Spatial Patterns in a Discrete-Time SIS Patch Model

November 2007

L.J.S. Allen[1], Y. Lou[2], and A.L. Nevai[3†]

Abstract

How do spatial heterogeneity, habitat connectivity, and different movement rates among subpopulations combine to influence the observed spatial patterns of an infectious disease? To find out, we formulated and analyzed a discrete-time SIS patch model. Patch differences in local disease transmission and recovery rates characterize whether patches are low-risk or high-risk, and these differences collectively determine whether the spatial domain, or habitat, is low-risk or high-risk. In low-risk habitats, the disease persists only when the mobility of infected individuals lies below some threshold value, but for high-risk habitats, the disease always persists. When the disease does persist, then there exists an endemic equilibrium (EE) which is unique and positive everywhere. This EE tends to a spatially inhomogeneous disease-free equilibrium (DFE) as the mobility of susceptible individuals tends to zero. The limiting DFE is nonempty on all low-risk patches and it is empty on at least one high-risk patch. Sufficient conditions for the limiting DFE to be empty on other high-risk patches are given in terms of disease transmission and recovery rates, habitat connectivity, and the infected movement rate. These conditions are also illustrated using numerical examples.

Keywords: Spatial heterogeneity, Dispersal, Habitat connectivity, Basic reproduction number, Disease-free equilibrium, Endemic equilibrium

Mathematics Subject Classification (2000): 92D30, 92D40, 92D50, 91D25

[1]Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409-1042.
[2]Department of Mathematics, The Ohio State University, Columbus, OH 43210.
[3]Mathematical Biosciences Institute, The Ohio State University, Columbus, OH 43210.
[†]Corresponding author: anevai@mbi.osu.edu.

1 Introduction

Recently, the influence of spatial heterogeneity on infectious disease outcomes has been explored for spatially-discrete habitats using continuous-time patch models [7, 14, 20] and for spatially-continuously habitats using both reaction-diffusion models [21, 26, 27] and integro-differential equation models [24]. One of the main drawbacks of any deterministic continuous-time model is that the multiple processes which combine to influence population dynamics (such as growth, infection, and dispersal) must operate together at all times. However, discrete-time models suffer no such limitation, and it is natural when using them to break time steps up into distinct stages so that each process occurs separately. Such a modeling approach has proved successful in studying the dynamics of many spatially-structured populations (e.g., [9, 17, 30, 31]), including those subject to the influence of infectious diseases [5, 10, 11].

In this paper, we model infection and movement among n patches over discrete time intervals. We assume that susceptible individuals (S) become infected (and infective, I) but do not develop immunity. Eventually, any individual will recover, and upon doing so they immediately become susceptible again. Thus, there are two compartments (susceptible and infected) in each patch and the dynamics of this population is described by a system of $4n$ difference equations which together constitute our discrete-time SIS epidemic model. We do not consider the effects of population growth and other changes in demography, such as disease-related deaths. In each time step, we assume that infection and recovery occur first (stage 1) and that individuals subsequently disperse among the n patches in the habitat (stage 2). In the first stage, individuals within a patch either contract the disease or recover from it based solely on local conditions within the patch. This produces an intermediate number of susceptible and infected individuals within each patch, who are then free to move between patches according to susceptible and infected dispersal matrices. These matrices need not have the same structure so long as every patch is accessible from every other patch. Characterizing different patches as low-risk or high-risk, and the habitat as a whole as low-risk or high-risk, allows us to determine various conditions under which the disease can invade or persist, and to describe the different equilibrium patterns that can arise.

In this investigation, we are particularly interested in studying equilibrium patterns of the discrete-time epidemic patch model as the susceptible dispersal rate diminishes to zero. This study therefore complements our prior research on a pair of related continuous-time epidemic models: one a patch model [1] and the other a reaction-diffusion model [2]. Here, we show that the discrete-time formulation of the disease model can exhibit dynamics and equilibrium patterns that are quite similar to those that can occur for its two continuous-time cousins.

