Journal of Mathematical Biology (2022) 85:23
https://doi.org/10.1007/s00285-022-01787-3

Mathematical Biology

An SIRS model with nonmonotone incidence and saturated treatment in a changing environment

Qin Pan[1] · Jicai Huang[1] · Hao Wang[2]

Received: 13 March 2022 / Revised: 7 June 2022 / Accepted: 10 June 2022 /
Published online: 20 August 2022

Abstract

Nonmonotone incidence and saturated treatment are incorporated into an SIRS model under constant and changing environments. The nonmonotone incidence rate describes the psychological or inhibitory effect: when the number of the infected individuals exceeds a certain level, the infection function decreases. The saturated treatment function describes the effect of infected individuals being delayed for treatment due to the limitation of medical resources. In a constant environment, the model undergoes a sequence of bifurcations including backward bifurcation, degenerate Bogdanov-Takens bifurcation of codimension 3, degenerate Hopf bifurcation as the parameters vary, and the model exhibits rich dynamics such as bistability, tristability, multiple periodic orbits, and homoclinic orbits. Moreover, we provide some sufficient conditions to guarantee the global asymptotical stability of the disease-free equilibrium or the unique positive equilibrium. Our results indicate that there exist three critical values r_1, r_2 and r_3 for the treatment rate r: **(i)** when $r \geq \max\{r_1, r_2\}$, the disease will disappear; **(ii)** when $r < \min\{r_1, r_3\}$, the disease will persist. In a changing environment, the infective population starts along the stable disease-free state (or an endemic state) and surprisingly continues tracking the unstable disease-free state (or a limit cycle) when the system crosses a bifurcation point, and eventually tends to the stable endemic state (or the stable disease-free state). This transient tracking of the unstable disease-free state when $\mathscr{R}_0 > 1$ predicts regime shifts that cause the delayed dis-

This work was partially supported by NSFC (No. 11871235) and NSERC (RGPIN-2020-03911 and RGPAS-2020-00090).

✉ Jicai Huang
hjc@mail.ccnu.edu.cn

✉ Hao Wang
hao8@ualberta.ca

[1] School of Mathematics and Statistics, Central China Normal University, Wuhan, Hubei 430079, People's Republic of China

[2] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada

ease outbreak in a changing environment. Furthermore, the disease can disappear in advance (or belatedly) if the rate of environmental change is negative and large (or small). The transient dynamics of an infectious disease heavily depend on the initial infection number and rate or the speed of environmental change.

Keywords SIRS model · Nonmonotone incidence rate · Saturated treatment rate · Backward bifurcation · Bogdanov-Takens bifurcation · Hopf bifurcation · Environmental change · Regime shifts · Transient dynamics

Mathematics Subject Classification 92D30 · 34C23

1 Introduction

Once again the COVID-19 pandemic raised the alarm for humans to prevent and control infectious diseases. A suitable epidemic model can provide helpful insights to understand, predict and control the transmission of an emerging infectious disease. In the epidemic model formulation, the incidence rate and the treatment rate play key roles in disease dynamics. We will incorporate a more realistic version of incidence and treatment into the classical susceptible-infectious-recovered-susceptible (SIRS) model with treatment:

$$\begin{aligned} \frac{dS}{dt} &= b - dS - f(I)S + \gamma R, \\ \frac{dI}{dt} &= f(I)S - (d+\mu)I - T(I), \\ \frac{dR}{dt} &= -(d+\gamma)R + \mu I + T(I), \end{aligned} \tag{1.1}$$

where the variables and parameters are listed in Table 1. Here the SIRS model assumes that new recruits are susceptible and recovered individuals have temporary immunity.

Many variations of the incidence rate were proposed in the literature. The simplest one is a bilinear incidence rate (Kermack and McKendrick 1927) $f(I)S = kIS$ (where $k > 0$ is the infection rate). Since it ignores the crowding effect of infected individuals, protection measures and awareness or inhibition, the bilinear incidence rate is not realistic enough for many infectious diseases. To describe the behavioral change and crowding effect of infected individuals, Capasso et al. (1977) and Capasso and Serio (1978) used a saturated incidence force $f(I) = kI/(1+\alpha I)$ to describe a "crowding effect" in modeling the cholera epidemics in Bari in 1973 (see Liu and Yang 2012; Yang et al. 2012; Zhou et al. 2014 for additional references). This function is a nonlinear increasing function in I but eventually tends to a saturation level k/α as I increases to infinity. Notice that both the bilinear and saturated incidence rates are monotone. However, the incidence rate exhibits nonmonotonicity in some epidemic diseases, induced by "psychological" effect (Capasso and Serio 1978): the contact rate and the infection probability usually increase when a new infectious disease emerges because people have little knowledge about the disease, however, when more and

Table 1 Definitions of variables and parameters in model (1.1)

Variables	Description
$S(t)$	The susceptible population at time t
$I(t)$	The infected population at time t
$R(t)$	The recovered population at time t
$f(I)$	The infectious force
$T(I)$	The treatment function
b	The recruitment rate of the population
d	The per capita natural death rate of the population
μ	The per capita natural recovery rate of the infective individuals
γ	The per capita rate of recovered individuals who lose immunity and return to the susceptible class

more individuals are infected by the disease, people usually tend to reduce the number of contacts. To incorporate the psychological effect, Xiao and Ruan (2007) proposed the nonmonotone incidence rate $f(I) = kI/(1 + \alpha I^2)$, which is increasing when $0 \le I \le 1/\sqrt{\alpha}$ but decreasing when $I > 1/\sqrt{\alpha}$.

Treatment is a pivotal methodology in the control of an infectious disease. Different types of treatment rate were proposed by many researchers for different epidemic scenarios. Traditional epidemic models assume the linear treatment rate $T(I) = rI$, which is reasonable given sufficient medical resources including vaccines, medicines, hospital beds, etc. However, the supply of these medical facilities is always limited once the disease spans rapidly and widely. For this reason, Wang and Ruan (2004) introduced the constant treatment function $T(I) = \begin{cases} H, I > 0 \\ 0, I = 0 \end{cases}$. Later Wang (2006) modified the constant treatment function to a piecewise-smooth treatment function $T(I) = \min\{rI, rI_0\}$, where r is a positive constant and I_0 is the infective level at which the health care system reaches its capacity. Furthermore, Zhang and Liu (2008) proposed a continuously differentiable treatment function $T(I) = \frac{rI}{1+\beta I}$, where $r > 0$ is the cure rate, $\beta \ge 0$ measures the extent of the effect of the infected population being delayed for treatment, and for small I, $T(I) \sim rI$, whereas for large I, $T(I) \sim \frac{r}{\beta}$. Moreover, when $\beta = 0$, the saturated treatment function returns to the linear one. There have been many studies (Zhang and Liu 2008; Zhang and Suo 2010; Ghosh et al. 2021; Jana et al. 2016; Zhou and Meng 2012) that used the saturated treatment function to characterize limited medical resources.

Ghosh et al. (2021) proposed an SIR model with nonmonotone incidence and saturated treatment as well as the disease-induced death rate and vaccination. They obtained a necessary and sufficient condition for backward bifurcation, and investigated saddle-node and Hopf bifurcations. Meanwhile, subcritical Hopf bifurcation was exhibited by numerical simulation. Pan et al. (2022) formulated an SIRS model by considering nonmonotone incidence and piecewise-smooth treatment, and found that the model can possess Bogdanov-Takens and subcritical Hopf bifurcations. In this paper, we replace the piecewise-smooth treatment function by a saturated treatment function as below:

$$\begin{aligned}
\frac{dS}{dt} &= b - dS - \frac{kIS}{1+\alpha I^2} + \gamma R,\\
\frac{dI}{dt} &= \frac{kIS}{1+\alpha I^2} - (d+\mu)I - \frac{rI}{1+\beta I},\\
\frac{dR}{dt} &= -(d+\gamma)R + \mu I + \frac{rI}{1+\beta I},
\end{aligned} \tag{1.2}$$

which is a specific form of model (1.1). In addition, model (1.2) extends the model in Wang (2006) that assumed a bilinear incidence rate and the model in Xiao and Ruan (2007) that assumed no treatment.

The response of populations to environmental changes has been and remains a cutting-edge area in ecological modeling (Arumugam et al. 2020, 2021; Xiang et al. 2022). Following the same logic in studying environmental changes in ecology, modeling a changing environment is equally important in epidemic research. Linking transient dynamics under a changing environment to stable and unstable states in a constant environment provides novel insights into understanding disease transmission mechanisms and outbreaks. The infection rate highly depends on nonpharmaceutical interventions, weather conditions, holidays, gatherings, and many other factors. To describe the impact of environmental changes on the infection rate, we assume that the infection rate is a linear function of time t as the simplest dependence. We formulate the following model in a changing environment:

$$\begin{aligned}
\frac{dS}{dt} &= b - dS - \frac{kIS}{1+\alpha I^2} + \gamma R,\\
\frac{dI}{dt} &= \frac{kIS}{1+\alpha I^2} - (d+\mu)I - \frac{rI}{1+\beta I},\\
\frac{dR}{dt} &= -(d+\gamma)R + \mu I + \frac{rI}{1+\beta I},\\
\frac{dk}{dt} &= u,
\end{aligned} \tag{1.3}$$

where u is the speed of environmental change. From the fourth equation of system (1.3), we have $k(t) = k_0 + ut$ (k_0 is the initial value), which represents the possible directional environmental change.

In a constant environment (i.e., $u = 0$), system (1.3) is reduced to system (1.2). We will study the dynamics of model (1.2) via rigorous bifurcation analysis. In a changing environment (i.e., $u \neq 0$), we will study how the speed of environmental change regulates the dynamics of system (1.3). Meanwhile, we will compare and contrast dynamics in the changing environment with those obtained by bifurcation analysis in the constant environment. In addition, the long-term dynamics and persistence of epidemic systems are predicted under continuous environmental changes.

The remaining paper is organized as follows. In Sect. 2, we simplify system (1.2) into a two-dimensional system based on the fact that the total population size is constant, then we discuss the stability and types of disease-free and endemic equilibria. Furthermore, we study the bifurcations and global dynamics of system (2.2). In Sect. 3,

we explore the impact of environmental change on the transient dynamics of system (1.3). We summarize the results and suggest future research directions in the last section.

2 Constant environment

Under a constant environment $u = 0$, we will study the dynamical behaviors of model (1.2) on the plane $S + I + R = \frac{b}{d}$ where the limit set of model (1.2) locates.

2.1 Model simplification

Model (1.2) assumes there is no disease-caused mortality, the total population size is hence completely determined by the vital dynamics in model (1.2). Therefore, the total population size eventually tends to a constant. This allows us to simplify system (1.2) into a two-dimensional system.

Lemma 2.1 *The plane $S + I + R = \frac{b}{d}$ is an invariant manifold of system (1.2), which is attracting in the first octant.*

Proof Summing up the three equations in (1.2) and denoting $N(t) = S(t)+I(t)+R(t)$, we have

$$\frac{dN}{dt} = b - dN.$$

It is clear that $N(t) = \frac{b}{d}$ is a solution and for any $N(t_0) \geq 0$, the general solution is

$$N(t) = \frac{1}{d}\left[b - (b - dN(t_0))e^{-d(t-t_0)}\right].$$

Thus

$$\lim_{t\to\infty} N(t) = \frac{b}{d},$$

which implies the conclusion. □

Lemma 2.1 implies that the limit set of system (1.2) is on the plane $S+I+R = \frac{b}{d}$. Thus we focus on the reduced system

$$\begin{aligned} \frac{dI}{dt} &= \frac{kI}{1+\alpha I^2}\left(\frac{b}{d} - I - R\right) - (d+\mu)I - \frac{rI}{1+\beta I}, \\ \frac{dR}{dt} &= -(d+\gamma)R + \mu I + \frac{rI}{1+\beta I}. \end{aligned} \tag{2.1}$$

We rescale system (2.1) by introducing $x = \frac{k}{d+\gamma} I,\ y = \frac{k}{d+\gamma} R,\ \tau = (d+\gamma)t$, then system (2.1) becomes (still denote τ by t)

$$\begin{aligned} \frac{dx}{dt} &= \frac{x}{1+px^2}(n-x-y) - mx - \frac{vx}{1+sx}, \\ \frac{dy}{dt} &= qx - y + \frac{vx}{1+sx}, \end{aligned} \tag{2.2}$$

where

$$\begin{aligned} &p = \alpha\left(\frac{d+\gamma}{k}\right)^2,\ n = \frac{kb}{d(d+\gamma)},\ m = \frac{d+\mu}{d+\gamma}, \\ &v = \frac{r}{\gamma+d},\ s = \frac{\beta(\gamma+d)}{k},\ q = \frac{\mu}{d+\gamma}, \end{aligned} \tag{2.3}$$

and

$$q < m < q+1,\ q,\ p,\ n,\ v,\ s > 0. \tag{2.4}$$

It is easy to see that the positive invariant and bounded region of system (2.2) is

$$D_1 = \left\{(x, y) | 0 \le x \le n,\ 0 \le y \le qn + \frac{vn}{1+sn}\right\}.$$

2.2 Equilibria and their types

To find the equilibria of system (2.2) in D_1, we set

$$\begin{aligned} \frac{x}{1+px^2}(n-x-y) - mx - \frac{vx}{1+sx} &= 0, \\ qx - y + \frac{vx}{1+sx} &= 0, \end{aligned}$$

which yield

$$\begin{aligned} &x[mpsx^3 + ((m+v)p + (1+q)s)x^2 \\ &\quad + (1+q+v+(m-n)s)x + m - n + v] = 0. \end{aligned} \tag{2.5}$$

Clearly, system (2.2) always has a disease-free equilibrium $E_0 = (0, 0)$. By the next generation matrix method in Driessche and Watmough (2002), we find the basic reproduction number

$$\mathscr{R}_0 = \frac{n}{m+v}. \tag{2.6}$$

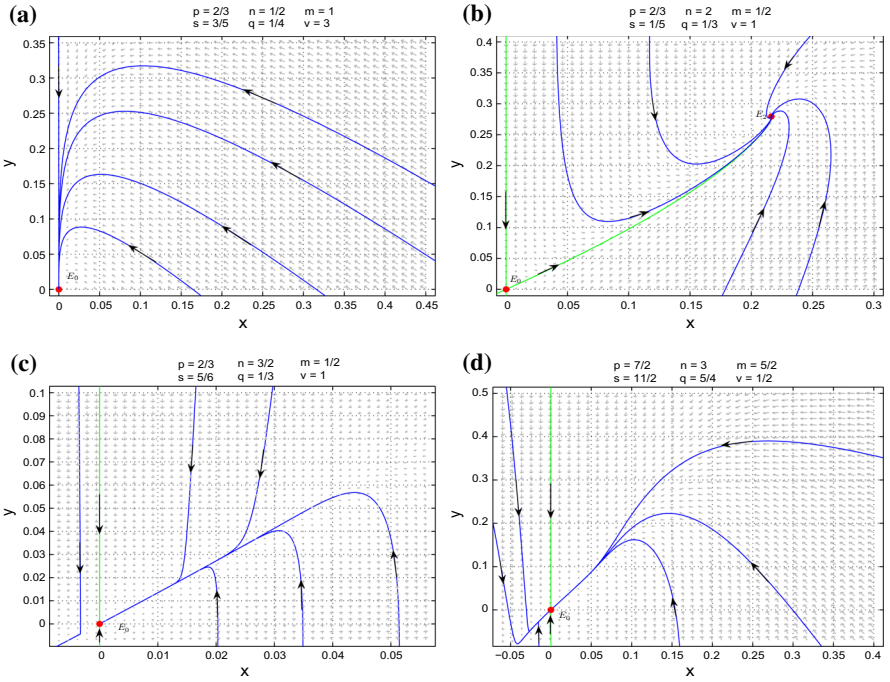


Fig. 1 Types of disease-free equilibrium $E_0(0,0)$ of system (2.2). **a** A hyperbolic stable node if $0 < \mathscr{R}_0 < 1$. **b** A hyperbolic saddle if $\mathscr{R}_0 > 1$. **c** A saddle-node with a stable parabolic sector in the right half plane if $\mathscr{R}_0 = 1$ and $0 < s < \frac{q+v+1}{v}$; **d** A stable degenerate node if $\mathscr{R}_0 = 1$ and $s = \frac{q+v+1}{v}$

Remark 2.1 From (2.3) and (2.6), using the original parameters, we have $\mathscr{R}_0 = \frac{kb}{d(d+\mu+r)}$, it is obvious that $\mathscr{R}_0$ decreases as r increases, which implies that we will overestimate the risk of the disease if we take $r = 0$ in system (2.1).

We have the following results for the types of E_0, which are important to determine the global dynamics in D_1.

Theorem 2.2 *The disease-free equilibrium $E_0(0,0)$ of system (2.2) is*

(I) *a hyperbolic stable node if $0 < \mathscr{R}_0 < 1$;*
(II) *a hyperbolic saddle if $\mathscr{R}_0 > 1$;*
(III) *a degenerate equilibrium if $\mathscr{R}_0 = 1$, more precisely,*

(i) *when $s \neq \frac{q+v+1}{v}$, $E_0(0,0)$ is a saddle-node with a stable (or unstable) parabolic sector in the right half plane if $0 < s < \frac{q+v+1}{v}$ (or $s > \frac{q+v+1}{v}$);*
(ii) *when $s = \frac{q+v+1}{v}$, $E_0(0,0)$ is a stable degenerate node.*

The phase portraits are given in Fig. 1.

