribution $\varepsilon \sim N(0, \sigma^2 I)$, where I is the identity matrix and $\theta \in R_{p\times 1}$ is the unknown parameter matrix. The likelihood function of Bayesian inference under the regression model can be expressed as

$$\begin{aligned} L(Y\left|X,\theta,\sigma^2\right.) &= (2\pi\sigma^2)^{-2n/2}\exp\left[-\frac{(Y-X\theta)^T(Y-X\theta)}{2\sigma^2}\right] \\ &\propto (\sigma^2)^{-2n/2}\exp\left[-\frac{(Y-X\theta)^T(Y-X\theta)}{2\sigma^2}\right] \\ &\propto (\sigma^2)^{-2n/2}\exp\left[-\frac{S+(\theta-\hat{\theta})X^TX(\theta-\hat{\theta})}{2\sigma^2}\right], \end{aligned} \tag{3.21}$$

where $\hat{\theta} = (X^TX)^{-1}(X^TY),\ S = (Y - X\hat{\theta})^T(Y - X\hat{\theta})$. Assuming that the joint prior distribution of parameters θ and σ^2 is the normal-inverse Gamma distribution $p(\theta, \sigma^2) = NIG(m, \sigma^2 V, a, b)$, i.e.

$$p(\theta,\sigma^2) = p(\theta\left|\sigma^2\right.)p(\sigma^2), \theta\left|\sigma^2\right. \sim N(m,\sigma^2V), \sigma^2 \sim IG(a,b). \tag{3.22}$$

Therefore

$$p(\theta,\sigma^2) \propto (\sigma^2)^{-(\frac{l}{2}+a+1)}\exp\left[-\frac{(\theta-m)^TV^{-1}(\theta-m)+2b}{2\sigma^2}\right], \tag{3.23}$$

where $l \triangleq p+1$. According to the Bayes' theorem, the posterior distribution of parameter (θ, σ^2) satisfies:

$$p(\theta,\sigma^2\left|Y\right.) = \frac{L(Y\left|\theta,\sigma^2\right.)p(\theta,\sigma^2)}{p(Y)}, \tag{3.24}$$

where $p(Y)$ is a constant. The posterior distribution of the parameter can be obtained by multiplying the prior distribution (3.23) with the likelihood function, which is given by

$$\begin{aligned} p(\theta,\sigma^2\left|Y\right.) \propto &(\sigma^2)^{-(\frac{l}{2}+a+1+\frac{2n}{2})}\exp\left[-\frac{S+2b}{2\sigma^2}\right] \\ &\times\exp\left[-\frac{(\theta-\hat{\theta})X^TX(\theta-\hat{\theta})+(\theta-m)^TV^{-1}(\theta-m)}{2\sigma^2}\right] \\ \propto&(\sigma^2)^{-\frac{l}{2}}\exp\left[-\frac{(\theta-m^*)^T(V^*)^{-1}(\theta-m^*)}{2\sigma^2}\right]\times(\sigma^2)^{-(a^*+1)}\exp\left[-\frac{b^*}{\sigma^2}\right] \\ \propto&N(m^*,\sigma^2V^*)IG(a^*,b^*), \end{aligned} \tag{3.25}$$

where

$$m^* = (V^{-1}+X^TX)^{-1}(V^{-1}m+X^TY),$$

$$V^* = (V^{-1} + X^T X)^{-1},$$
$$a^* = a + \frac{2n}{2},$$
$$b^* = b + \left[m^T V^{-1} m + Y^T Y - (m^*)(V^*)^{-1} m^* \right] /2.$$

As shown in equation (3.25), the posterior distribution of parameter vector θ follows a normal distribution, while the posterior distribution of parameter σ^2 follows an inverse Gamma distribution,i.e.

$$\theta \left| Y \right. \sim N(m^*, \sigma^2 V^*), \tag{3.26}$$
$$\sigma^2 \left| Y \right. \sim IG(a^*, b^*). \tag{3.27}$$

When we take the prior distribution parameter $m = \mathrm{O}_{p\times 1}, V = \mathrm{O}_{p\times p}, a = 0, b = 0$ is a special case of Bayesian estimation. From the formula for the posterior conditional distribution of the parameters of the linear model given by Zellner [34], the posterior marginal distribution of the parameters is then given as

$$\theta \left| Y \right. \sim N(\hat{\theta}, \sigma^2 (X^T X)^{-1}), \tag{3.28}$$
$$\sigma^2 \left| Y \right. \sim IG(\frac{n}{2}, \frac{1}{2}(Y - X\theta)^T (Y - X\theta)), \tag{3.29}$$

the mean of the posterior distribution of the parameters has the same expression as the least squares estimator.