When $A = (a_{jk})$ is a matrix and $u = (u_j)$ is a vector, we will use the notation $(Au)_j$ to denote the j-th entry of the vector Au, i.e., $(Au)_j = \sum_{k\in\Omega} a_{jk}u_k$. Let 1_n denote the vector $[1, 1, \ldots, 1]^t$ of n ones and $\mathcal{I}_n$ the $n \times n$ identity matrix.

1.1 Infection and recovery

Let $S_j = S_j(t)$ and $I_j = I_j(t)$ denote the number of susceptible and infected individuals in patch $j \in \Omega = \{1, 2, \ldots, n\}$ at time $t \in \mathfrak{T} = \{0, \Delta t, 2\Delta t, \ldots\}$. We assume that

(A1) $S_j(0) \geq 0$ and $I_j(0) \geq 0$ for $j \in \Omega$, and $\sum_{j\in\Omega} I_j(0) > 0$.

In the time interval $[t, t+\Delta t]$, let $\hat{S}_j = \hat{S}_j(t)$ and $\hat{I}_j = \hat{I}_j(t)$ be the number of susceptible and infected individuals in patch j after infection and recovery but before dispersal. We assume that

$$\begin{aligned} \hat{S}_j &= (1-\Theta_j)S_j + \gamma_j I_j, \qquad j \in \Omega, \\ \hat{I}_j &= \Theta_j S_j + (1-\gamma_j)I_j, \qquad j \in \Omega. \end{aligned} \tag{1.1}$$

Here, $\Theta_j \in [0,1]$ is the probability that a susceptible individual in patch j contracts the disease in time Δt, and $\gamma_j \in [0,1]$ is the probability that an infected individual in patch j recovers from the disease in time Δt. We assume that disease transmission is frequency-dependent, i.e.,

$$\Theta_j = \Theta_j\left(\frac{\beta_j I_j}{S_j + I_j}\right) \tag{1.2}$$

for some patch-dependent transmission parameter $\beta_j \geq 0$ and function $\Theta_j(x)$. Moreover, we assume that

(A2) $\Theta_j(x)$ is twice continuously differentiable on $[0, \beta_j]$ and satisfies

(i) $\Theta_j(0) = 0$ and $\Theta_j'(0) = 1$,
(ii) $0 \leq \Theta_j'(x) \leq \Theta_j(x)/x$ for $0 < x \leq \beta_j$,
(iii) $\Theta_j(x) \leq x$ for $0 \leq x \leq \beta_j$,
(iv) $\Theta_j(\beta_j) \leq 1$.

We remark that property (iii) actually follows directly from properties (i) and (ii), however, because we use it repeatedly in later arguments, we list it separately.

Particular functional forms for Θ_j satisfying (A2) appear in the literature. For example, if the number of infective encounters that each susceptible individual has in patch j follows a Poisson distribution then

$$\Theta_j = 1 - \exp\left(-\frac{\beta_j I_j}{S_j + I_j}\right), \tag{1.3}$$

(e.g., [3, 4, 18]). Letting $x_j = \beta_j I_j/(S_j + I_j)$ represent the average number of such encounters in patch j, then $\sigma_j(i) = (e^{-x_j}x_j^i)/i!$ is the probability of $i \in \{0, 1, 2, \dots\}$ such encounters. Since successful encounters result in infection, the probability that a susceptible individual contracts the disease is simply the probability of having at least one infective encounter, or $1 - \sigma_j(0) = 1 - e^{-x_j} = \Theta_j$. A second functional form for Θ_j satisfying (A2) is

$$\Theta_j = 1 - (1 - \overline{\alpha}_j)^{z_j},$$

where $\overline{\alpha}_j$ is the probability that a susceptible individual contracts the disease after a single infective encounter in patch j and $z_j = \underline{\alpha}_j I_j/(S_j + I_j)$ is the expected number of infective encounters [22]. Formally, this case corresponds to the previous one with $\beta_j = \ln(1 - \overline{\alpha}_j)^{-\underline{\alpha}_j}$. A third functional form for Θ_j satisfying (A2) is