Proof The Jacobian matrix of system (2.2) at the equilibrium $E_0(0,0)$ is

$$J(E_0) = \begin{pmatrix} n-m-v & 0 \\ q+v & -1 \end{pmatrix}.$$

It is easy to see that the two eigenvalues of $J(E_0)$ are -1 and $n-m-v$. Obviously, $n-m-v<0$ when $0<\mathscr{R}_0<1$, then $E_0(0,0)$ is a hyperbolic stable node.

When $\mathscr{R}_0=1$, we have $\mathrm{Det}(J(E_0))=0$, $\mathrm{Tr}(J(E_0))=-1$. We first make the following transformation

$$x=\frac{1}{q+v}Y,\ y=X+Y,\ t=-\tau,$$

still denote τ by t, and the Taylor expansions of system (2.2) around origin are as follows:

$$\begin{aligned}\frac{dX}{dt}&=X-XY+\tilde{a}_{02}Y^2+\tilde{a}_{03}Y^3+\tilde{a}_{13}XY^3+\tilde{a}_{04}Y^4+o(|X,Y|^4),\\ \frac{dY}{dt}&=\tilde{b}_{02}Y^2+XY+\tilde{b}_{03}Y^3+\tilde{b}_{13}XY^3+\tilde{b}_{04}Y^4+o(|X,Y|^4),\end{aligned}\tag{2.7}$$

where

$$\begin{aligned}&\tilde{a}_{02}=-\frac{(q+v+1)(q-sv+v)}{(q+v)^2},\ \tilde{a}_{03}=-\frac{p(m+v)(q+v)+s^2v(q+v+1)}{(q+v)^3},\\ &\tilde{a}_{13}=\frac{p}{(q+v)^2},\\ &\tilde{a}_{04}=\frac{(q+v+1)\left(p(q+v)+s^3v\right)}{(q+v)^4},\ \tilde{b}_{02}=\frac{q-sv+v+1}{q+v},\\ &\tilde{b}_{03}=\frac{p(m+v)+s^2v}{(q+v)^2},\ \tilde{b}_{13}=-\frac{p}{(q+v)^2},\ \tilde{b}_{04}=-\frac{p(q+v+1)+s^3v}{(q+v)^3}.\end{aligned}\tag{2.8}$$

By Theorem 7.1 of Zhang et al. (1992), we know that $E_0(0,0)$ is a saddle-node if $s\neq\frac{q+v+1}{v}$.

If $s=\frac{q+v+1}{v}$, then $\tilde{b}_{02}=0$, by the center manifold theorem, we assume $X=m_{11}Y^2+m_{12}Y^3+o(|Y|^3)$ and substitute it into the first equation of (2.7). By using the second equation of system (2.7), we have

$$\begin{aligned}m_{11}&=-\frac{q+v+1}{(q+v)^2},\\ m_{12}&=\frac{v^2(p(m+q)+q+2)+v(q(mp+2q+5)+3)+pv^3+(q+1)^3}{v(q+v)^3},\end{aligned}$$

and substitute $X=m_{11}Y^2+m_{12}Y^3+o(|Y|^3)$ into the second equation of system (2.7), then the reduced equation restricted to the center manifold is described as

$$\frac{dY}{dt}=\frac{v(p(m+v)+1)+q^2+q(v+2)+1}{v(q+v)^2}Y^3+o(|Y|^3).\tag{2.9}$$

Notice that $\frac{v(p(m+v)+1)+q^2+q(v+2)+1}{v(q+v)^2} > 0$, by Theorem 7.1 of Zhang et al. (1992), and we have made a time transformation $\tau = -t$, then $E_0(0, 0)$ is a stable degenerate node. □

To find the endemic equilibria of system (2.2), we set

$$f(x) \triangleq mpsx^3 + ((m+v)p + (1+q)s)x^2 + (1+q+v+(m-n)s)x + m - n + v, \tag{2.10}$$

then an endemic equilibrium of system (2.2) is given by $(x, qx + \frac{vx}{1+sx})$, where x is a positive real root of the cubic equation $f(x) = 0$. Then, we discuss the existence of the positive real roots of $f(x) = 0$ (i.e., the existence of the endemic equilibria of (2.2)) according to the properties of (2.10). First of all, we notice that

$$\lim_{x\to-\infty} f(x) = -\infty, \quad \lim_{x\to+\infty} f(x) = +\infty, \quad f(0) = m - n + v, \tag{2.11}$$

and

$$f'(x) = 3mpsx^2 + 2((m+v)p + (1+q)s)x + 1 + q + v + (m-n)s. \tag{2.12}$$

If $0 < \mathscr{R}_0 < 1$ (i.e., $m - n + v > 0$) and $1 + q + v + (m - n)s \geq 0$, all coefficients of $f(x)$ are positive. Thus, $f(x) = 0$ has no positive root due to Descartes' Rule of signs (or the facts $f(0) \geq 0$ and $f'(x) > 0$ for all $x \geq 0$).

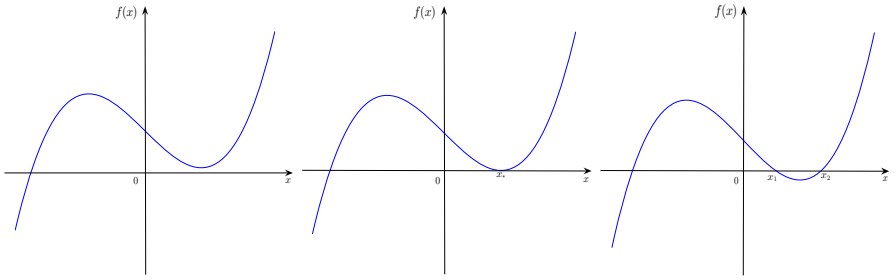
If $0 < \mathscr{R}_0 < 1$ (i.e., $m - n + v > 0$) and $1 + q + v + (m - n)s < 0$, $f'(x) = 0$ has two roots

$$\begin{aligned} \overline{x}_1 &= -\frac{(m+v)p + (1+q)s + \sqrt{((m+v)p + (1+q)s)^2 - 3mps(q+v+1+(m-n)s)}}{3mps} < 0, \\ \overline{x}_2 &= \frac{-(m+v)p + (1+q)s + \sqrt{((m+v)p + (1+q)s)^2 - 3mps(q+v+1+(m-n)s)}}{3mps} > 0. \end{aligned} \tag{2.13}$$

Consequently, $f'(x) > 0$ (i.e., $f(x)$ is an increase function) for $x \in (-\infty, \overline{x}_1) \cup (\overline{x}_2, +\infty)$, but $f'(x) < 0$ (i.e., $f(x)$ is a decreasing function) for $x \in (\overline{x}_1, \overline{x}_2)$. Combining these and the property (2.11), we can see that $f(x) = 0$ has at most two positive real roots (see Fig. 2).

If $\mathscr{R}_0 = 1$ (i.e., $m - n + v = 0$) and $1 + q + v + (m - n)s \geq 0$, it is easy to see that $f(x) = 0$ has no positive root. If $\mathscr{R}_0 = 1$ (i.e., $m - n + v = 0$) and $1 + q + v + (m - n)s < 0$, then $f(x) = 0$ has a unique positive root.

If $\mathscr{R}_0 > 1$ (i.e., $m - n + v < 0$), whether $1 + q + v + (m - n)s \geq 0$ or < 0, $f(x) = 0$ only has a single positive root. In fact, from $\mathscr{R}_0 > 1$, we can get $f(0) < 0$. If $1 + q + v + (m - n)s \geq 0$, then $f'(x) > 0$ for $x > 0$ (i.e., $f(x)$ is increasing in $(0, +\infty)$), and by (2.11), $f(x) = 0$ only has a positive root. If $1+q+v+(m-n)s < 0$, then $f'(x)$ has two roots $\overline{x}_1 < 0 < \overline{x}_2$. Moreover, $f(x)$ is a decreasing function for $x \in (\overline{x}_1, \overline{x}_2)$, then $f(\overline{x}_2) < f(0) < 0$. Since $f(x)$ is increasing in $(\overline{x}_2, +\infty)$, then by (2.11), $f(x) = 0$ only has a positive root.



(a) **(b)** **(c)**

Fig. 2 The positive real roots of $f(x) = 0$ when $0 < \mathscr{R}_0 < 1$ and $1 + q + v + (m - n)s < 0$: **a** no positive root. **b** A double positive root x_* (i.e., $\overline{x}_2$). **c** Two single positive roots x_1, x_2

We further discuss the stability of endemic equilibria. The Jacobian matrix of system (2.2) at any equilibrium $E(x, y)$ is

$$J(E) = \begin{pmatrix} -m - \frac{v}{(sx+1)^2} + \frac{-npx^2+n+px^2y-2x-y}{(px^2+1)^2} & -\frac{x}{px^2+1} \\ \frac{q(sx+1)^2+v}{(sx+1)^2} & -1 \end{pmatrix}.$$

From $f(x) = 0$, we have

$$n = \frac{m\left(px^2+1\right)(sx+1) + v\left(px^2+x+1\right) + (q+1)x(sx+1)}{sx+1}, \tag{2.14}$$

then

$$\begin{aligned} \mathrm{Det}(J(E)) &= m + \frac{v}{(sx+1)^2} - \frac{-npx^2+n+px^2y-2x-y}{\left(px^2+1\right)^2} + \frac{x\left(q(sx+1)^2+v\right)}{\left(px^2+1\right)(sx+1)^2} \\ &= \frac{x\left((sx+1)^2(2mpx+1) + v\left(psx^2+2px-s+1\right) + q(sx+1)^2\right)}{\left(px^2+1\right)(sx+1)^2} \\ &= \frac{x}{\left(px^2+1\right)(sx+1)} f^{'}(x), \end{aligned} \tag{2.15}$$

$$\begin{aligned} \mathrm{Tr}(J(E)) &= -m - \frac{v}{(sx+1)^2} + \frac{-npx^2+n+px^2y-2x-y}{\left(px^2+1\right)^2} - 1 \\ &= -\frac{1}{\left(px^2+1\right)(sx+1)^2}(2m+1)ps^2x^4 + sx^3(p(4m+v+2)+s) \\ &\quad + x^2(p(2m+2v+1)+s(s+2)) + x(1-s(v-2)) + 1. \end{aligned} \tag{2.16}$$

It implies that $E(x, y)$ is a hyperbolic saddle if $f^{'}(x) < 0$, an elementary equilibrium if $f^{'}(x) \neq 0$, and a degenerate equilibrium if $f^{'}(x) = 0$. We have the following results.

Lemma 2.3 *Let $f(x)$ and $\overline{x}_2$ be given by (2.10) and (2.13), respectively. System (2.2) has at most two positive equilibria. Moreover,*

(I) *when $0 < \mathscr{R}_0 < 1$, we have*

(i) *if $n \le \frac{1+q+v+ms}{s}$, or $n > \frac{1+q+v+ms}{s}$ and $f(\overline{x}_2) > 0$, then system (2.2) has no positive equilibrium;*

(ii) *if $n > \frac{1+q+v+ms}{s}$ and $f(\overline{x}_2) = 0$, then system (2.2) has a unique positive equilibrium $E_*(x_*, y_*)$, which is degenerate and $0 < x_* < n$, $y_* = qx_* + \frac{vx_*}{1+sx_*}$;*

(iii) *if $n > \frac{1+q+v+ms}{s}$ and $f(\overline{x}_2) < 0$, then system (2.2) has two positive equilibria $E_1(x_1, y_1)$ and $E_2(x_2, y_2)$, which are all elementary and E_1 is a hyperbolic saddle. $0 < x_1 < x_2 < n$, $y_1 = qx_* + \frac{vx_1}{1+sx_1}$, $y_2 = qx_2 + \frac{vx_2}{1+sx_2}$;*

(II) *when $\mathscr{R}_0 = 1$, we have*

(i) *if $n \le \frac{1+q+v+ms}{s}$, then system (2.2) has no positive equilibrium;*

(ii) *if $n > \frac{1+q+v+ms}{s}$, then system (2.2) has a positive equilibrium $E_2(x_2, y_2)$;*

(III) *when $\mathscr{R}_0 > 1$, then system (2.2) has a positive equilibrium $E_2(x_2, y_2)$.*

Proof From (2.15) and Fig. 2, it is easy to see that $\mathrm{Det}(J(E_1)) < 0$, $\mathrm{Det}(J(E_2)) > 0$ and $\mathrm{Det}(J(E_*)) = 0$, then E_1 and E_2 are all elementary equilibria and only E_1 is a hyperbolic saddle, and E_* is a degenerate equilibrium. □

We first consider **case (I)(ii) of Lemma** 2.3. In this case, system (2.2) has a unique positive equilibrium $E_*(x_*, y_*)$, which is degenerate. From $f(x_*)=f'(x_*)=0$, v and n can be expressed by m, p, s, q and x_* as

$$\begin{aligned} v = v_0 &\triangleq -\frac{(sx_*+1)^2(2mpx_*+q+1)}{s\left(px_*^2-1\right)+2px_*+1}, \\ n = n_0 &\triangleq \\ &-\frac{m\left(s\left(p^2x_*^4+2p(x_*+1)x_*^2+1\right)+px_*^2-1\right)+(q+1)\left(-px_*^2+s(x_*+2)x_*+1\right)}{s\left(px_*^2-1\right)+2px_*+1}. \end{aligned} \tag{2.17}$$

Moreover, from $\mathrm{Tr}(J(E_*)))=0$ and (2.17), m can be expressed by p, s, q and x_* as

$$m = m_0 \triangleq -\frac{2p^2x_*^3+s\left(px_*^2-1\right)\left(px_*^2-qx_*+1\right)+x_*^2(p-2pq)+2px_*+x_*+1}{2px_*^2}. \tag{2.18}$$

Let

$$q_0 = \frac{s^2x_*\left(p^2x_*^4-2px_*^2-3\right)+s\left(3p^2x_*^4+(2p+1)x_*^2+2x_*-1\right)+x_*+2}{sx_*\left(sx_*\left(px_*^2-3\right)+3px_*^2-1\right)}, \tag{2.19}$$

we have the following results.

Theorem 2.4 *If* $0 < \mathscr{R}_0 < 1, n > \frac{1+q+v+ms}{s}$ *and the conditions in* (2.4) *and* (2.17) *are satisfied, then system* (2.2) *has a unique positive equilibrium* $E_*(x_*, y_*)$. *Moreover,*

(I) *if* $m \neq m_0$, *then* E_* *is a saddle-node, which includes a stable parabolic sector (an unstable parabolic sector) if* $m < m_0$ $(m > m_0)$;
(II) *if* $m = m_0$, *then* E_* *is a cusp. Moreover,*

(i) *if* $q \neq q_0$, *then* (x_*, y_*) *is a cusp of codimension 2;*
(ii) *if* $q = q_0$, *then* (x_*, y_*) *is a cusp of codimension 3.*

Proof The Proof of Theorem 2.4 is given in Appendix A. □

Next we discuss the **case (I)(iii)**, **(II)(ii)** and **(III) of Lemma** 2.3. Let

$$\begin{aligned}
\tilde{m}_0 &= -\frac{x_2\left(s^2x_2\left(px_2^2+x_2+1\right)+s\left(p(v+2)x_2^2-v+2x_2+2\right)+p\left(2vx_2+x_2\right)+1\right)+1}{2px_2^2\left(sx_2+1\right)^2},\\
\tilde{v}_0 &= -\frac{\left(px_2^2+x_2+1\right)\left(sx_2+1\right)^2}{x_2\left(s\left(px_2^2-1\right)+2px_2\right)},\\
\tilde{s}_0 &= \frac{2px_2}{1-px_2^2}.
\end{aligned} \tag{2.20}$$

Theorem 2.5 *If equilibria* E_1 *and* E_2 *exist, then* $E_1(x_1, y_1)$ *is always a hyperbolic saddle, and*

(I) *when* $p \geq \frac{1}{x_2^2}$, $E_2(x_2, y_2)$ *is a hyperbolic stable node or focus,*
(II) *when* $0 < p < \frac{1}{x_2^2}$,

(i) $E_2(x_2, y_2)$ *is a hyperbolic unstable node or focus if* $m < \tilde{m}_0$, $v > \tilde{v}_0$ *and* $s > \tilde{s}_0$;
(ii) $E_2(x_2, y_2)$ *is a hyperbolic stable node or focus if* $m > \tilde{m}_0$;
(iii) $E_2(x_2, y_2)$ *is a weak focus or center if* $m = \tilde{m}_0$, $v > \tilde{v}_0$ *and* $s > \tilde{s}_0$.

Proof From (2.15) and Fig. 2, it is easy to see that $\mathrm{Det}(J(E_1)) < 0$ and $\mathrm{Det}(J(E_2)) > 0$, then E_1 and E_2 are all elementary equilibria and E_1 is a hyperbolic saddle. From (2.16), we have

$$\begin{aligned}
\mathrm{Tr}(J(E_2)) = -\frac{1}{(px_2^2+1)(sx_2+1)^2}&\left[2mpx_2^2(sx_2+1)^2+vx_2(s(px_2^2-1)+2px_2)\right.\\
&\left.+(px_2^2+x_2+1)(sx_2+1)^2\right].
\end{aligned} \tag{2.21}$$

If $p \geq \frac{1}{x_2^2}$, then $\mathrm{Tr}(J(E_2)) < 0$; if $0 < p < \frac{1}{x_2^2}$, it is easy to see that $\mathrm{Tr}(J(E_2)) > 0$, $\mathrm{Tr}(J(E_2)) = 0$ and $\mathrm{Tr}(J(E_2)) < 0$ if $m < \tilde{m}_0$, (the conditions $v > \tilde{v}_0$ and $s > \tilde{s}_0$ guarantee $\tilde{m}_0 > 0$), $m = \tilde{m}_0$ and $m > \tilde{m}_0$, respectively, leading to the conclusions. □

2.3 Bifurcation analysis

In this subsection, we discuss different kinds of bifurcations for system (2.2) in depth.