Based on the above analysis the sample sequence of parameters can be obtained using the Gibbs sample algorithm, which proceeds as follows:

1. Given a vector of initial values $\theta_{(0)}$ and an initial value ${\sigma^2}_{(0)}$.
2. Given an initial time $t = 0$ and a sufficiently large positive integer M.
3. Loop: While $t = 1, 2, \cdots, M$
   (1) Draw $\theta_{(t+1)}$ from the distribution of $p(\theta \left| {\sigma^2}_{(t)}, Y \right.)$.
   (2) Draw ${\sigma^2}_{(t+1)}$ from the distribution of $p(\sigma^2 \left| \theta_{(t)}, Y \right.)$.
   (3) $t = t + 1$.
   (4) Store the parameter samples of $\theta_{(t+1)}$ and ${\sigma^2}_{(t+1)}$ at each step.
4. End the loop when $t = M$.

Assuming that the number of steps in the generated Markov chain is M and the obtained sample sequence is $\theta = (\theta_{(1)}, \theta_{(2)}, \cdots, \theta_{(M)}), \sigma^2 = ({\sigma^2}_{(1)}, {\sigma^2}_{(2)}, \cdots, {\sigma^2}_{(M)})$. We discard the first 10% of the samples as burn-in to ensure that the Markov chain is smooth, and find its average value as the estimated value of the model parameters, i.e.:

$$\hat{\theta} = \sum_{i=\frac{M}{10}+1}^{M} \theta_{(i)},\ \hat{\sigma}^2 = \sum_{i=\frac{M}{10}+1}^{M} {\sigma^2}_{(i)}.$$

The Bayesian posterior mean coincides with the least squares estimator under the adopted prior specification, using Y^2 and X^2 as variables for estimating parameter β is more effective. Therefore, in the Bayesian estimation, we first use second equation of (3.2) to estimate the parameter β, and then we substitute the estimated value of $\hat{\beta}$ back into first equation of (3.2) to estimate the parameter $\hat{\theta}^{1'} = (r,\ \mu,\ \delta)^T$.

Table 4. Bayesian posterior mean estimation results.

	$\hat{\beta}$	$\hat{r}$	$\hat{\mu}$	$\hat{\delta}$	$\hat{\eta}$
Sample A $(T = 20, \Delta t = 0.01)$	0.468	0.201	0.101	0.171	0.568
Sample B $(T = 20, \Delta t = 0.0025)$	0.515	0.193	0.095	0.204	0.609
Sample C $(T = 80, \Delta t = 0.01)$	0.478	0.200	0.100	0.175	0.575

Table 5. Bayesian posterior mean estimation of absolute error.

	$\|\hat{\beta} - \beta\|$	$\|\hat{r} - r\|$	$\|\hat{\mu} - \mu\|$	$\|\hat{\delta} - \delta\|$	$\|\hat{\eta} - \eta\|$
Sample A $(T = 20, \Delta t = 0.01)$	0.032	0.001	0.001	0.029	0.032
Sample B $(T = 20, \Delta t = 0.0025)$	0.015	0.008	0.005	0.004	0.009
Sample C $(T = 80, \Delta t = 0.01)$	0.023	0.000	0.000	0.025	0.025

Example 3.5. Using the same sample data A, B, and C as in Example 3.2, estimate the parameters using Bayesian posterior distribution means and calculate the absolute error between the parameter estimates and the true values.

The numerical simulation results (Tables 4 and 5) show that the estimation accuracy for β, δ and η improves as the sampling interval Δt becomes smaller, whereas the estimation accuracy for r and μ does not improve significantly. In fact, when Δt is very small, the estimation accuracy for r and μ may deteriorate. In particular, increasing the observation time horizon T improves the estimation performance for all parameters.