$$\Theta_j = 1 - \left[1 - \frac{1}{m_j}\left(\frac{\beta_j I_j}{S_j + I_j}\right)\right]^{m_j},$$

where an individual in patch j has at most $m_j \in \{1, 2, 3, \dots\}$ infective encounters in each time step and $\beta_j \leq m_j$. In this case, $\Theta_j(x) = 1 - (1 - x/m_j)^{m_j}$, and in particular, if $m_j = 1$ then $\Theta_j(x) = x$. As $m_j \to \infty$ then the probability that a susceptible individual has at least one infective encounter is again given by (1.3).

We say that in a *low-risk patch* disease transmission occurs at a lower rate than disease recovery ($\beta_j < \gamma_j$) and that in a *high-risk patch* the reverse situation holds ($\beta_j > \gamma_j$). Let

$$H^- = \{j \in \Omega : \beta_j < \gamma_j\} \quad \text{and} \quad H^+ = \{j \in \Omega : \beta_j > \gamma_j\}$$

denote the set of these low-risk and high-risk patches, respectively. We assume that

(A3) H^- and H^+ are nonempty and $H^- \cup H^+ = \Omega$.

We define the *virulence* of a patch to be $\nu_j = 1 + \beta_j - \gamma_j$. Thus, the virulence of a low-risk patch is less than one and the virulence of a high-risk patch is greater than one. We define a *sanctuary* to be a patch with zero virulence, i.e., a patch in which the disease can never be transmitted ($\beta_j = 0$) and recovery in one time step is always certain ($\gamma_j = 1$). We assume that the habitat has no sanctuaries, that is

(A4) $\nu_j > 0$ for every $j \in \Omega$.

We define the *activity* of a patch to be $\beta_j + \gamma_j$. We will sometimes assume that no patch has activity greater than unity, i.e.,

(B1) $\beta_j + \gamma_j \leq 1$ for every $j \in \Omega$.

Finally, let

$$\mathcal{B} = \frac{n}{\sum_{j\in\Omega} 1/\nu_j} \tag{1.4}$$

be the harmonic mean of the patch virulences. Since there are no sanctuaries, $\mathcal{B}$ is well-defined. We say that Ω is a *low-risk habitat* if $\mathcal{B} < 1$ and a *high-risk habitat* if $\mathcal{B} \geq 1$.

1.2 Movement

Let $S = (S_j)$, $I = (I_j)$, $\hat{S} = (\hat{S}_j)$, $\hat{I} = (\hat{I}_j)$, $S' = S(t + \Delta t)$, and $I' = I(t + \Delta t)$. We assume that after infection and recovery occur, then individuals redistribute themselves throughout the habitat according to

$$S' = P\hat{S} \quad \text{and} \quad I' = Q\hat{I}, \quad t \in \mathfrak{T}. \tag{1.5}$$

Here, the entries of the matrices $P = (p_{jk})$ and $Q = (q_{jk})$ are constants and they represent the fraction of susceptible and infected individuals in patch k that move to patch j, respectively, in the time interval Δt.

We define a *dispersal matrix* to be a constant matrix which is doubly stochastic (i.e., stochastic, symmetric, and nonnegative) and irreducible. We assume that

(A5) P and Q are dispersal matrices.

The irreducibility assumption implies that the system of patches considered as a directed graph with patches as the vertices is strongly connected [25]. Other characterizations of irreducibility are given in Allen et al. (2007a, Appendix A), and we will make use of them as needed.