2.3.1 Saddle-node bifurcation

From Theorem 2.4, we know that the surface

$$\begin{aligned}\mathcal{SN} = \Big\{&(p,\ n,\ m,\ v,\ s,\ q) : 0 < p < \frac{1}{x_*^2},\ m - n + v > 0,\\ &1 + q + v + (m - n)s < 0, q < m < 1 + q, v = v_0, n = n_0, m \neq m_0\}\end{aligned} \tag{2.22}$$

is a *saddle-node bifurcation surface*. When the parameters vary to cross the surface from one side to the other, the number of positive equilibria of system (2.2) change from zero to two, the saddle-node bifurcation yields two positive equilibria.

2.3.2 Backward bifurcation

From Wang (2006) (see also Lu et al. (2021) and Zhang and Liu (2008)), backward bifurcation is an interesting and crucial topic in epidemic models. For some classical epidemic models, the basic reproduction number $\mathscr{R}_0 = 1$ serves as a threshold in the sense that a disease is persistent if $\mathscr{R}_0 > 1$ and goes extinct if $\mathscr{R}_0 < 1$, where the transition from a disease-free equilibrium to an endemic equilibrium is called a forward bifurcation. While this bifurcation is backward if an endemic equilibrium occurs when $\mathscr{R}_0 < 1$. From Lemma 2.3, we know that system (2.2) can have one or two positive equilibria when $0 < \mathscr{R}_0 < 1, n > \frac{1+q+v+ms}{s}$ and $f(\overline{x}_2) \leq 0$. Then we have the following result.

Theorem 2.6 *System (2.2) admits a backward bifurcation as $\mathscr{R}_0$ crosses one if $n > \frac{1+q+v+ms}{s}$ and $f(\overline{x}_2) \leq 0$.*

2.3.3 Degenerate Bogdanov-Takens bifurcation of codimension three

The case (II)(ii) of Theorem 2.4 indicates that system (2.2) may exhibit a degenerate Boganov-Takens bifurcation of codimension 3 around $E_*(x_*, y_*)$. In this subsection, we want to make sure if such a bifurcation can be fully unfolded inside the class of system (2.2).

Let

$$\begin{aligned}\Gamma \triangleq \Big\{&(m,\ n,\ p,\ v,\ s,\ q,\ x_*) : v = v_0,\ n = n_0,\ m = m_0,\ q = q_0,\ 0 < x_* < \frac{2}{5},\\ &0 < p < \frac{1}{x_*^2},\ s > s_1,\ \max\{0,\ q_2\} < q_0 < q_1\Big\},\end{aligned}$$

where v_0, n_0, m_0, q_0, s_1, q_1 and q_2 are given in (2.17), (2.18), (2.19) and (A10) of Appendix A, respectively.

Firstly, we choose n, m and q as bifurcation parameters and obtain the following unfolding system:

$$\begin{aligned}\frac{dx}{dt} &= \frac{x}{1+px^2}(n+r_1-x-y)-(m+r_2)x-\frac{vx}{1+sx},\\ \frac{dy}{dt} &= (q+r_3)x-y+\frac{vx}{1+sx},\end{aligned} \tag{2.23}$$

where $(m,\ n,\ p,\ v,\ s,\ q)\in\Gamma$ and $(r_1, r_2, r_3)\sim(0, 0, 0)$. If we can transform, by a series of near-identity transformations, the unfolding system (2.23) into the following versal unfolding of a cusp of codimension 3:

$$\begin{aligned}\frac{dx}{dt} &= y,\\ \frac{dy}{dt} &= \gamma_1+\gamma_2 y+\gamma_3 xy+x^2\pm x^3 y+R(x,y,r),\end{aligned} \tag{2.24}$$

where

$$\begin{aligned}R(x,y,r) = {}& y^2O(|x,y|^2)+O(|x,y|^5)\\ &+O(r)((O(y^2)+O(|x,y|^3))+O(r^2)O(|x,y|),\end{aligned} \tag{2.25}$$

and $\left|\frac{\partial(\gamma_1,\gamma_2,\gamma_3)}{\partial(r_1,r_2,r_3)}\right|_{r=0}\neq 0$, then we can claim that system (2.23) (i.e., system (2.2)) undergoes a degenerate Bogdanov-Takens bifurcation of codimension three (see Dumortier et al. (1987) and Chow et al. (1994)). In fact, we have the following theorem.

Theorem 2.7 *When* $(m,\ n,\ p,\ v,\ s,\ q)\in\Gamma$ *and* $p\neq\frac{-3s^2(x_*+1)x_*^2+4sx_*+1}{x_*^2(s^2x_*^2+(4s+3)x_*+9)}$, *system* (2.2) *has a unique positive equilibrium* $E_*(x_*, y_*)$, *which is a nilpotent cusp of codimension* 3. *If we choose* n, m *and* q *as bifurcation parameters, then system* (2.2) *can undergo a degenerate Bogdanov-Takens bifurcation of codimension* 3 *in a small neighborhood of* E_*. *Hence, system* (2.2) *can exhibit the coexistence of a stable homoclinic loop and an unstable limit cycle, coexistence of two limit cycles (the inner one unstable), and a semi-stable limit cycle for different sets of parameters.*

Proof Following the procedure in Li et al. (2015), we use several steps to transform system (2.23) into the versal unfolding of a Bogdanov-Takens singularity (cusp case) of codimension three. We make the following transformations successively:

$$\begin{aligned}&X=x-x_*,\quad Y=y-y_*;\quad X_1=X,\quad Y_1=\frac{dX}{dt};\\ &X_1=X_2+\frac{c_{02}}{2}X_2^2,\quad Y_1=Y_2+c_{02}X_2Y_2;\\ &X_2=X_3+\frac{d_{12}}{6}X_3^3,\quad Y_2=Y_3+\frac{d_{12}}{2}X_3^2Y_3;\\ &X_3=X_4+\frac{e_{22}}{12}X_4^4,\quad Y_3=Y_4+\frac{e_{22}}{3}X_4^3Y_4;\\ &X_4=X_5-\frac{f_{30}}{4f_{20}}X_5^2+\frac{15f_{30}^2-16f_{20}f_{40}}{80f_{20}^2}X_5^3,\quad Y_4=Y_5,\end{aligned}$$

$$t = \left(1 - \frac{f_{30}}{2f_{20}}X_5 + \frac{45f_{30}^2 - 48f_{20}f_{40}}{80f_{20}^2}X_5^2\right)\tau_1;$$

$$X_5 = X_6,\ \ Y_5 = Y_6 + \frac{g_{21}}{3g_{20}}Y_6^2 + \frac{g_{21}^2}{36g_{20}^2}Y_6^3,\ \ \tau_2 = \left(1 + \frac{g_{21}}{3g_{20}}Y_6 + \frac{g_{21}^2}{36g_{20}^2}Y_6^2\right)\tau_1;$$

$$X_6 = h_{20}^{\frac{1}{5}}h_{31}^{-\frac{2}{5}}X_7,\ \ Y_6 = -h_{20}^{\frac{4}{5}}h_{31}^{-\frac{3}{5}}Y_7,\ \ \tau_2 = -h_{20}^{-\frac{3}{5}}h_{31}^{\frac{1}{5}}\tau_3;\ \ \ X_7 = X_8 - \frac{i_{10}}{2},\ \ Y_7 = Y_8,$$

where the expressions of $c_{02}, d_{12}, e_{22}, f_{20}, f_{30}, f_{40}, g_{20}, g_{21}, h_{20}$ and h_{31} are given in supplementary materials. Then system (2.23) becomes (still denote τ_3 by t)

$$\begin{aligned}\frac{X_8}{dt} &= Y_8,\\ \frac{Y_8}{dt} &= \mu_1 + \mu_2 Y_8 + \mu_3 X_8 Y_8 + X_8^2 - X_8^3 Y_8 + R_3(X_8, Y_8, r),\end{aligned} \tag{2.26}$$

where $R_3(X_8, Y_8, r)$ has the property of (2.25). To save space, we omit the coefficient μ_1, μ_2 and μ_3 here. With the help of Mathematica software, we obtain that

$$\begin{aligned}&\left|\frac{\partial(\mu_1,\mu_2,\mu_3)}{\partial(r_1,r_2,r_3)}\right|_{r=0}\\ &= \frac{16\cdot(2)^{2/5}px_*^5\left[px_*^2(s^2x_*^2+(4s+3)x_*+9)+(3s^2x_*^3+3s^2x_*^2-4sx_*-1)\right]k_{11}^{\frac{4}{5}}k_{13}^{\frac{8}{5}}}{x_*^{\frac{2}{5}}(px_*^2+1)^{\frac{11}{5}}(sx_*+1)^{\frac{13}{5}}k_{12}^{\frac{21}{5}}}\\ &\neq 0\end{aligned} \tag{2.27}$$

since $p \neq \frac{-3s^2(x_*+1)x_*^2+4sx_*+1}{x_*^2(s^2x_*^2+(4s+3)x_*+9)}$, $k_{11} < 0$ and $k_{12}\cdot k_{13} < 0$, which have been shown in the Proof of Theorem 2.4, and k_{11}, k_{12} and k_{13} are given in (A19) of Appendix A. System (2.26) is exactly in the form of system (2.24). Therefore, system (2.23) undergoes a degenereate Bogdanov-Takens of codimension 3 in a small neighborhood of E_*.

Following Figs. 2-5 in Dumortier et al. (1987), we next describe the bifurcation diagram and phase portraits in system (2.26), the bifurcation diagram has the conical structure in $\mathbb{R}^3$, starting from (μ_1, μ_2, μ_3)=(0, 0, 0). It can be best shown by drawing its intersection with the half sphere

$$S = \{(\mu_1, \mu_2, \mu_3) | \mu_1^2 + \mu_2^2 + \mu_3^2 = \epsilon, \epsilon > 0 \text{ small enough}\}.$$

To clearly see the traces of the intersections, we draw the projections of traces onto the (μ_2, μ_3)-plane in Fig. 3.

In Fig. 3, both curves H and C are tangent to ∂S (the boundary of S) at the points $b_1 = (0, 0, \epsilon)$ and $b_2 = (0, 0, -\epsilon)$, and the two curves cross each other at point d. Along ∂S, the circle $\mu_1^2 + \mu_2^2 = \epsilon^2$, excluding the points b_1 and b_2, is a saddle-node bifurcation curve, while b_1 and b_2 correspond to Bogdanov-Takens bifurcation points of codimension 2; H is the Hopf bifurcation curve, on which h_2 is a Hopf bifurcation

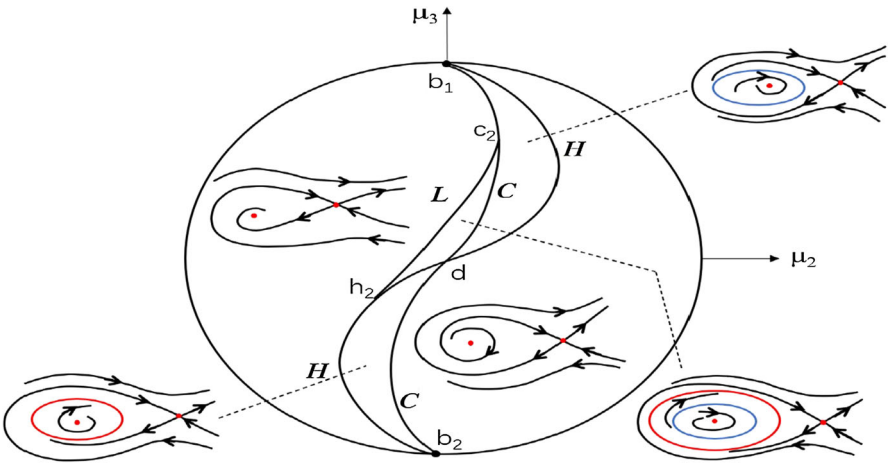


Fig. 3 The projection of bifurcation diagram for system (2.26) on S

point of codimension 2; C is the homoclinic bifurcation curves, on which c_2 is a homoclinic bifurcation point of codimension 2; L is the saddle-node bifurcation curve of limit cycles, which is tangent to curves H and C at h_2 and c_2, respectively. □

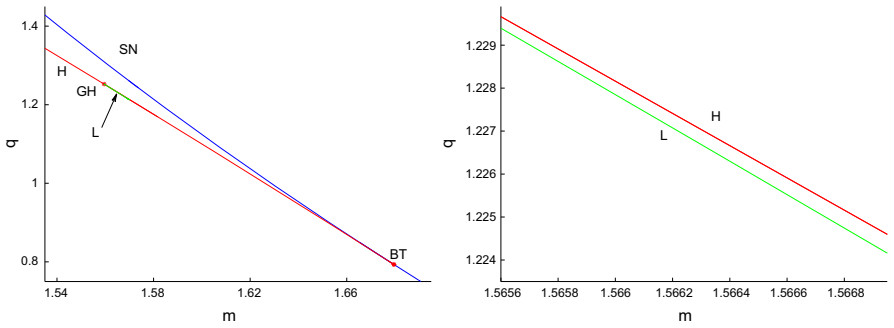
We next present some numerical simulations plotted by *Matcont* to illustrate the existence of almost all bifurcations appearing in Fig. 3. Firstly, we obtain the numerical bifurcation diagram of system (2.2) in (m, q) plane as shown in Fig. 4 by fixing $p = 0.35$, $n = 8.2$, $s = 1.9$, $v = 7$. While we could not numerically plot homoclinic bifurcation curve C, we will give evidences for its existence in Fig. 4. Indeed, we plot some phase portraits in Figs. 5, 6 and 7, where parameter values and corresponding dynamical behaviors are given in Tables 2, 3 and 4, respectively. The parameter values are taken from three parallel lines $q = 1.225$, $q = 1.236$ and $q = 1.27$ in Fig. 4, respectively.

From Fig. 5 and Table 2, we can see that when m decreases, system (2.2) undergoes successively saddle-node bifurcation, subcritical Hopf bifurcation, attracting homoclinic bifurcation and saddle-node bifurcation of limit cycles.

From Fig. 6 and Table 3, as m decreases, system (2.2) undergoes successively saddle-node bifurcation, attraction homoclinic bifurcation, subcritical Hopf bifurcation and saddle-node bifurcation of limit cycles;

From Fig. 7 and Table 4, as m decreases, system (2.2) undergoes successively attraction homoclinic bifurcation, supercritical Hopf bifurcation.

From the phase portraits in Figs. 5, 6 and 7, we can infer that there exists a curve C in Fig. 4. As m decreases, the curve C is tangent to the curve SN at BT point, intersects the curve H at d point and finally intersects the curve L at c_2 point.



(a) **(b)**

Fig. 4 **a** Bifurcation diagram for system (2.2) in $(m,\ q)$ plane when $p = 0.35, n = 8.2, s = 1.9, v = 7$. BT and GH denote Bogdanov-Takens bifurcation point and degenerate Hopf bifurcation point, respectively. Blue, red and green curves denote saddle-node bifurcation SN, Hopf bifurcation H, saddle-node bifurcation L of limit cycles, respectively. **b** The local enlarged view of (a) (colour figure online)

Table 2 $p = 0.35, n = 8.2, s = 1.9, v = 7$ and $q = 1.225$

m	Positive equilibria and types	Closed orbits and homoclinic orbits
1.59	No	No(Fig. 5a)
1.57	E_1(saddle), E_2(unstable focus)	No (Fig. 5b)
1.566815	E_1(saddle), E_2(stable focus)	Unstable limit cycle (Fig. 5c)
1.566758405	E_1(saddle), E_2(stable focus)	Unstable limit cycle, a homoclinic orbit (Fig. 5d)
1.5667506	E_1(saddle), E_2(stable focus)	Unstable limit cycle, a stable limit cycle (Fig. 5e)
1.56	E_1(saddle), E_2(stable focus)	No (Fig. 5f)

Table 3 $p = 0.35, n = 8.2, s = 1.9, v = 7$ and $q = 1.236$

m	Positive equilibria and types	Closed orbits and homoclinic orbits
=1.58	No	No (Fig. 6a)
1.569	E_1(saddle), E_2(unstable focus)	No (Fig. 6b)
1.563985	E_1(saddle), E_2(unstable focus)	A homoclinic orbit (Fig. 6c)
1.563956	E_1(saddle), E_2(unstable focus)	A stable limit cycle (Fig. 6d)
1.56389	E_1(saddle), E_2(stable focus)	A stable limit cycle, unstable limit cycle (Fig. 6e)
1.562	E_1(saddle), E_2(stable focus)	No (Fig. 6f)

Table 4 $p = 0.35$, $n = 8.2$, $s = 1.9$, $v = 7$ and $q = 1.27$

m	Positive equilibria and types	Closed orbits and homoclinic orbits
1.56	E_1(saddle), E_2(unstable focus)	No (Fig. 7a)
1.555476	E_1(saddle), E_2(unstable focus)	A homoclinic orbit (Figure 7b)
1.55495	E_1(saddle), E_2(unstable focus)	A stable limit cycle (Fig. 7c)
1.55	E_1(saddle), E_2(stable focus)	No (Fig. 7d)

2.3.4 Hopf bifurcation

Now we discuss Hopf bifurcation around $E_2(x_2, y_2)$ in system (2.2). Let

$$\begin{aligned}
L_0 &= 2((sx_2+1)^2(ps^3x_2^7(pv+s)+ps^2x_2^6(4pv+3s^2+4s)-s^2x_2^5(p(7sv-12s-6)+s^2)\\
&\quad -sx_2^4(2p(4sv-9s-2)+(s+4)s^2)+x_2^3(p(-7sv+12s+1)-2s^2(2s+3))\\
&\quad +x_2^2(3p+2s(2sv-3s-2))+x_2(s(v-4)-1)-1)),\\
K_0 &= (3p^2x_2^4+p(8p+1)x_2^3+4px_2^2+(8p+1)x_2+1)(sx_2+1)^6+sv^2x_2(p^3s^3x_2^8\\
&\quad +6p^3s^2x_2^7+p^2sx_2^6(9p-5s^2)-2p^2s(7s+1)x_2^5+x_2^2(p(14-5s)+3(3s-2)s^2)\\
&\quad +psx_2^4(3p+s(3s+10))+2ps(10-7s)x_2^3+2s(3s-5)x_2+s-2)\\
&\quad +v(sx_2+1)^2(3p^3s^3x_2^8+2p^2s^2x_2^7(6p+s)+p^2sx_2^6(3p+s(8-11s))\\
&\quad +2psx_2^5(p(2s+3)-2s^2)-psx_2^4(3p+23s^2+2s+2)-2x_2^3(p(6s^2+2s+1)\\
&\quad +(3s-2)s^2)-x_2^2(p(7s+6)+s(9s^2+2s-6))+(-4s^2+6s+2)x_2-s+2).
\end{aligned}$$

Theorem 2.8 *When* $m = \tilde{m}_0$, $q > \frac{s^2(px_2^4+x_2^2)+2s(px_2^3+x_2)+px_2^2-vx_2+1}{x_2(sx_2+1)^2}$, $L_0 \neq 0$, $q \neq \frac{K_0}{L_0}$ *and the conditions in* (2.4) *are satisfied, then* $E_2(x_2, y_2)$ *is a weak focus with multiplicity one for system* (2.2). *Where* $\tilde{m}_0$ *is given in* (2.20).