Example 3.6. The analytical exploration of the numerical simulation and parameter estimation results in Examples 3.2, 3.4 and 3.5 shows that when given the same parameters r, μ, δ, β, ε and σ, the generation of simulated data at different time intervals and time lengths T is likely to have an impact on the estimation effect, and therefore parameter estimation is performed on simulated data at different time intervals and time lengths, respectively, and the estimation results are analysed. First assume that the parameters are

$$\beta = 0.5, \ \mu = 0.1, \ K = 1, \ r = 0.2, \ \varepsilon = 0.3, \ \delta = 0.2, \ \sigma = 0.02,$$

and the initial value is $(S(0), I(0)) = (0.8, 0.4)^T$.T $= 20$ and the time interval Δt is taken as 0.1, 0.05 and 0.01, respectively, to analyse the effect of the time interval Δt on the estimation results, which are shown in Table 6.

The results of parameter estimation are shown in Table 6. Firstly, it can be observed that both the least squares method and the Bayesian posterior mean estimation method perform well in estimating the parameters. On the other hand, the pseudo-maximum likelihood method is not very effective in estimating the parameters μ and β. Overall, there is no significant difference in the estimation performance between the least squares method and the Bayesian posterior mean estimation method, and both methods outperform the pseudo-maximum likelihood parameter estimation method.

Secondly, it can be seen that as the time interval Δt becomes smaller and smaller, the estimation of most parameters basically does not improve, but the estimation accuracy deteriorates when Δt becomes excessively small. This may be because the estimation procedure becomes more sensitive to numerical errors and discretization effects in the approximate model when Δt

Table 6. Estimation results at different time intervals Δt.

	Δt	Least squares estimation	Pseudo-maximum likelihood estimation	Bayesian posterior distribution mean estimation
$\lvert\hat{r}-r\rvert$	$\Delta t = 0.1$	0.0010	0.0199	0.0010
	$\Delta t = 0.05$	0.0077	0.0238	0.0077
	$\Delta t = 0.01$	0.0117	0.0451	0.0115
$\lvert\hat{\beta}-\beta\rvert$	$\Delta t = 0.1$	0.0284	0.0869	0.0282
	$\Delta t = 0.05$	0.0064	0.0430	0.0066
	$\Delta t = 0.01$	0.0270	0.2901	0.0272
$\lvert\hat{\mu}-\mu\rvert$	$\Delta t = 0.1$	0.0327	0.0836	0.0327
	$\Delta t = 0.05$	0.0158	0.0513	0.0161
	$\Delta t = 0.01$	0.0189	0.2214	0.0192
$\lvert\hat{\delta}-\delta\rvert$	$\Delta t = 0.1$	0.0010	0.0118	0.0011
	$\Delta t = 0.05$	0.0052	0.0147	0.0052
	$\Delta t = 0.01$	0.0077	0.0220	0.0075
$\lvert\hat{\varepsilon}-\varepsilon\rvert$	$\Delta t = 0.1$	0.0102	0.0226	0.0104
	$\Delta t = 0.05$	0.0157	0.0300	0.0158
	$\Delta t = 0.01$	0.0184	0.0086	0.0181

is too small. In order to verify the above conclusions, several simulations are carried out to obtain consistent results.

Example 3.7. In view of the above analysis, since the parameters are already well estimated at $\Delta t = 0.01$, the parameters are set to be

$$\beta = 0.5,\ \mu = 0.1,\ K = 1,\ r = 0.2,\ \varepsilon = 0.3,\ \delta = 0.2,\ \sigma = 0.02,$$

and the initial value is $(S(0), I(0)) = (0.8, 0.4)^T$. The time length T was taken as 20, 40 and 80 and the effect of time length T on the estimation results was analysed and the estimation results are shown in Table.

Table 7 indicates that as T increases, the estimation effectiveness of the parameters gradually improves, indicating that T is the primary factor influencing the estimation effectiveness. When T is sufficiently large, there is no significant disparity in the estimation effectiveness among different estimation methods.

4. Conclusion

In this paper, we studied a stochastic SIS model with logistic population input and degenerate diffusion, where environmental fluctuations are introduced through perturbation of the contact rate. We established the existence and uniqueness of the global positive solution and obtained a threshold theorem that characterizes disease extinction and ergodicity of the stationary distribution. These results show that the stochastic threshold determines the long-term behavior of the epidemic under environmental noise.