Proof When $m = \tilde{m}_0$, we have $q > \frac{s^2(px_2^4+x_2^2)+2s(px_2^3+x_2)+px_2^2-vx_2+1}{x_2(sx_2+1)^2}$ from $\mathrm{Det}(J(E_2)) > 0$. We make the following transformations successively:

$$\begin{aligned}
&\hat{X} = x - x_2, \quad \hat{Y} = y - y_2;\\
&\hat{X} = \hat{X}_1, \quad \hat{Y} = \frac{1+px_2^2}{x_2}\hat{X}_1 - \frac{\omega(1+px_2^2)}{x_2}\hat{Y}_1, \quad \hat{\tau} = \omega t,
\end{aligned} \tag{2.28}$$

where $y_2 = qx_2 + \frac{vx_2}{1+sx_2}$ and $\omega = \sqrt{\frac{x_2(-s(px_2^2+1)(sx_2+2)-px_2+q(sx_2+1)^2+v)-1}{(px_2^2+1)(sx_2+1)^2}}$, then system (2.2) becomes as (still denote $\hat{X}_1$, $\hat{Y}_1$ and $\hat{\tau}$ by x, y and t, respectively)

$$\begin{aligned}
\frac{dx}{dt} &= y + s_{20}x^2 + s_{11}xy + s_{02}y^2 + s_{30}x^3 + s_{21}x^2y\\
&\quad +s_{12}xy^2 + s_{03}y^3 + s_{40}x_4 + s_{31}x_3y + s_{22}x_2y_2\\
&\quad +s_{13}xy^3 + s_{04}y^4 + s_{50}x_5 + s_{41}x_4y + s_{32}x_3y_2
\end{aligned}$$

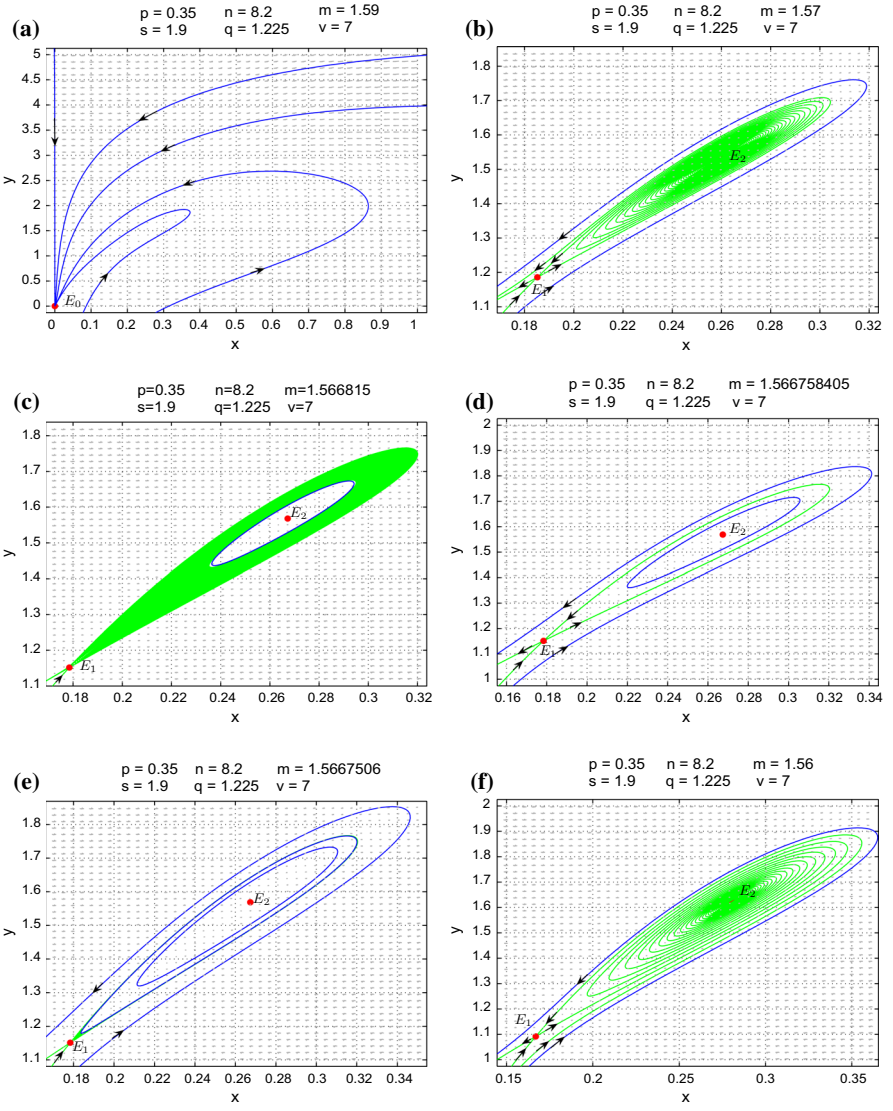


Fig. 5 Phase portraits of system (2.2) with $p = 0.35$, $n = 8.2$, $s = 1.9$, $v = 7$ and $q = 1.225$. **a** $m = 1.59$; **b** $m = 1.57$; **c** $m = 1.566815$; **d** $m = 1.566758405$; **e** $m = 1.5667506$; **f** $m = 1.56$. The detailed dynamical behaviors are described in Table 2

$$\begin{aligned}
&+s_{23}x^2y^3 + s_{14}xy^4 + s_{05}y^5 + o(|x, y|^5),\\
\frac{dy}{dt} = &-x + r_{20}x^2 + r_{11}xy + r_{02}y^2 + r_{30}x^3\\
&+r_{21}x^2y + r_{12}xy^2 + r_{03}y^3 + r_{40}x_4 + r_{31}x_3y + r_{22}x_2y_2\\
&+r_{13}xy^3 + r_{04}y^4 + r_{50}x_5 + r_{41}x_4y + r_{32}x_3y_2\\
&+r_{23}x^2y^3 + r_{14}xy^4 + r_{05}y^5 + o(|x, y|^5),
\end{aligned} \tag{2.29}$$

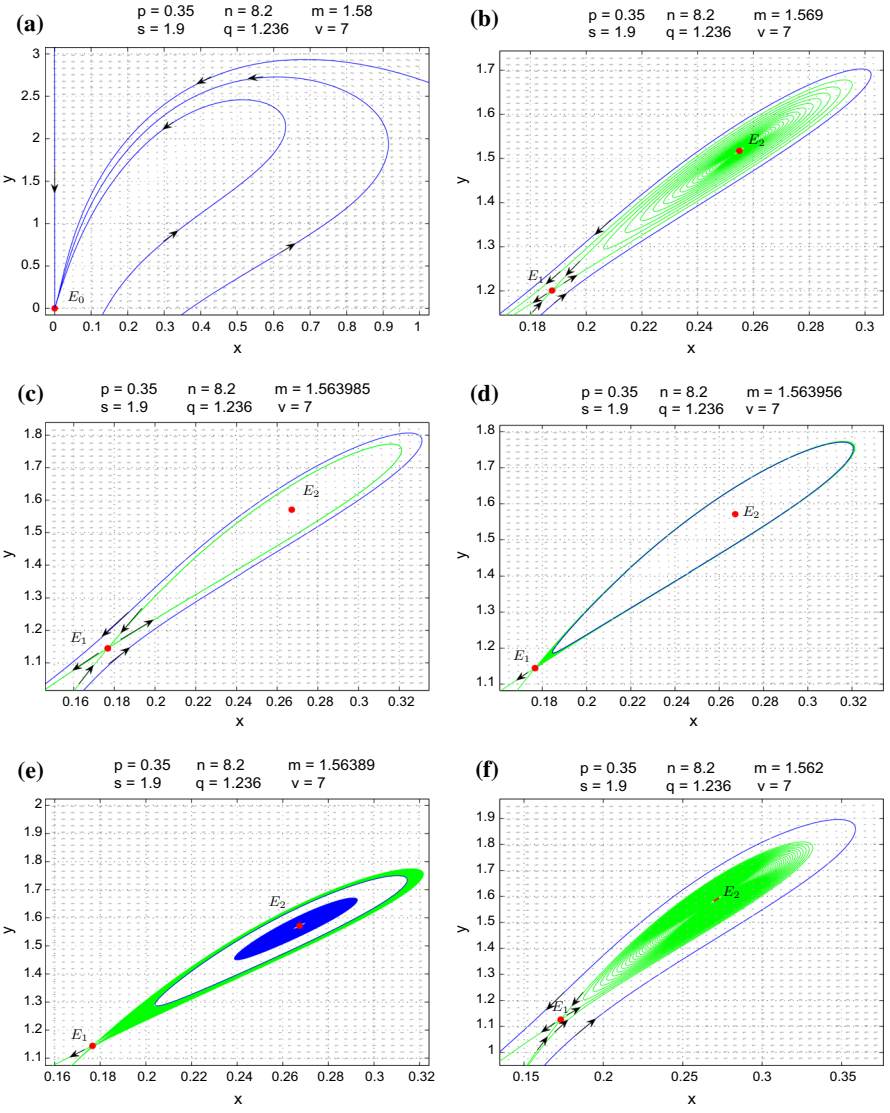


Fig. 6 Phase portraits of system (2.2) with $p = 0.35$, $n = 8.2$, $s = 1.9$, $v = 7$ and $q = 1.236$. **a** $m = 1.58$; **b** $m = 1.569$; **c** $m = 1.563985$; **d** $m = 1.563956$; **e** $m = 1.56389$; **f** $m = 1.562$. The detailed dynamical behaviors are described in Table 3

where r_{ij} and s_{ij} $(i, j = 0, 1, 2, 3, 4, 5)$ are given in supplementary materials.

Using the formal series method in Zhou et al. (2014) and *Mathematica* software, we obtain the first Lyapunov coefficient

$$\sigma_1 = -\frac{K_0 - qL_0}{8x_2\omega\left(px_2^2+1\right)^2(sx_2+1)^4\{x_2\left(s\left(px_2^2+1\right)(sx_2+2)+px_2-q(sx_2+1)^2-v\right)+1\}}. \tag{2.30}$$

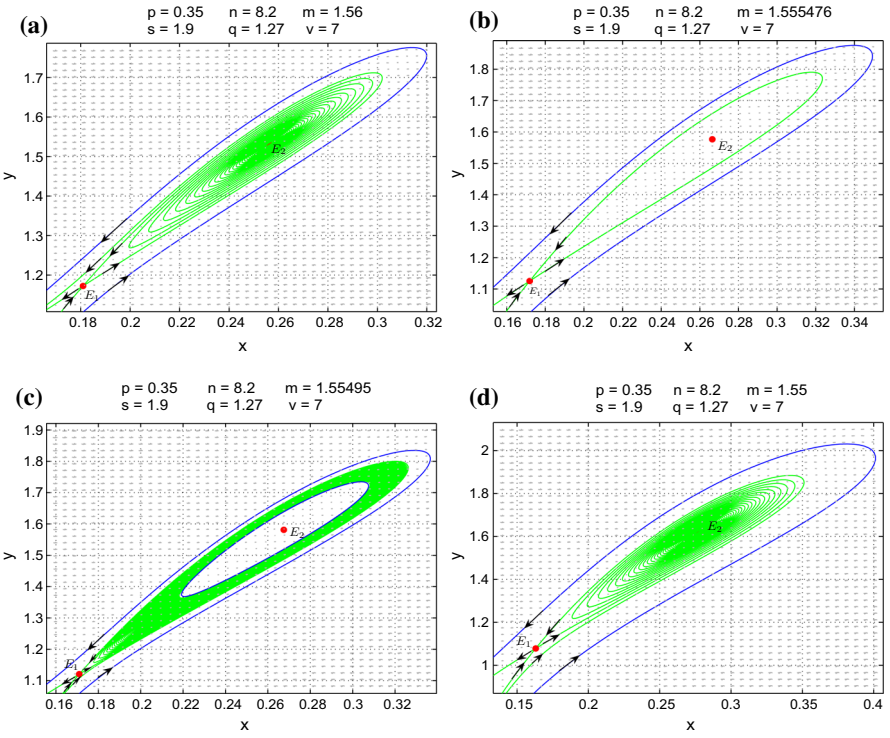


Fig. 7 Phase portraits of system (2.2) with $p = 0.35$, $n = 8.2$, $s = 1.9$, $v = 7$ and $q = 1.27$. **a** $m = 1.56$; **b** $m = 1.555476$; **c** $m = 1.55495$; **d** $m = 1.55$. The detailed dynamical behaviors are described in Table 4

Since $q > \frac{s^2(px_2^4+x_2^2)+2s(px_2^3+x_2)+px_2^2-vx_2+1}{x_2(sx_2+1)^2}$, then the denominator of σ_1 is negative, therefore the sign of σ_1 is determined by $K_0 - qL_0$. If $L_0 \neq 0$ and $q \neq \frac{K_0}{L_0}$, then $\sigma_1 \neq 0$, which leads to the conclusion. □

There exists a set of parameter values $(v, s, p, q, x_2) = (6, 6, \frac{13}{16}, \frac{532848289161357}{51529374176000}, \frac{1}{8})$ that satisfies the conditions of Theorem 2.8 such that $\sigma_1 \doteq -0.0264 < 0$. On the other hand, there exists a set of parameter values $(v, s, p, q, x_2) = (6, 6, \frac{13}{16}, \frac{20371974606517}{2061174967040}, \frac{1}{8})$ that also satisfies the conditions of Theorem 2.8 such that $\sigma_1 \doteq 2.0185 > 0$. Therefore, there exists an open set V_1 in the parameter space (m, n, v, s, p, q, x_2) such that $\sigma_1 < 0$, i.e.,

$$V_1 = \left\{(m, n, v, s, p, q, x_2) : n, v, s, x_2 > 0,\ f(x_2) = 0,\ m = \tilde{m}_0,\ 0 < p < \frac{1}{x_2^2},\right.$$
$$\left. q < \tilde{m}_0 < 1 + q,\ q > \frac{s^2\left(px_2^4 + x_2^2\right) + 2s\left(px_2^3 + x_2\right) + px_2^2 - vx_2 + 1}{x_2\left(sx_2 + 1\right)^2},\ \sigma_1 < 0\right\}.$$

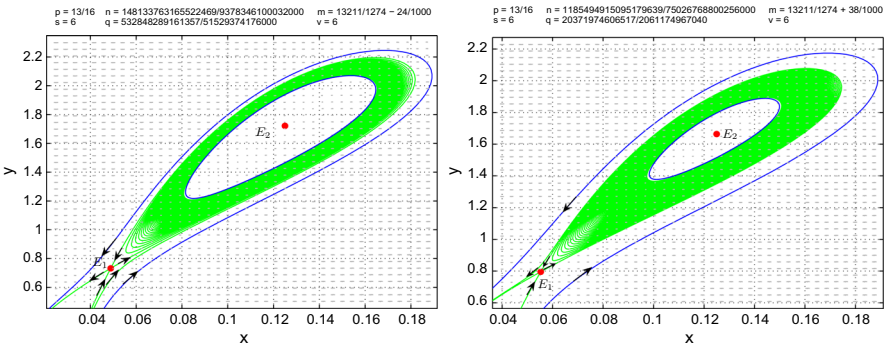


Fig. 8 **a** A stable limit cycle arising from supercritical Hopf bifurcation. **b** An unstable limit cycle arising from subcritical Hopf bifurcation

And there exists another open set V_2 in the parameter space $(m,\ n,\ v,\ s,\ p,\ q,\ x_2)$ such that $\sigma_1 > 0$, i.e.,

$$V_2 = \Big\{(m,\ n,\ v,\ s,\ p,\ q,\ x_2) : n,\ v,\ s,\ x_2 > 0,\ f(x_2) = 0,\ m = \tilde{m}_0,\ 0 < p < \frac{1}{x_2^2},$$
$$q < \tilde{m}_0 < 1 + q,\ q > \frac{s^2\left(px_2^4 + x_2^2\right) + 2s\left(px_2^3 + x_2\right) + px_2^2 - vx_2 + 1}{x_2\left(sx_2 + 1\right)^2},\ \sigma_1 > 0\Big\}.$$

Summarizing the above discussion, we have the following results.

Theorem 2.9 (i) *If $m = \tilde{m}_0$ and $(m,\ n,\ v,\ s,\ p,\ q,\ x_2) \in V_1$, then the equilibrium $E_2(x_2,\ y_2)$ of system (2.2) is a stable weak focus with multiplicity one. There exists a stable limit cycle arising from supercritical Hopf bifurcation (see Fig. 8a);*

(ii) *If $m = \tilde{m}_0$ and $(m,\ n,\ v,\ s,\ p,\ q,\ x_2) \in V_2$, then the equilibrium $E_2(x_2,\ y_2)$ of system (2.2) is an unstable weak focus with multiplicity one. There exists an unstable limit cycle arising from subcritical Hopf bifurcation (see Fig. 8b).*

Using the formal series method in Zhou et al. (2014) and Mathematica software, when $L_0 \neq 0$ and $q = \frac{K_0}{L_0}$, we can obtain the second Lyapunov coefficient σ_2, which is given in supplementary materials.

Next, we present an example to show that the positive equilibrium $E_2(x_2,\ y_2)$ of system (2.2) is a stable weak focus of multiplicity two, system (2.2) can undergo a degenerate Hopf bifurcation around E_2 and two limit cycles occur.

Theorem 2.10 *When $(m,\ n,\ v,\ s,\ p,\ q) = (\frac{13211}{1274},\ \frac{2373655918142609}{150053537600512},\ 6,\ 6,\ \frac{13}{16},\ \frac{4259900668337}{412234993408})$, system (2.2) has a positive equilibrium $E_2(x_2,\ y_2)$ which is a stable weak focus of multiplicity two. There exist two limit cycles arising from degenerate Hopf bifurcation around E_2, the repelling cycle is surrounded by an attracting one (see Fig. 9).*

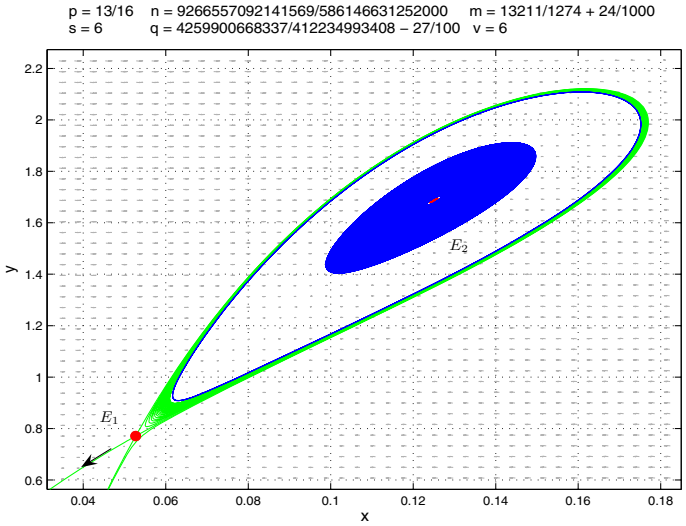


Fig. 9 Two limit cycles (the inner one is unstable) around E_2 in system (2.2)

Proof We first fix $x_2 = \frac{1}{8}, s = 6, v = 6$ and $p = \frac{13}{16}$, from the expression of σ_2, we get $\sigma_2 \doteq -216.829 < 0$, furthermore, from $\sigma_1 = 0$, $\mathrm{Tr}(J(E_2)) = 0$ and $f(x_2) = 0$, we get $q = \frac{4259900668337}{412234993408}, m = m_0 = \frac{13211}{1274}$ and $n = \frac{2373655918142609}{150053537600512}$, respectively. Therefore, the positive equilibrium $E_2(x_2, y_2)$ is a stable weak focus of multiplicity two when $(m, n, v, s, p, q) = (\frac{13211}{1274}, \frac{2373655918142609}{150053537600512}, 6, 6, \frac{13}{16}, \frac{4259900668337}{412234993408})$. Next, we can check by Mathematica that

$$\begin{aligned}
&\left|\frac{\partial(\mathrm{Tr}(J(E_2)), \sigma_1)}{\partial(m, q)}\right|_{(m,n,v,s,p,q)=(\frac{13211}{1274}, \frac{2373655918142609}{150053537600512}, 6, 6, \frac{13}{16}, \frac{4259900668337}{412234993408})} \\
&= \frac{45085509536147335610368\sqrt{\frac{460083698}{11663047457}}}{9441154578973488519803 9} \\
&\neq 0,
\end{aligned}$$

which means that $\mathrm{Tr}(J(E_2))$ and σ_1 with respect to m, q have full rank 2. We first perturb q such that q decreases to $q - \frac{27}{100}$, then E_2 becomes an unstable weak focus with multiplicity one, a stable limit cycle occurs around E_2 which is the outer limit cycle in Fig. 9. Secondly, we perturb m such that m increases $m + \frac{24}{1000}$, then E_2 becomes a stable hyperbolic focus, another unstable limit cycle occurs around E_2, which is the inner limit cycle in Fig. 9. □

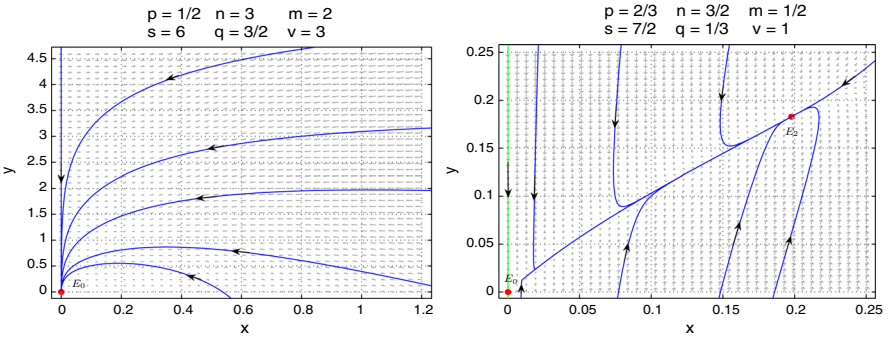


Fig. 10 **a** $E_0(0,0)$ of system (2.2) is globally asymptotically stable when $0 < \mathscr{R}_0 < 1, n > \frac{1+q+v+ms}{s}$ and $f(\overline{x}_2) > 0$. **b** E_2 of system (2.2) is globally asymptotically stable when $\mathscr{R}_0 = 1, s > \frac{q+v+1}{v}$ and $v \leq \frac{1+2s}{s}$

2.4 Global dynamics of system (2.2)

In this subsection, we discuss the global asymptotical stability of the unique boundary equilibrium $E_0(0,0)$ of system (2.2) (i.e., the disease-free euilibrium $(\frac{d}{b},0,0)$ or the unique positive equilibrium E_2 of system (1.2)).

Theorem 2.11 *The unique boundary equilibrium $E_0(0,0)$ of system (2.2) (i.e., the disease-free euilibrium $(\frac{d}{b},0,0)$ of system (1.2)) is globally asymptotically stable if one of the following conditions holds:*

(i) $0 < \mathscr{R}_0 \leq 1$ *and* $n \leq \frac{1+q+v+ms}{s}$ *(see Fig. 1a and c);*

(ii) $0 < \mathscr{R}_0 < 1, n > \frac{1+q+v+ms}{s}$ *and* $f(\overline{x}_2) > 0$ *(see Fig. 10a).*

Proof Lemma 2.1 implies that the stability of disease-free equilibrium $(\frac{d}{b},0,0)$ of system (1.2) in the interior $\mathbb{R}^3_+$ is equivalent to that of the equilibrium $E_0(0,0)$ of system (2.2) in $\mathbb{R}^2_+$, thus we only need to discuss the stability of equilibrium $E_0(0,0)$ of system (2.2) in $\mathbb{R}^2_+$.

On one hand, $D_1 \subset \mathbb{R}^2_+$ is a positive invariant and bounded region and $x = 0$ is an invariant line. On the other hand, from Theorem 2.2, we know that system (2.2) has no positive equilibrium and only has a boundary equilibrium $E_0(0,0)$ when the conditions in (i) or (ii) are satisfied. Thus, Poincaré-Bendixson Theorem implies that $E_0(0,0)$ of system (2.2) is globally asymptotically stable under the condition (i) or (ii). □

Remark 2.2 Using the original parameters, we have $\mathscr{R}_0 < 1 \Longleftrightarrow r > r_1$; $n \leq \frac{1+q+v+ms}{s} \Longleftrightarrow r \geq r_2$, where

$$\begin{aligned} r_1 &= \frac{bk}{d} - d - \mu, \\ r_2 &= \frac{b\beta dk + b\beta\gamma k - \beta d^3 - \beta\gamma d^2 - \beta d^2\mu - d^2 k - \beta\gamma d\mu - \gamma dk - dk\mu}{dk}. \end{aligned} \tag{2.31}$$

From Theorem 2.11, we can see that the disease will disappear for all positive initial populations if one of the following cases holds:

(I.1) $r \geq \max\{r_1, r_2\}$;
(I.2) $r_1 < r < r_2$ and $f(\overline{x}_2) > 0$.

Theorem 2.12 *If the conditions in* (2.4) *and* $v \leq \frac{1+2s}{s}$ *hold, then system* (2.2) *does not have nontrivial periodic orbits in the interior of* $\mathbb{R}^2_+$.

Proof Taking a Dulac function $D(x, y) = \frac{1+px^2}{x}$, we have

$$\begin{aligned}&\frac{\partial(DP)}{\partial x} + \frac{\partial(DQ)}{\partial y}\\ &= -\frac{1}{x(sx+1)^2}\{(2m+1)ps^2x^4 + sx^3(p(4m+v+2)+s)\\ &\quad + x^2(p(2m+2v+1)+s(s+2))\\ &\quad + x(1-s(v-2))+1\},\end{aligned}$$

for $v \leq \frac{1+2s}{s}$ and $x > 0$, we can immediately get

$$\frac{\partial(DP)}{\partial x} + \frac{\partial(DQ)}{\partial y} < 0.$$

Thus, we obtain the conclusion by Dulac's criteria. □

Theorem 2.13 *The unique positive equilibrium* E_2 *is globally asymptotically stable if one of the following conditions is satisfied:*

(I) $\mathscr{R}_0 = 1$, $s > \frac{q+v+1}{v}$ *and* $v \leq \frac{1+2s}{s}$ *(see Fig. 10b);*
(II) $\mathscr{R}_0 > 1$ *and* $v \leq \frac{1+2s}{s}$ *(see Fig. 1b).*

Proof From the Theorem 2.2, Lemma 2.3 and Theorem 2.12, we can see that **(I)** if $\mathscr{R}_0 = 1, s > \frac{q+v+1}{v}$ and $v \leq \frac{1+2s}{s}$, then the unique positive equilibrium E_2 of system (2.2) is globally asymptotically stable since the unique boundary equilibrium E_0 is a saddle-node with a stable parabolic sector lying in the left half plane of $\mathbb{R}^2_+$ and D_1 is a positive invariant and bounded region. **(II)** if $\mathscr{R}_0 > 1$ and $v \leq \frac{1+2s}{s}$, then the unique positive equilibrium E_2 is globally asymptotically stable since the unique boundary equilibrium E_0 is a hyperbolic saddle and D_1 is a positive invariant and bounded region. We use simulations to illustrate the results. We take a set of parameter values: $p = \frac{2}{3}, n = \frac{3}{2}, m = \frac{1}{2}, s = 3, q = \frac{1}{3}, v = 1$ such that $\mathscr{R}_0 = 1, s > \frac{q+v+1}{v}$ and $v \leq \frac{1+2s}{s}$ for case (I), as shown in Fig. 10b: E_2 is globally asymptotically stable. We take a set of parameter values: $p = \frac{2}{3}, n = 2, m = \frac{1}{2}, s = \frac{1}{5}, q = \frac{1}{3}, v = 1$ such that $\mathscr{R}_0 > 1$ and $v \leq \frac{1+2s}{s}$ for case (II), as shown in Fig. 1b: E_2 is globally asymptotically stable. □

Remark 2.3 Using the original parameters, we have $v \leq \frac{1+2s}{s} \Longleftrightarrow r \leq r_3$, where

$$r_3 = 2\gamma + 2d + \frac{k}{\beta}. \tag{2.32}$$

Theorem 2.13 indicates that the disease will persist for all positive initial populations if one of the following cases holds:

(I.1) $r = r_1$ and $r_1 < \min\{r_2, r_3\}$;
(I.2) $r < \min\{r_1, r_3\}$.

Remark 2.4 If the cure rate $r = 0$ (i.e., no treatment), then $v = 0$ and $\mathscr{R}_0 = \frac{n}{m}$. From Theorem 2.2, Lemma 2.3 and Theorem 2.12, if $r = 0$, we have the following conclusions, which are the results in Xiao and Ruan (2007):

(i) When $0 < \mathscr{R}_0 \leq 1$, system (2.2) has a unique disease-free equilibrium E_0, which is a global attractor in the first octant.
(ii) When $\mathscr{R}_0 > 1$, system (2.2) has two equilibria: a disease-free equilibrium E_0 and an endemic equilibrium $E_2(x_2, y_2)$, where E_2 is a global attractor in the interior of the first octant.

3 Changing environment

In this section, we study how the speed of environmental change regulates the dynamics of system (1.3). In detail, we first use the bifurcation software Matcont to plot the one-parameter bifurcation diagram in $k - I$ plane for system (1.2), then plot some "representative" trajectories (time series) for system (1.3) and include them into the bifurcation diagram (see Figs. 11, 12 and 13). In Figs. 11, 12 and 13, the stable and unstable steady states for system (1.2) are described by blue solid and dashed curves, respectively. The maximum and minimum values of stable and unstable oscillations are described by blue filled and open circles, respectively. The "representative" trajectories (time series) for system (1.3) are plotted by red or green solid curves.

In system (1.3), the infection rate $k(t)$ continuously increases over time when $u > 0$, and continuously decreases when $u < 0$. According to the bifurcations and dynamics of system (1.2) given in the previous sections, we classify the effect of k in system (1.3) as the two cases.

Case (I): $\mathscr{R}_0 < 1$ and system (1.2) has no endemic equilibrium.

In this case, the basic reproduction number $\mathscr{R}_0 = 1$ is a threshold in the sense that a disease is persistent if $\mathscr{R}_0 > 1$ and goes extinct if $\mathscr{R}_0 < 1$ in system (1.2).

In Fig. 11a, we choose $b = 1.1557$, $d = 0.121528$, $\alpha = 0.1$, $\beta = 1$, $\mu = 0.03125$, $\gamma = 0.1$, $r = 2$ in system (1.2), then we get $k_{BP} \doteq 0.226376$, $k_{H_1} \doteq 0.235410$ and $k_{H_2} \doteq 0.265639$. We can see that when $0 < k < k_{BP}$, system (1.2) only has a stable disease-free equilibrium $\widetilde{E}_0(S_0, I_0, R_0)$; when $k = k_{BP}$ (i.e., $\mathscr{R}_0 = 1$), system (1.2) undergoes transcritical bifurcation, the disease-free equilibrium $\widetilde{E}_0$ becomes unstable and a stable endemic equilibrium $\widetilde{E}_2(S_2, I_2, R_2)$ occurs when $k > k_{BP}$. As k further increases, $\widetilde{E}_2$ becomes unstable, system (1.2) exhibits supercritical Hopf bifurcation

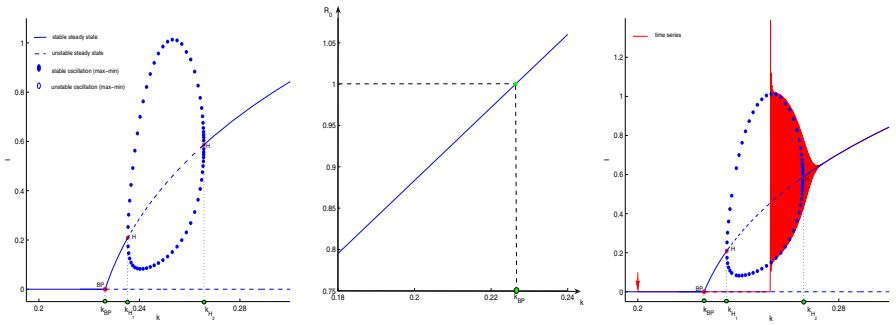


Fig. 11 **a** Bifurcation diagram in $k - I$ plane for system (1.2), where $k_{BP} \doteq 0.226376$, $k_{H_1} \doteq 0.235410$ and $k_{H_2} \doteq 0.265639$. **b** The curve of $\mathscr{R}_0 - k$. **c** Time series in model (1.3) with $u = 0.00001$ and the initial point: $(S(0),\ I(0),\ R(0),\ k_0) = (1,\ 0.1,\ 0.3,\ 0.2)$. Other parameters: $b = 1.1557$, $d = 0.121528$, $\alpha = 0.1$, $\beta = 1$, $\mu = 0.03125$, $\gamma = 0.1$, $r = 2$,

around the unique endemic equilibrium $\widetilde{E}_2$ at $k = k_{H_1}$ or $k = k_{H_2}$, thus system (1.2) has a stable limit cycle when $k_{H_1} < k < k_{H_2}$. In Fig. 11b, we can see that $\mathscr{R}_0$ increases as k increases, and $\mathscr{R}_0 = 1$ if $k = k_{BP}$. In Fig. 11c, we choose $u = 0.00001 > 0$ in system (1.3), the infective population $I(t)$ (red curve) of system (1.3) starts along the stable disease-free state $\widetilde{E}_0$ of system (1.2), tracks the unstable disease-free state $\widetilde{E}_0$ when $k(t) > k_{BP}$ (i.e., $\mathscr{R}_0 > 1$), then tends to the stable oscillation and finally to the stable endemic state $\widetilde{E}_2$.