Table 7. Estimation results for different lengths of time T.

	T	Least squares estimation	Pseudo-maximum likelihood estimation	Bayesian posterior distribution mean estimation
$\lvert\hat{r}-r\rvert$	$T=20$	0.0078	0.0064	0.0080
	$T=40$	0.0043	0.0031	0.0043
	$T=80$	0.0027	0.0003	0.0027
$\left\lvert\hat{\beta}-\beta\right\rvert$	$T=20$	0.0308	0.0402	0.0312
	$T=40$	0.0410	0.0481	0.0411
	$T=80$	0.0117	0.0331	0.0117
$\lvert\hat{\mu}-\mu\rvert$	$T=20$	0.0285	0.0247	0.0289
	$T=40$	0.0407	0.0459	0.0407
	$T=80$	0.0134	0.0289	0.0135
$\left\lvert\hat{\delta}-\delta\right\rvert$	$T=20$	0.0048	0.0026	0.0049
	$T=40$	0.0023	0.0017	0.0023
	$T=80$	0.0014	0.0002	0.0014
$\lvert\hat{\varepsilon}-\varepsilon\rvert$	$T=20$	0.0037	0.0026	0.0038
	$T=40$	0.0030	0.0026	0.0030
	$T=80$	0.0039	0.0012	0.0039

We also studied parameter estimation for the proposed model under discrete observations by comparing least squares estimation, pseudo-MLE, and Bayesian posterior mean estimation. The numerical results indicate that the total observation time plays a major role in estimation accuracy, while an excessively small observation step size does not necessarily improve finite-sample performance. Overall, the estimation part complements the dynamical analysis by providing a comparative study of parameter identification for the same stochastic SIS model.

Future work may consider more general stochastic perturbations, higher-dimensional epidemic systems, and data-driven identification of stochastic epidemic models from observed sample paths.

Funding. The work was supported by Science and Technology Development Project of Jilin Province (YDZJ202401359ZYTS).

Data availability. Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

Conflict of interest. The authors declare that they have no conflict of interest.

References

[1] M. A. Achterberg, B. Prasse and P. Van Mieghem, *Analysis of continuous-time Markovian ε-SIS epidemics on networks*, Phys. Rev. E, 2022, 105(5), 054305.

[2] R. Anderson and R. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.

[3] D. R. Bell, *The Malliavin Calculus*, Courier Corporation, New York, 2012.

[4] J. P. N. Bishwal, *Parameter Estimation in Stochastic Differential Equations*, Springer, Berlin, 2007.

[5] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, John Wiley & Sons, New York, 2011.

[6] F. Brauer, *Mathematical approaches to problems in resource management and epidemiology*, Proceedings of a Conference Held at Ithaca, 1987, 109–123.

[7] J. Cao and L. Hu, *Asymptotic properties of a pseudo-MLE for CIR model*, in: Proc. 2010 Int. Conf. IIGSS-CPS, Nanjing, China, 2010, 28–31.

[8] L. Q. Gao and H. W. Hethcote, *Disease transmission models with density-dependent demographics*, J. Math. Biol., 1992, 30(7), 717–731.

[9] A. Gray, D. Greenhalgh, L. Hu, et al., *A stochastic differential equation SIS epidemic model*, SIAM J. Appl. Math., 2011, 71(3), 876–902.

[10] S. He, S. Tang, Y. Cai, et al., *A stochastic epidemic model coupled with seasonal air pollution: Analysis and data fitting*, Stoch. Environ. Res. Risk Assess., 2020, 34(12), 2245–2257.

[11] S. He and Z. Xiang, *Dynamics of stochastic SIQS model based on nonlinear incidence: Disease extinction and stationary distribution under degenerate diffusion*, Electron. Res. Arch., 2025, 33(7), 4259–4283.

[12] S. M. Iacus, *Simulation and Inference for Stochastic Differential Equations: With R Examples*, Springer, New York, 2008.

[13] M. C. M. de Jong, O. Diekmann, et al., *How does transmission of infection depend on population size?*, Epidemic Models: Their Structure and Relation to Data, 1995, 84–94.

[14] R. Khasminskii, *Stochastic Stability of Differential Equations*, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands, 1980.