This transient tracking on the unstable disease-free state $\widetilde{E}_0$ when $\mathscr{R}_0 > 1$ predicts regime shifts that cause the delayed disease outbreak under environmental changes.

Case (II): $\mathscr{R}_0 < 1$ and system (1.2) has endemic equilibria.

In this case, it is not the basic reproduction number $\mathscr{R}_0 = 1$ but a subthreshold $\mathscr{R}_0 = \mathscr{R}_0^*(< 1)$ that determines whether the disease persists in system (1.2), i.e., the disease will disappear when $\mathscr{R}_0 < \mathscr{R}_0^*$ and persist when $\mathscr{R}_0 > \mathscr{R}_0^*$ in system (1.2).

Case (II.1): a limit cycle when $\mathscr{R}_0 < 1$.

In Fig. 12a, we choose $b = 1$, $d = 0.10604$, $\alpha = 2.55$, $\beta = 10.0532$, $\mu = 2$, $\gamma = 0.0963$, $r = 1.8$ in system (1.2), then we get $k_{LP} \doteq 0.385620$, $k_H \doteq 0.395537$ and $k_{BP} \doteq 0.414196$. System (1.2) only has a stable disease-free equilibrium $\widetilde{E}_0$ when $0 < k < k_{LP}$, and exhibits backward bifurcation (i.e., saddle-node bifurcation) at $k = k_{LP}$, where two endemic equilibria $\widetilde{E}_1(S_1, I_1, R_1)$ and $\widetilde{E}_2(S_2, I_2, R_2)$ occur and exist when $k_{LP} < k < k_{BP}$. Furthermore, system (1.2) exhibits subcritical Hopf bifurcation around $\widetilde{E}_2$ at $k = k_H$, which implies the coexistence of two stable equilibria $\widetilde{E}_0$, $\widetilde{E}_2$ and an unstable limit cycle. When $k > k_{BP}$, $\widetilde{E}_0$ becomes unstable, and system (1.2) has a unique endemic equilibrium $\widetilde{E}_2$ which is stable. Figure 12b implies $\mathscr{R}_0 = \mathscr{R}_0^*$ if $k = k_{LP}$, and $\mathscr{R}_0 = 1$ if $k = k_{BP}$. In Fig. 12c, it is shown that the outcome of a disease spread for system (1.3) can heavily depend on the initial infection number $I(0)$ and the initial infection rate k_0 under the same rate u of environmental changes. We can see that $I(t)$ (green curve) of system (1.3) tracks the stable disease-free state $\widetilde{E}_0$, then the unstable disease-free state $\widetilde{E}_0$ when $k(t) > k_{BP}$ (i.e., $\mathscr{R}_0 > 1$), finally the stable endemic state $\widetilde{E}_2$; while $I(t)$ (red curve) of system (1.3) tracks the unstable

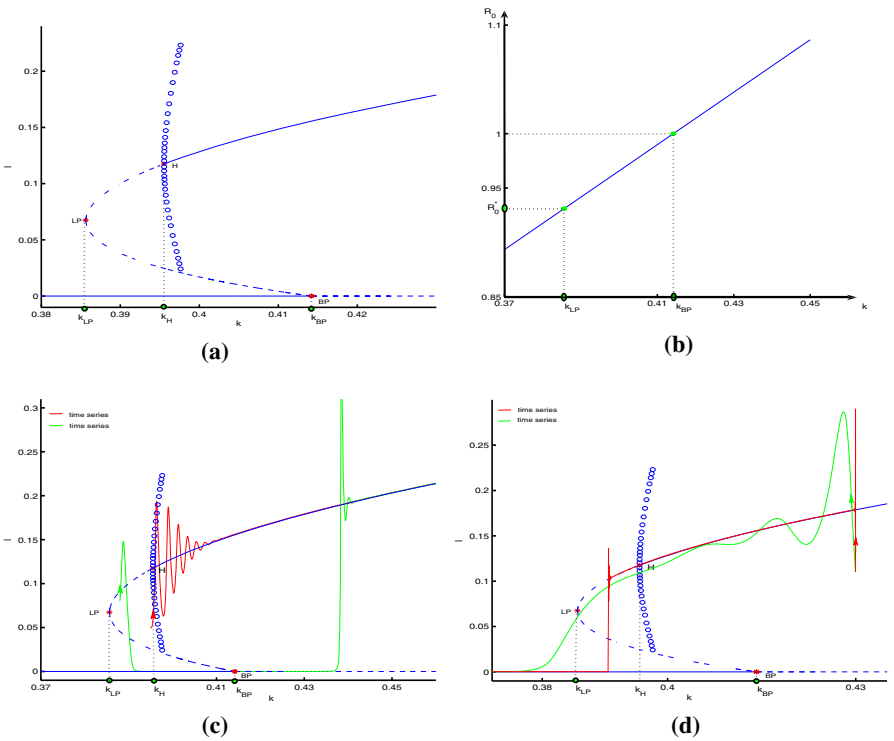


Fig. 12 **a**) Bifurcation diagram in $k - I$ plane for system (1.2) ($k_{LP} \doteq 0.385620$, $k_H \doteq 0.395537$ and $k_{BP} \doteq 0.414196$). **b** The curve of $\mathscr{R}_0 - k$ ($\mathscr{R}_0^* \doteq 0.931006$). **c** Time series in model (1.3) with $u = 0.0001$, different initial points: $(S(0),\ I(0),\ R(0),\ k_0) = (0.5,\ 0.08,\ 1,\ 0.388)$ (green curve) and $(S(0),\ I(0),\ R(0),\ k_0) = (0.5,\ 0.05,\ 1,\ 0.395)$ (red curve). **d** Time series in model (1.3) with $u = -0.0006$ (green curve) or $u = -0.000001$ (red curve), and the same initial point: $(S(0),\ I(0),\ R(0),\ k_0) = (0.5,\ 0.1,\ 1.1,\ 0.43)$. Other parameters: $b = 1, d = 0.10604, \alpha = 2.55, \beta = 10.0532, \mu = 2, \gamma = 0.0963$, $r = 1.8$ (colour figure online)

oscillation, and finally the stable endemic state $\widetilde{E}_2$. In Fig. 12d, we choose $u < 0$ indicating that the infection rate decreases over time in the changing environment (e.g., nonpharmaceutical interventions), the disease will eventually disappear. For $u = -0.000001$, $I(t)$ (red curve) of system (1.3) reaches the stable disease-free state $\widetilde{E}_0$ earlier compared to that of system (1.2), while for $u = -0.0006$, $I(t)$ (green curve) has a delay in disappearance.

Case (II.2): two limit cycles when $\mathscr{R}_0 < 1$.

In Fig. 13a, we choose $b = 1$, $d = 0.10604$, $\alpha = 0.872576$, $\beta = 10.5232$, $\mu = 2$, $\gamma = 0.0963142$, $r = 1.24$ in system (1.2), and get $k_{LP} \doteq 0.336158$, $k_H \doteq 0.344038$ and $k_{BP} \doteq 0.354814$. System (1.2) undergoes degenerate Hopf bifurcation around $\widetilde{E}_2$ at $k = k_H$, and exhibits the coexistence of tristability (two stable equilibria $\widetilde{E}_2$, $\widetilde{E}_0$ and a stable big limit cycle) and an unstable small limit cycle (see Fig. 13a). In Fig. 13b, it is shown that $\mathscr{R}_0 = \mathscr{R}_0^*$ if $k = k_{LP}$, and $\mathscr{R}_0 = 1$ if $k = k_{BP}$. In Fig. 13c, we choose $u > 0$ in system (1.3) indicating that the infection rate increases over time in the changing environment (e.g., more gatherings). $I(t)$ (green curve) of system (1.3)

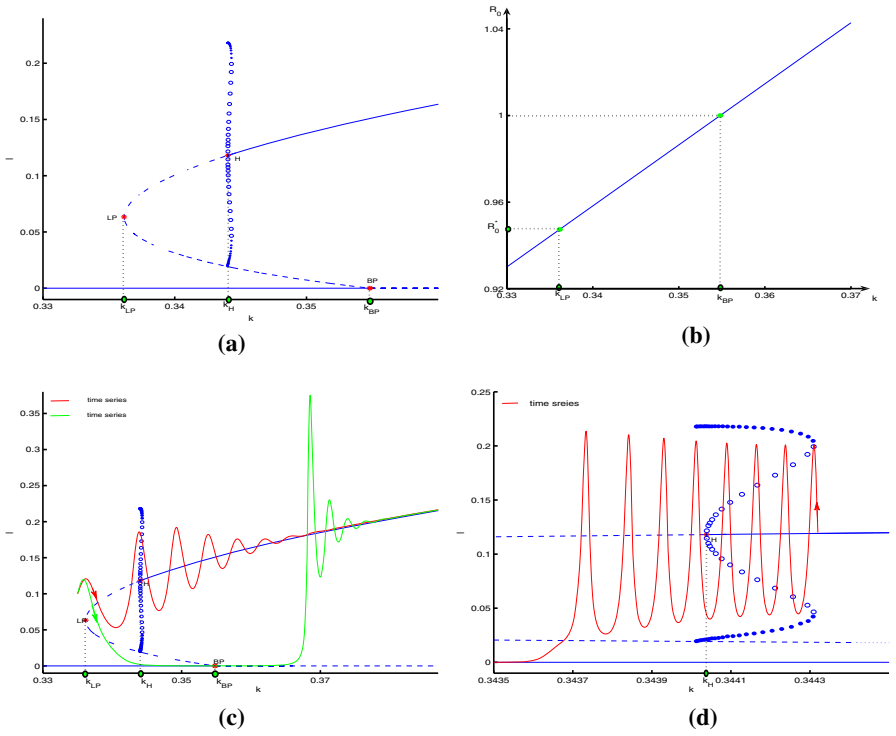


Fig. 13 **a**) Bifurcation diagram in $k - I$ plane for system (1.2) ($k_{LP} \doteq 0.336158$, $k_H \doteq 0.344038$ and $k_{BP} \doteq 0.354814$). **b** The curve of $\mathscr{R}_0 - k$ ($\mathscr{R}_0^* \doteq 0.947421$). **c** Time series in model (1.3) with $u = 0.00015$ (green curve) or $u = 0.0002$ (red curve) and the same initial point: $(S(0),\ I(0),\ R(0),\ k_0) = (6,\ 0.1,\ 1,\ 0.335)$. **d** Time series in model (1.3) with $u = -0.000002$ and the initial point: $(S(0),\ I(0),\ R(0),\ k_0) = (6,\ 0.1,\ 1,\ 0.34432)$. Other parameters: $b = 1$, $d = 0.10604$, $\alpha = 0.872576$, $\beta = 10.5232$, $\mu = 2$, $\gamma = 0.0963142$, $r = 1.24$, (colour figure online)

for smaller u starts along the stable disease-free state $\widetilde{E}_0$ of system (1.2), and then continues following the unstable disease-free state $\widetilde{E}_0$ when $k(t) > k_{BP}$ (i.e., $\mathscr{R}_0 > 1$), and finally tends to the stable endemic state $\widetilde{E}_2$ in an oscillatory way. However, from the same initial state, $I(t)$ (red curve) of system (1.3) for larger u tracks the unstable small cycle first, and then tends to the unique stable endemic equilibrium $\widetilde{E}_2$. We can observe that the rate u affects the transient dynamics of system (1.3). In Fig. 13d, for $u < 0$, $I(t)$ (red curve) of system (1.3) tracks the unstable cycle, and finally tends to the disease-free state $\widetilde{E}_0$. Moreover, $I(t)$ reaches in advance the stable disease-free state $\widetilde{E}_0$, where the corresponding $k(t) > k_{LP}$ (i.e., $\mathscr{R}_0 > \mathscr{R}_0^*$).

4 Concluding remarks

In epidemiology, the incidence rate and effective treatment are two most important covariates in determining the disease transmission and control. To explore the significance of the awareness factors, crowing effect, limited medical resources, and

intervention policies on the dynamics of emerging infectious diseases, we proposed and studied an SIRS model with nonmonotone incidence and saturated treatment. Firstly, we simplified the model into a two-dimensional system (2.2). Then we analyzed the stability and types of disease-free and endemic equilibria as well as different bifurcations of system (2.2) in detail. We found that system (2.2) could have at most two positive equilibria when $\mathscr{R}_0 < 1$. The model exhibits rich dynamics, such as bistability (a disease-free equilibrium and an endemic equilibrium), tristability (a disease-free equilibrium, an endemic equilibrium and a big limit cycle), homoclinic orbits, the coexistence of a stable homoclinic loop and an unstable limit cycle, the coexistence of two limit cycles (the inner one unstable and the outer stable), and a semi-stable limit cycle for different sets of parameters. In particular, as parameters vary, the model undergoes saddle-node bifurcation, backward bifurcation, degenerate Bogdanov-Takens bifurcation of codimension three, Hopf bifurcation and degenerated Hopf bifurcation with codimension at least two. Numerical simulations are provided to demonstrate our results.

In Pan et al. (2022), Pan *et al.* consider an SIRS model with the same nonmonotone incidence rate and a piecewise-smooth treatment rate, where the piecewise-smooth treatment rate describes the situation where the community has limited medical resources, treatment rises linearly with I until the treatment capacity is reached, after which constant treatment (i.e., the maximum treatment) is taken. Their results indicate that there exists a critical value $\widetilde{I_0}$ for the infective level I_0 at which the health care system reaches its capacity such that: **(i)** When $I_0 \geq \widetilde{I_0}$, $\mathscr{R}_0 = 1$ separates disease persistence from disease eradication. **(ii)** When $I_0 < \widetilde{I_0}$, their model can exhibit multiple endemic equilibria, periodic orbits, homoclinic orbits, Bogdanov-Takens bifurcations, and subcritical Hopf bifurcation. In this paper, for system (2.2) with saturated treatment rate, we showed that there exist three critical values r_1, r_2 and r_3 for the treatment rate r: **(i)** when $r \geq \max\{r_1, r_2\}$, the disease will disappear; **(ii)** when $r < \min\{r_1, r_3\}$, the disease will persist. Compared the results in Pan et al. (2022) with ours, system (2.2) can exhibit more complex bifurcation phenomena: degenerate Bogdanov-Takens bifurcation of codimension three and degenerated Hopf bifurcation with codimension at least two. While the system in Pan et al. (2022) can have three endemic equilibria. On the other hand, compared with no treatment in Xiao and Ruan (2007), our results can cover their ones.

The researchers in infectious disease modeling suggested and studied many different incidence functions and treatment functions according to different scenarios. Some incorporated vaccination and even boosted vaccination, which are clearly significant in mitigating and control disease epidemics. The model formulated in this paper only considers treatment control, while vaccination and its boosted version will be mechanistically modeled in our next work. In addition, for mathematical simplicity we assumed no disease-caused mortality in our model. This assumption is reasonable for most diseases with a low death probability due to infection such as flu and COVID-19. It would also be interesting to explore complicated dynamics of model (1.1) by applying other incidence rate functions such as $f(I)S = kIS/(1 + \beta I + \alpha I^2)$ (Xiao and Zhou 2006; Lu et al. 2021) and $f(I)S = kI^2S/(1 + \beta I + \alpha I^2)$ (Lu et al. 2019).

Linking transient dynamics driven by environmental change to key states in a constant environment can be insightful in epidemiological studies. The transmissi-

bility depends on nonpharmaceutical interventions, weather conditions, etc. Instead of assuming the infection rate as an explicit function of time, future research efforts should focus on the mechanistic modeling of transmissibility versus social behaviors, policies, and environmental factors.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00285-022-01787-3.

Appendix A. The Proof of Theorem 2.4

Proof **Case (I)**: firstly, let $X = x - x_*$, $Y = y - y_*$, where $y_* = qx_* + \frac{vx_*}{1+sx_*}$, then system (2.2) becomes (still denote X, Y by x, y, respectively)

$$\begin{aligned} \frac{dx}{dt} &= a_1 x + a_2 y + a_3 x^2 + a_4 xy + o(|x, y|^2), \\ \frac{dy}{dt} &= b_1 x - y + b_2 x^2 + o(|x, y|^2), \end{aligned} \tag{A1}$$

where

$$\begin{aligned} a_1 &= \frac{x_* \left(-2mpx_* + q\left(s\left(px_*^2 - 1\right) + 2px_*\right) - 1\right)}{\left(px_*^2 + 1\right)\left(s\left(px_*^2 - 1\right) + 2px_* + 1\right)}, \ a_2 = -\frac{x_*}{px_*^2 + 1}, \ a_4 = \frac{px_*^2 - 1}{\left(px_*^2 + 1\right)^2}, \\ b_1 &= q - \frac{2mpx_* + q + 1}{psx_*^2 + 2px_* - s + 1}, \ b_2 = \frac{s\left(2mpx_* + q + 1\right)}{(sx_* + 1)\left(s\left(px_*^2 - 1\right) + 2px_* + 1\right)}, \\ a_3 &= -\frac{px_*\left(px_*^2 - 3\right)\left(m\left(psx_*^2 + s - 1\right) + (q + 1)(sx_* + 1)\right)}{\left(px_*^2 + 1\right)^2\left(s\left(px_*^2 - 1\right) + 2px_* + 1\right)} \\ &\quad - \frac{s\left(2mpx_* + q + 1\right)}{(sx_* + 1)\left(s\left(px_*^2 - 1\right) + 2px_* + 1\right)} \\ &\quad + \frac{2px_*^2}{\left(px_*^2 + 1\right)^2} - \frac{1}{px_*^2 + 1}. \end{aligned}$$

For $m \neq m_0$, let $x = \frac{-psx_*^2 - 2px_* + s - 1}{2px_*(m-q) - pqsx_*^2 + qs + 1} X + \frac{x_*}{px_*^2+1} Y$, $y = X + Y$ and $\tau = \frac{1}{(px_*^2+1)(psx_*^2+2px_*-s+1)} \{-px_*^2(2m-2q+1) - p^2sx_*^4 + px_*^3(qs-2p) - x_*(2p+qs+1) + s - 1\}t$, then system (A1) becomes (still denote X, Y, τ by x, y, t, respectively)

$$\begin{aligned} \frac{dx}{dt} &= a_5 x^2 + a_6 xy + a_7 y^2 + o(|x, y|^2), \\ \frac{dy}{dt} &= y + b_3 x^2 + b_4 xy + b_5 y^2 + o(|x, y|^2), \end{aligned} \tag{A2}$$

where

$$
\begin{aligned}
a_5 &= \frac{x_*\left(px_*^2+1\right)\left(s\left(px_*^2-1\right)+2px_*+1\right)^2 S_0}{(sx_*+1)S_1^2 S_2},\\
S_0 &= mp^2 s x_*^2\left(sx_*+3\right)-p\left(m(s-1)\left(3sx_*+1\right)+q+1\right)+(-q-1)(s-1)s,\\
S_1 &= px_*^2(2m-2q+1)+2p^2x_*^3+s\left(px_*^2-1\right)\left(px_*^2-qx_*+1\right)+2px_*+x_*+1,\\
S_2 &= -2mpx_*+q\left(s\left(px_*^2-1\right)+2px_*\right)-1,
\end{aligned}
\tag{A3}
$$

we omit the expression for a_i (i=6, 7) and b_i (i=3, 4, 5) for brevity. In what follows, we will prove $a_5 < 0$ for $S_0 < 0$, $S_1 \neq 0$, $S_2 < 0$.

Step 1. We prove $S_0 < 0$. From the expression of v_0 defined in (2.17), then $0 < p < \frac{1}{x_*^2}$ and $s > \frac{2px_*+1}{1-px_*^2}$ (i.e., $s\left(px_*^2-1\right)+2px_*+1 < 0$). By direct calculation, $m+v_0-n_0 = \frac{x_*^2\phi_1}{s(px_*^2-1)+2px_*+1}$, where

$$
\phi_1 = mp^2 s x_*^2 - p\left(m(s-1)\left(2sx_*+1\right)+q+1\right)+(-q-1)(s-1)s. \tag{A4}
$$

Since $0 < \mathscr{R}_0 < 1$ (i.e., $m+v_0-n_0 > 0$), then $\phi_1 < 0$. Furthermore, do the calculation

$$
S_0-\phi_1 = mpsx_*[s\left(px_*^2-1\right)+2px_*+1] < 0,
$$

we have $S_0 < \phi_1 < 0$.

Step 2. We prove $S_1 \neq 0$. From $m \neq m_0$, it is obviously that $S_1 \neq 0$.

Step 3. We prove $S_2 < 0$. For $s > \frac{2px_*+1}{1-px_*^2}$, we have $S_2 < 0$.

In conclusion, $a_5 < 0$. According to the Theorem 7.1 in Zhang et al. (1992), the equilibrium (x_*, y_*) is a saddle-node.

Case (II)(i): we make the following transformations successively

$$
\begin{aligned}
&\widetilde{X} = x-x_*,\ \widetilde{Y} = y-y_*;\\
&\widetilde{X} = \frac{x_*}{px_*^2+1}\widetilde{X}_1,\ \widetilde{Y}_1 = \widetilde{X}_1-\widetilde{Y}_1,
\end{aligned}
$$

then (2.2) becomes (still denote $\widetilde{X}_1$, $\widetilde{Y}_1$ by x, y, respectively)

$$
\begin{aligned}
\frac{dx}{dt} &= y + c_4x^2 + c_5xy + o(|x,y|^2),\\
\frac{dy}{dt} &= d_3x^2 + d_4xy + o(|x,y|^2),
\end{aligned}
\tag{A5}
$$

we omit the expression of c_i (i=4, 5) and d_i (i=3, 4) for brevity.

By Remark 1 of Sect. 2.13 in Perko (2001), we obtain an equivalent system of (A5) in the small neighborhood of (0, 0) as follows:

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= Dx^2 + Exy + o(|x, y|^2), \end{aligned} \tag{A6}$$

where

$$D = d_3 = \frac{N_1(x_*,\ p,\ s,\ q)}{2\left(px_*^2+1\right)^2(sx_*+1)}, \quad E = d_4 + 2c_4 = \frac{N_2(x_*,\ p,\ s,\ q)}{\left(px_*^2+1\right)^2(sx_*+1)}, \tag{A7}$$

and

$$\begin{aligned} N_1(x_*,\ p,\ s,\ q) &= p^2s^2x_*^5 + psx_*^4(3p - qs) \\ &\quad + psx_*^3(-3q - 2s + 3) + x_*^2(2ps + p + s(3qs - 2q + 1)) \\ &\quad + x_*\left((q + 3)s - 3s^2 + 1\right) - s + 1, \\ N_2(x_*,\ p,\ s,\ q) &= A_1 - qsx_*A_2, \\ A_1 &= x_*\left(sx_*\left(px_*\left(sx_* + 3\right) - 2s\right) + 2p + 1\right) + s(2 - 3s) + 1) - s + 2, \\ A_2 &= psx_*^3 + 3px_*^2 - 3sx_* - 1. \end{aligned} \tag{A8}$$

We next prove $D < 0$ (i.e., $N_1(x_*,\ p,\ s,\ q) < 0$), and $E \neq 0$ (i.e., $N_2(x_*,\ p,\ s,\ q) \neq 0$) if $q \neq q_0$.

Step 1. We prove $N_1(x_*,\ p,\ s,\ q) < 0$. Let

$$\begin{aligned} \psi_1 &= 2s^2x_*\left(px_*^2 - qx_* + 1\right) - s\left(px_*^2 + 2x_* - 1\right)\left(px_*^2 - qx_* + 1\right) - px_*^2 - x_* - 1, \\ \psi_2 &= -(2p^2x_*^3 + s\left(px_*^2 - 1\right)\left(px_*^2 - qx_* + 1\right) + px_*^2 + 2px_* + x_* + 1). \end{aligned}$$

Since $0 < \mathscr{R}_0 < 1$, (i.e., $m_0 - n_0 + v_0 > 0$), we have $m_0 - n_0 + v_0 = \frac{1}{2}\psi_1 > 0$, which implies $\psi_1 > 0$. Since $m_0 - q > 0$, we have $m_0 - q = \frac{\psi_2}{2px_*^2} > 0$, which implies $\psi_2 > 0$. By direct calculation, we have $N_1(x_*,\ p,\ s,\ q) + \psi_1 = sx_*(-2pqx_*^2 - \psi_2) < 0$, since $\psi_2 > 0$. Then we can get $N_1(x_*,\ p,\ s,\ q) < 0$, since $\psi_1 > 0$. Thus $D < 0$.

Step 2. We prove $N_2(x_*,\ p,\ s,\ q) \neq 0$ if $q \neq q_0$. From the expression of v_0 defined in (2.17), then $0 < p < \frac{1}{x_*^2}$ and $s > \frac{2px_*+1}{1-px_*^2}$. Since $s > \frac{2px_*+1}{1-px_*^2}$, we can easily get $A_2 < 0$. Thus, we conclude that $N_2(x_*,\ p,\ s,\ q) \neq 0$ if $q \neq q_0$.

Hence, by the result in Perko (2001), we know that E_* is a cusp of codimension 2 if $q \neq q_0$.

Case (II)(ii): when $m = m_0$, the conditions $0 < \mathscr{R}_0 < 1$, $n > \frac{1+q+v+ms}{s}$, (2.4) and (2.17) are equivalent to $(x_*,\ p,\ s,\ q) \in \Omega_1 \cup \Omega_2$, where

$$\begin{aligned}
\Omega_1 &:= \{(x_*,\ p,\ s,\ q) | x_* > 0,\ 0 < p < \frac{1}{x_*^2},\ s_1 < s < s_2,\ 0 < q < q_1\},\\
\Omega_2 &:= \{(x_*,\ p,\ s,\ q) | x_* > 0,\ 0 < p < \frac{1}{x_*^2},\ s \geq s_2,\ q_2 < q < q_1\},
\end{aligned} \tag{A9}$$

and

$$\begin{aligned}
s_1 &= \frac{2p^2x_*^3 + px_*^2 + 2px_* + x_* + 1}{1 - p^2x_*^4},\\
q_1 &= \frac{s\left(p^2x_*^4 - 1\right) + x_*\left(p\left(2px_*^2 + x_* + 2\right) + 1\right) + 1}{sx_*\left(px_*^2 - 1\right)},\\
s_2 &= \frac{2p^2x_*^3 + 3px_*^2 + 2px_* + x_* + 1}{1 - p^2x_*^4},\\
q_2 &= \frac{s\left(p^2x_*^4 - 1\right) + (2px_* + 1)\left(px_*^2 + x_* + 1\right)}{sx_*\left(px_*^2 - 1\right)}.
\end{aligned} \tag{A10}$$

Obviously, $\Omega_1 \cup \Omega_2 = \Omega_{11} \cup \Omega_{21} \cup \Omega_{12} \cup \Omega_{22}$, where

$$\begin{aligned}
\Omega_{11} &:= \left\{(x_*,\ p,\ s,\ q) | x_* \geq \frac{2}{5},\ 0 < p < \frac{1}{x_*^2},\ s_1 < s < s_2,\ 0 < q < q_1\right\},\\
\Omega_{21} &:= \left\{(x_*,\ p,\ s,\ q) | x_* \geq \frac{2}{5},\ 0 < p < \frac{1}{x_*^2},\ s \geq s_2,\ q_2 < q < q_1\right\},\\
\Omega_{12} &:= \left\{(x_*,\ p,\ s,\ q) | 0 < x_* < \frac{2}{5},\ 0 < p < \frac{1}{x_*^2},\ s_1 < s < s_2,\ 0 < q < q_1\right\},\\
\Omega_{22} &:= \left\{(x_*,\ p,\ s,\ q) | 0 < x_* < \frac{2}{5},\ 0 < p < \frac{1}{x_*^2},\ s \geq s_2,\ q_2 < q < q_1\right\}.
\end{aligned} \tag{A11}$$

We next prove the conclusion in Case (II)(ii) in the following steps.

Step 1. We prove when $(x_*,\ p,\ s,\ q) \in \Omega_{11} \cup \Omega_{21}$ then $N_2(x_*,\ p,\ s,\ q) < 0$, that is to say if $N_2(x_*,\ p,\ s,\ q) = 0$, then $(x_*,\ p,\ s,\ q)$ must lies in $\Omega_{12} \cup \Omega_{22}$. In the following, we take $N_2(x_*,\ p,\ s,\ q) < 0$ in Ω_{11} for an example.

Step 1.1. We check $N_2(x_*,\ p,\ s,\ 0) < 0$ and $N_2(x_*,\ p,\ s,\ q_1) < 0$ when $x_* \geq \frac{2}{5}$, $0 < p < \frac{1}{x_*^2}$ and $s_1 < s < s_2$.

Since

$$
\begin{aligned}
&N_2(x_*,\ p,\ s,\ 0) = \Phi_1(x_*,\ p,\ s),\ N_2(x_*,\ p,\ s,\ q_1) = \frac{1}{px_*^2-1}\Psi_1(x_*,\ p,\ s),\\
&\Phi_1(x_*,\ p,\ s_1) = -\frac{1}{\left(px_*^2-1\right)^2\left(px_*^2+1\right)}\Phi_2(x_*,\ p),\\
&\Phi_2(x_*,0) = 2x_*^3+3x_*^2+x_*-1,\ \ \Phi_2\left(x_*,\frac{1}{x_*^2}\right) = \frac{2\left(x_*^2+2x_*+4\right)^2}{x_*},\\
&L_0(\overline{p}) := (1+\overline{p})^5\cdot\Phi_2\left(x_*,\frac{1}{x_*^2(1+\overline{p})}\right)\\
&\quad = \frac{1}{x_*}(\overline{p}^5x_*(2x_*^3+3x_*^2+x_*-1)\\
&\quad\ +\overline{p}^4(10x_*^4+21x_*^3+18x_*^2+3x_*+2)+\overline{p}^3(20x_*^4\\
&\quad\ +53x_*^3+73x_*^2+40x_*+16)+\overline{p}^2(20x_*^4+63x_*^3+120x_*^2+100x_*+48)\\
&\quad\ +2\overline{p}(5x_*^4+18x_*^3+44x_*^2+48x_*+32)+2(x_*^2+2x_*+4)^2),
\end{aligned}
\tag{A12}
$$

where we omit the expression of $\Phi_1(x_*,\ p,\ s)$ and $\Phi_2(x_*,\ p)$ for brevity. Obviously, $\Phi_2(x_*,0)>0$, $\Phi_2(x_*,\frac{1}{x_*^2})>0$ and $L_0(\overline{p})=0$ has no positive root when $x_*\geq\frac{2}{5}$. By Lemma 3.1 of Yang (1999), $\Phi_2(x_*,p)=0$ with p as a variable has no root in the interval $(0,\frac{1}{x_*^2})$ if parameter $x_*\geq\frac{2}{5}$. Since $\Phi_2(x_*,0)>0$, we immediately obtain $\Phi_2(x_*,p)>0$, thus $\Phi_1(x_*,\ p,\ s_1)<0$. Similarly, we can get $\Phi_1(x_*,\ p,\ s_2)<0$. Furthermore, let

$$
\begin{aligned}
L_1(\overline{s}) &:= (1+\overline{s})^2\cdot\Phi_1\left(x_*,\ p,\ \frac{s_2+s_1\cdot\overline{s}}{1+\overline{s}}\right)\\
&= -\frac{1}{\left(px_*^2-1\right)^2\left(px_*^2+1\right)}h_1(x_*,\ p,\ \overline{s}),
\end{aligned}
\tag{A13}
$$

where we omit the expression of $h_1(x_*,\ p,\ \overline{s})$, then we obtain

$$
\begin{aligned}
&h_1(x_*,\ 0,\ \overline{s}) = (\overline{s}+1)^2\left(2x_*^3+3x_*^2+x_*-1\right),\\
&h_1\left(x_*,\ \frac{1}{x_*^2},\ \overline{s}\right) = \frac{2\left(\overline{s}\left(x_*^2+2x_*+4\right)+(x_*+2)^2\right)^2}{x_*},\\
&L_2(\hat{p}) := (1+\hat{p})^5\cdot h_1\left(x_*,\ \frac{1}{x_*^2(1+\hat{p})},\ \overline{s}\right) = h_2(x_*,\ \hat{p},\ \overline{s}).
\end{aligned}
$$

Since $x_*\geq\frac{2}{5}$, $\hat{p}>0$ and $\overline{s}>0$, it is obvious that $h_1(x_*,\ 0,\ \overline{s})>0$, $h_1(x_*,\ \frac{1}{x_*^2},\ \overline{s})>0$ and $h_2(x_*,\ \hat{p},\ \overline{s})>0$, which is given in supplementary materials. Then by Lemma 3.1 of Yang (1999), $h_1(x_*,\ p,\ \overline{s})=0$ with p as a variable has no root in the interval $(0,\frac{1}{x_*^2})$ if $x_*\geq\frac{2}{5}$ and $\overline{s}>0$. Thus $L_1(\overline{s})=0$ has no root in $(0,+\infty)$ for parameters $x_*\geq\frac{2}{5}$ and $0<p<\frac{1}{x_*^2}$. Then, $\Phi_1(x_*,\ p,\ s)=0$ has no root with s as a variable in (s_1,s_2) and parameters $x_*>\frac{2}{5}$, $0<p<\frac{1}{x_*^2}$. Since $\Phi_1(x_*,\ p,\ s_1)<0$, which

implies $\Phi_1(x_*,\ p,\ s) < 0$. We get $N_2(x_*,\ p,\ s,\ 0) < 0$ in Ω_{11}. Similarly, by using the same steps to $\Psi_1(x_*,\ p,\ s)$ as $\Phi_1(x_*,\ p,\ s)$, we can get $N_2(x_*,\ p,\ s,\ q_1) < 0$ in Ω_{11}.