[15] N. R. Kristensen, H. Madsen and S. B. Jørgensen, *Parameter estimation in stochastic grey-box models*, Automatica, 2004, 40(2), 225–237.

[16] S. Li and S. Guo, *Permanence and extinction of a stochastic SIS epidemic model with three independent Brownian motions*, Discrete Contin. Dyn. Syst. B, 2021, 26(5), 2693–2719.

[17] Q. Liu, D. Jiang, T. Hayat, et al., *Threshold dynamics of a stochastic SIS epidemic model with nonlinear incidence rate*, Phys. A, 2019, 526, 120946.

[18] Q. Liu, D. Jiang, T. Hayat, et al., *A stochastic SIRS epidemic model with logistic growth and general nonlinear incidence rate*, Phys. A, 2020, 551, 124152.

[19] X. Mao, *Stochastic Differential Equations and Applications*, Elsevier, Horwood Publishing, 2007.

[20] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.

[21] H. Nemoto and Y. Shimizu, *Statistical inference for discretely sampled stochastic functional differential equations with small noise*, Stat. Inference Stoch. Process., 2024, 27(2), 427–456.

[22] J. Pan, A. Gray, D. Greenhalgh, et al., *Parameter estimation for the stochastic SIS epidemic model*, Stat. Inference Stoch. Process., 2014, 17(1), 75–98.

[23] K. Pichór and R. Rudnicki, *Stability of Markov semigroups and applications to parabolic systems*, J. Math. Anal. Appl., 1997, 215(1), 56–74.

[24] J. O. Rawlings, S. G. Pantula and D. A. Dickey, *Applied Regression Analysis: A Research Tool*, Springer, New York, 1998.

[25] R. Rudnicki, *Long-time behaviour of a stochastic prey-predator model*, Stochastic Process. Appl., 2003, 108(1), 93–107.

[26] R. Rudnicki and K. Pichór, *Influence of stochastic perturbation on prey-predator systems*, Math. Biosci., 2007, 206(1), 108–119.

[27] R. Rudnicki, K. Pichór and M. Tyran-Kamińska, *Markov semigroups and their applications*, Springer, Berlin, 2002, 215–238.

[28] J. Timmer, *Parameter estimation in nonlinear stochastic differential equations*, Chaos Solitons Fractals, 2000, 11(15), 2571–2578.

[29] H. Wang, G. Zhang, T. Chen, et al., *Threshold analysis of a stochastic SIRS epidemic model with logistic birth and nonlinear incidence*, Mathematics, 2023, 11(7), 1737.

[30] L. Wang, C. Gao, R. Rifhat, et al., *Stationary distribution and bifurcation analysis for a stochastic SIS model with nonlinear incidence and degenerate diffusion*, Chaos Solitons Fractals, 2024, 182, 114872.

[31] Z. Xu and L. Wang, *Stationary distribution and extinction of a stochastic SEI epidemic model with logistic growth and nonlinear perturbation*, AIMS Math., 2025, 10(12), 28488–28513.

[32] Z. Xu, L. Wang and D. Jiang, *The threshold of a stochastic SIR epidemic model with logistic growth*, J. Appl. Anal. Comput., 2026, 16(3), 1187–1206.

[33] T. Xue, X. Fan and Z. Chang, *Dynamics of a stochastic SIRS epidemic model with standard incidence and vaccination*, Math. Biosci. Eng., 2022, 19(10), 10618–10636.

[34] A. Zellner, *Applications of Bayesian analysis in econometrics*, J. R. Stat. Soc. Ser. D Stat., 1983, 32(1), 23–34.

[35] X. Zhang, Z. Zhang, J. Tong, et al., *Ergodicity of stochastic smoking model and parameter estimation*, Adv. Differ. Equ., 2016, 2016(1), 1–20.

[36] Y. Zhao, D. Jiang, X. Mao, et al., *The threshold of a stochastic SIRS epidemic model in a population with varying size*, Discrete Contin. Dyn. Syst. B, 2015, 20(4), 1277–1295.

[37] C. Zhou and Z. Li, *Parameter estimation of stochastic SEIR epidemic model using particle MCMC*, Chaos, 2025, 35(5), 053139.

Received January 2024; Accepted May 2026; Available online June 2026.