Step 1.2. We regard q as a variable in $N_2(x_*,\ p,\ s,\ q)$ with parameters $x_* \geq \frac{2}{5}$, $0 < p < \frac{1}{x_*^2}$ and $s_1 < s < s_2$. Since $N_2(x_*,\ p,\ s,\ 0) < 0$, $N_2(x_*,\ p,\ s,\ q_1) < 0$ when $x_* \geq \frac{2}{5}$, $0 < p < \frac{1}{x_*^2}$ and $s_1 < s < s_2$, then let

$$\begin{aligned} g(\tilde{q}) &:= (1+\tilde{q}) \cdot N_2\left(x_*,\ p, s,\ \frac{q_1}{1+\tilde{q}}\right) = \frac{1}{px_*^2-1}\tilde{h}_1(x_*,\ p,\ s,\ \tilde{q}), \\ g_1(\tilde{s}) &:= (1+\tilde{s})^2 \cdot \tilde{h}_1\left(x_*,\ p,\ \frac{s_2+s_1\cdot\tilde{s}}{1+\tilde{s}},\ \tilde{q}\right) = \frac{1}{1-p^2x_*^4}\tilde{h}_2(x_*,\ p,\ \tilde{s},\ \tilde{q}), \\ g_2(\tilde{p}) &:= (1+\tilde{p})^5 \cdot \tilde{h}_2\left(x_*,\ \frac{1}{x_*^2(1+\tilde{p})},\tilde{s},\ \tilde{q}\right) = -\frac{1}{x_*}\tilde{h}_3(x_*,\ \tilde{p},\ \tilde{s},\ \tilde{q}). \end{aligned} \tag{A14}$$

We can check $\tilde{h}_1(x_*,\ p,\ s_1,\ \tilde{q}) < 0$ and $\tilde{h}_1(x_*,\ p,\ s_2,\ \tilde{q}) < 0$ when $x_* \geq \frac{2}{5}, 0 < p < \frac{1}{x_*^2}$ and $\tilde{q} > 0$. $\tilde{h}_2(x_*,\ 0,\ \tilde{s},\ \tilde{q}) < 0$ and $\tilde{h}_2(x_*,\ \frac{1}{x_*^2},\ \tilde{s},\ \tilde{q}) < 0$ when $x_* \geq \frac{2}{5}, \tilde{s} > 0$ and $\tilde{q} > 0$. We make the transformation $x_0 = x_* - \frac{2}{5}$ to transform the problem of determining the sign of $\tilde{h}_3(x_*,\ \tilde{p},\ \tilde{s},\ \tilde{q})$ with x_* as variable in the interval $(\frac{2}{5}, +\infty)$ to the issue of determining $\tilde{h}_4(x_0,\ \tilde{p},\ \tilde{s},\ \tilde{q})$ with x_0 as variable in the interval $(0, +\infty)$ and parameters $\tilde{p} > 0, \tilde{s} > 0$ and $\tilde{q} > 0$, where $\tilde{h}_4(x_0,\ \tilde{p},\ \tilde{s},\ \tilde{q}) = h_3(x_0 + \frac{2}{5},\ \tilde{p},\ \tilde{s},\ \tilde{q})$, and $\tilde{h}_4(x_0,\ \tilde{p},\ \tilde{s},\ \tilde{q}) > 0$ is given in supplementary materials. Then $h_3(x_*,\ \tilde{p},\ \tilde{s},\ \tilde{q}) > 0$ when $x_* \geq \frac{2}{5}$, $\tilde{p} > 0$, $\tilde{s} > 0$ and $\tilde{q} > 0$, thus the sign of $g_2(\tilde{p})$ does not change. We immediately obtain the sign of $g_1(\tilde{s})$ does not change, which implies the sign of $g_(\tilde{q})$ does not change. Hence, $N_2(x_*,\ p,\ s,\ q) = 0$ has no root with q as a variable in $(0, q_1)$. Since $N_2(x_*,\ p,\ s,\ 0) < 0$, then $N_2(x_*,\ p,\ s,\ q) < 0$ in Ω_{11}.

Similarly, we can prove $N_2(x_*,\ p,\ s,\ q) < 0$ in Ω_{21}. In conclusion, when $N_2(x_*,\ p,\ s,\ q) = 0$ parameters must satisfy $q = q_0$ and $(x_*,\ p,\ s,\ q_0) \in \Omega_{12} \cup \Omega_{22}$.

Step 2. When $q = q_0$ and $(x_*,\ p,\ s,\ q_0) \in \Omega_{12} \cup \Omega_{22}$, We make the following transformations successively

$$\begin{aligned} &\overline{X} = x - x_*,\ \overline{Y} = y - y_*; \\ &\overline{X}_1 = \overline{X},\ \overline{Y}_1 = \frac{d\overline{X}}{dt}; \\ &t = (1 - \hat{c}_{02}\overline{X}_1)\bar{\tau}_1,\ \overline{X}_2 = \overline{X}_1,\ \overline{Y}_2 = \overline{Y}_1(1 - \hat{c}_{02}\overline{X}_1); \\ &\overline{X}_3 = -\overline{X}_2,\ \overline{Y}_3 = \frac{-\overline{Y}_2}{\sqrt{-\hat{d}_{20}}},\ \bar{\tau}_2 = \sqrt{-\hat{d}_{20}}\bar{\tau}_1, \end{aligned} \tag{A15}$$

then system (2.2) can be rewritten as (for simplicity, we denote $\overline{X}_3, \overline{Y}_3, \bar{\tau}_2$ by x, y, t, respectively)

$$\begin{aligned}\frac{dx}{dt} &= y\\ \frac{dy}{dt} &= x^2 + \hat{e}_{30}x^3 + \hat{e}_{21}x^2y + \hat{e}_{12}xy^2 + \hat{e}_{40}x^4 + \hat{e}_{31}x^3y + \hat{e}_{22}x^2y^2 + o(|x,\, y|^4),\end{aligned} \tag{A16}$$

where $\hat{c}_{02} = \frac{1-px_*^2}{px_*^3+x_*}$, $\hat{d}_{20} = \frac{p^2sx_*^5+3p^2x_*^4-4psx_*^3-2(2p+1)x_*^2+(3s-4)x_*+1}{2x_*(px_*^2+1)(sx_*(px_*^2-3)+3px_*^2-1)} < 0$. In fact, the expression of $\hat{d}_{20}$ is the same as D when $q = q_0$, where D is defined in (A7). The expression of $\hat{e}_{30}$, $\hat{e}_{21}$, $\hat{e}_{12}$, $\hat{e}_{40}$, $\hat{e}_{31}$ and $\hat{e}_{22}$ are given in supplementary materials.

By the Proposition 5.3 in Lamontagne et al. (2008) (see also Lemma 2 in Huang et al. (2016)), we obtain the equivalent system of (A16) as follows:

$$\begin{aligned}\frac{dx}{dt} &= y,\\ \frac{dy}{dt} &= x^2 + Mx^3y + o(|x,\, y|^4),\end{aligned} \tag{A17}$$

where

$$M = \hat{e}_{31} - \hat{e}_{30}\hat{e}_{21} = \frac{-\sqrt{2}k_{11}}{x_*\left(px_*^2+1\right)^2(sx_*+1)^2k_{12}k_{13}\sqrt{\frac{-k_{12}}{2x_*(px_*^2+1)k_{13}}}}, \tag{A18}$$

and

$$\begin{aligned}k_{11} &= p^4x_*^6(2s^4x_*^4 + s^2(16s+7)x_*^3 + 2s(29s+12)x_*^2 + (96s+9)x_* + 36)\\ &\quad + p^3x_*^4(7s^4x_*^5 + 24s^3x_*^4 - 4s^2(8s+5)x_*^3 - 2(94s^2+24s+3)x_*^2\\ &\quad - (208s+51)x_* - 108) - p^2x_*^2(s^3(29s+2)x_*^5\\ &\quad + (-88s^2+52s+42)x_*^2 + 4s(7s^3+13s^2+6s+2)x_*^4\\ &\quad + (-16s^3+78s^2+54s+6)x_*^3 + (61-128s)x_* - 12)\\ &\quad + p(-2s^3(3s+4)x_*^6 - 3s^2(s^2+8s+2)x_*^5\\ &\quad + 4s(2s^2-2s+1)x_*^4 + (4s^2+40s+6)x_*^3\\ &\quad + (-52s^2+64s+28)x_*^2 + (31-16s)x_* - 4)\\ &\quad - s(x_*+2)(3s^3x_*^2(2x_*+1) - 2s^2x_*^2(2x_*+1) + s(-6x_*^2-4x_*+1)+2),\\ k_{12} &= p^2sx_*^5 + 3p^2x_*^4 - 4psx_*^3 - 2(2p+1)x_*^2 + (3s-4)x_* + 1,\\ k_{13} &= sx_*\left(px_*^2-3\right) + 3px_*^2 - 1.\end{aligned} \tag{A19}$$

Step 3. We will prove $M < 0$ in $\Omega_{21} \cup \Omega_{22}$. We already proved that $D < 0$, which is defined in (A7). Furthermore, when $q = q_0$ then $D = \frac{k_{12}}{2x_*(px_*^2+1)k_{13}}$, which implies $k_{12} \cdot k_{13} < 0$. We then prove $k_{11} < 0$ in $\Omega_{21} \cup \Omega_{22}$. We regard x_* and p as parameters in k_{11}, and we can check $k_{11}(x_*,\ p,\ s_1) < 0$ and $k_{11}(x_*,\ p,\ s_2) < 0$ when $0 < x_* < \frac{2}{5}$ and $0 < p < \frac{1}{x_*^2}$. Let

$$
\begin{aligned}
P_1(\bar{\bar{s}}) &:= (1+\bar{\bar{s}})^4 \cdot k_{11}\left(x_*,\ p,\ \frac{s_2+s_1\cdot\bar{\bar{s}}}{1+\bar{\bar{s}}}\right) = \frac{1}{\left(p^2{x_*}^4-1\right)^4} f_1(x_*,\ p,\ \bar{\bar{s}}),\\
P_2(\bar{\bar{p}}) &:= (1+\bar{\bar{p}})^{12} \cdot f_1\left(x_*,\ \frac{1}{x_*^2(1+\bar{\bar{p}})},\ \bar{\bar{s}}\right) = -\frac{1}{x_*^2} f_2(x_*,\ \bar{\bar{p}},\ \bar{\bar{s}}),\\
P_3(\widetilde{x}_*) &:= (1+\widetilde{x}_*)^{10} \cdot f_2\left(\frac{2}{5(1+\widetilde{x}_*)},\ \bar{\bar{p}},\ \bar{\bar{s}}\right).
\end{aligned}
\tag{A20}
$$

We check $f_1(x_*,\ 0,\ \bar{\bar{s}}) < 0$ and $f_1(x_*,\ \frac{1}{x_*^2},\ \bar{\bar{s}}) < 0$ when $0 < x_* < \frac{2}{5}$ and $\bar{\bar{s}} > 0$, $f_2(0,\ \bar{\bar{p}},\ \bar{\bar{s}}) > 0$ and $f_2(\frac{2}{5},\ \bar{\bar{p}},\ \bar{\bar{s}}) > 0$ when $\bar{\bar{p}} > 0$ and $\bar{\bar{s}} > 0$. the expression of $P_3(\widetilde{x}_*)$ given in supplementary materials and obviously the sign of $P_3(\widetilde{x}_*)$ does not change in the interval $(0, +\infty)$, with parameters $\bar{\bar{p}} > 0$ and $\bar{\bar{s}} > 0$. Thus $f_2(x_*,\ \bar{\bar{p}},\ \bar{\bar{s}}) > 0$, implying the sign of $P_2(\bar{\bar{p}})$ does not change in the interval $(0, +\infty)$, with parameters $0 < x_* < \frac{2}{5}$ and $\bar{\bar{s}} > 0$. Thus, we can get the sign of $f_1(x_*,\ p,\ \bar{\bar{s}})$ does not change in the interval $(0, +\infty)$, with parameters $0 < x_* < \frac{2}{5}$ and $0 < p < \frac{1}{x_*^2}$. Then we can get the sign of $P_1(\bar{\bar{s}})$ does not change in the interval $(0, +\infty)$, with parameters $0 < x_* < \frac{2}{5}$ and $0 < p < \frac{1}{x_*^2}$. Then $k_{11}(x_*,\ p,\ s) < 0$ in (s_1, s_2) with parameters $0 < x_* < \frac{2}{5}$ and $0 < p < \frac{1}{x_*^2}$. Hence $M < 0$ in Ω_{21}.

In Ω_{22}, we regard x_* and p as parameters in k_{11}, and make the transformation $s = s_0 + s_2$ to transform the problem of determining the sign of k_{11} in the interval $(s_2, +\infty)$ with parameters $0 < x_* < \frac{2}{5}$ and $0 < p < \frac{1}{x_*^2}$ to the issue of determining the sign of $f_3(s_0)$ in the interval $(0, +\infty)$, with parameters $0 < x_* < \frac{2}{5}$ and $0 < p < \frac{1}{x_*^2}$, where $f_3(s_0) = k_{11}(x_*, p, s_0 + s_2)$. Then by the similarly steps of k_{11} in Ω_{21}, we can get $k_{11} < 0$ in Ω_{22}. Hence $M < 0$ in Ω_{22}.

In conclusion, we have $M < 0$ in $\Omega_{21} \cup \Omega_{22}$, then the equilibrium E_* is a cusp of codimension three (Dumortier et al. 1987). □

References

Arumugam R, Guichard F, Lutscher F (2020) Persistence and extinction dynamics driven by the rate of environmental change in a predator-prey metacommunity. Theo Eco 13:629–643

Arumugam R, Lutscher F, Guichard F (2021) Tracking unstable states: Ecosystem dynamics in a changing world. Oikos 130:525–540

Capasso V, Crosso E, Serio G (1977) I modelli matematici nella indagine epidemiologica. I Applicazione all'epidemia di colera verificatasi in Bari nel 1973. Ann Sclavo 19:193–208

Capasso V, Serio G (1978) A generalization of the Kermack-Mckendrick deterministic epidemic model. Math Biosci 42:43–61

Chow S, Li C, Wang D (1994) Normal Forms and Bifurcation of Planar Vector Fields. Cambridge University Press, Cambridge

Dumortier F, Roussarie R, Sotomayor J (1987) Generic 3-parameter families of vector fields on the plane, unfolding a singularity with nilpotent linear part. The cusp case of codimension 3. Ergo The Dyn Sys 7(3):375–413

Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180:29–48

Ghosh JK, Majumdar P, Ghosh U (2021) Qualitative analysis and optimal control of an SIR model with logistic growth, non-monotonic incidence and saturated treatment. Mathe Model Nat Phenom 16:13

Ghosh U, Sarkar S, Ghosh J (2021) Three dimensional epidemic model with non-monotonic incidence and saturated treatment: a case study of SARS infection of Hong Kong 2003 Scenario. https://doi.org/10.21203/rs.3.rs-310291/v1
Huang J, Liu S, Ruan S, Zhang X (2016) Bogdanov-Takens bifurcation of codimension 3 in a predator-prey model with constant-yield predator harvesting, Commun. Pur. Appl Anal 15:1041–1055
Jana S, Nandi S, Kar T (2016) Complex dynamics of an SIR epidemic model with saturated incidence rate and treatment. Acta Biotheor 64(1):1–20
Kermack W, McKendrick A (1927) A contribution to the mathematical theory of epidemics. Proc Roal Soc Lond 115:700–721
Li C, Li J, Ma Z (2015) Codimension 3 B-T bifurcations in an epidemic model with a nonlinear incidence. Discrete Cont Dyn Ser B 20(4):1107–1116
Lamontagne Y, Coutu C, Rousseau C (2008) Bifurcation analysis of a predator-prey system with generalised Holling type III functional response. J Dyn Differ Equ 20(3):535–571
Lu M, Huang J, Ruan S, Yu P (2021) Global dynamics of a susceptible-infectious-recovered epidemic model with a generalized nonmonotone incidence rate. J Dyn Differ Equ 33(4):1625–1661
Lu M, Huang J, Ruan S, Yu P (2019) Bifurcation analysis of a SIRS epidemic model with a generalized nonmonotone and saturated incidence rate. J Differ Equ 267(3):1859–1898
Liu X, Yang L (2012) Stability analysis of an SEIQV epidemic model with saturated incidence rate. Nonlinear Anal Real 13(6):2671–2679
Pan Q, Huang J, Huang Q (2022) Global dynamics and bifurcations in a SIRS epidemic model with a nonmonotone incidence rate and a piecewise-smooth treatment rate. Discrete Cont Dyn Ser B 27(7):3533–3561
Perko L (2001) Differential Equations and Dynamical Systems, 3rd edn. Springer, New York
Wang W (2006) Backward bifurcation of an epidemic model with treatment. Math Biosci 201(1–2):58–71
Wang W, Ruan S (2004) Bifurcation in an epidemic model with constant removal rate of infectives. J Math Anal Appl 291:775–793
Xiang C, Huang J, Wang H (2022) Linking bifurcation analysis of Holling-Tanner model with generalist predator to a changing environment. Stud Appl Math 149(1):124–163
Xiao D, Ruan S (2007) Global analysis of a epidemic model with nonmonotone incidence rate. Math Biosci 208:419–429
Xiao D, Zhou Y (2006) Qualitative analysis of an epidemic model. Can Appl Math Q 14(4):469–492
Yang L (1999) Recent advances on determining the number of real roots of parametric polynomials. J Symb Comput 28(1–2):225–242
Yang Q, Jiang D, Shi N, Ji C (2012) The ergodicity and extinction of stochastically perturbed SIR and SEIR epidemic models with saturated incidence. J Math Anal Appl 388(1):248–271
Zhou L, Meng F (2012) Dynamics of an SIR epidemic model with limited medical resources revisited. Nonlinear Anal Real 13(1):312–324
Zhang Z, Suo Y (2010) Qualitative analysis of a SIR epidemic model with saturated treatment rate. J Appl Math Comput 34:177–194
Zhang Z, Ding T, Huang W, Dong Z (1992) Qualitative Theory of Differential Equations, Transl. Math. Monogr, vol 101. Amer. Math. Soc, Providence, RI
Zhang X, Liu X (2008) Backward bifurcation of an epidemic model with saturated treatment function. J Math Anal Appl 348:433–443
Zhou T, Zhang W, Lu Q (2014) Bifurcation analysis of an SIS epidemic model with saturated incidence rate and saturated treatment function. Appl Math Comput 226(1):288–305