The assumptions on the parameters β_j, γ_j, p_{jk}, and q_{jk} and the functions $\Theta_j(x)$ in (1.2) imply that solutions $(\hat{S}, \hat{I}, S, I)$ to (1.1) and (1.5) are nonnegative for all time $t \in \mathfrak{T}$. Observe from (1.1) that $\hat{S}_j + \hat{I}_j = S_j + I_j$ for $j \in \Omega$. Let

$$N(t) = \sum_{j\in\Omega} \big[S_j(t) + I_j(t)\big] = \sum_{j\in\Omega} \big[\hat{S}_j(t) + \hat{I}_j(t)\big] \tag{1.6}$$

be the total population size at time t, and let $N = N(0)$. Assumption (A1) implies that N is positive. It follows from (1.1), (1.5), and the properties of P and Q that

$$N(t) \equiv N, \quad t \in \mathfrak{T}. \tag{1.7}$$

For an arbitrary patch $j \in \Omega$, it will be convenient to define

$$Q_j = \sum_{k\in\Omega} q_{jk}, \quad Q_j^- = \sum_{k\in H^-} q_{jk}, \quad \text{and} \quad Q_j^+ = \sum_{k\in H^+} q_{jk}.$$

For infected individuals, these sums denote the probability of moving from patch j to an arbitrary patch (Q_j), to a low-risk patch (Q_j^-), and to a high-risk patch (Q_j^+), respectively. The properties of Q imply that $Q_j = Q_j^- + Q_j^+ = 1$ for all $j \in \Omega$.

1.3 One parameter families of dispersal matrices

Let $R = (r_{jk})$ be a dispersal matrix. Define

$$R(\alpha) = \alpha R + (1-\alpha)\mathcal{I}_n, \quad \alpha \geq 0,$$

and $\mathfrak{A}(R) = \{R(\alpha) : R(\alpha) \text{ is a dispersal matrix}\}$. We say that R is a *baseline* dispersal matrix for the family of dispersal matrices $\mathfrak{A}(R)$. Of course, $R = R(1) \in \mathfrak{A}(R)$ but $\mathcal{I}_n = R(0) \notin \mathfrak{A}(R)$ because $\mathcal{I}_n$ is not irreducible. In view of the definition of a dispersal matrix, we see that $R(\alpha) \in \mathfrak{A}(R)$ if and only if $0 < \alpha \leq \min_{j,k\in\Omega}\{\pi_{jk}\}$, where we set

$$\pi_{jk} = \begin{cases} \frac{1}{r_{jk}}, & j \neq k \\ \frac{1}{1-r_{jj}}, & j = k. \end{cases}$$

If $r_{jk} = 0$ or $r_{jj} = 1$ then we define the corresponding π_{jk} to be infinite. Observe that $\pi_{jk} \geq 1$ for all $j, k \in \Omega$ because R is stochastic. Therefore, $\{R(\alpha) : 0 < \alpha \leq 1\} \subseteq \mathfrak{A}(R)$.

Two interesting cases arise. The first is when the dispersal matrices P and Q are coupled to the same baseline dispersal matrix and susceptibles diffuse slower than infected individuals. Thus, we sometimes assume that

(B2) $P, Q \in \mathfrak{A}(R)$ where R is a dispersal matrix, $P = R(d_S)$, $Q = R(d_I)$, and $0 < d_S \leq d_I$.

If this condition holds, then

$$\begin{aligned} P &= d_S R + (1-d_S)\mathcal{I}_n, \\ Q &= d_I R + (1-d_I)\mathcal{I}_n. \end{aligned} \tag{1.8}$$

We will refer to the parameters d_S and d_I as *diffusion coefficients*.

The second interesting case arises when Q belongs to a baseline dispersal matrix R whose entries satisfy

$$r_{jk} = \begin{cases} \mathcal{O}(\epsilon), & j \neq k \\ 1 - \mathcal{O}(\epsilon), & j = k, \end{cases} \tag{1.9}$$

because in this case $\mathfrak{A}(R) = \{R(\alpha) : \alpha > 0\}$. For simplicity, we will say that R is *slight* when its entries satisfy (1.9) since it closely approximates the matrix representing no dispersal ($\mathcal{I}_n$). Thus, we sometimes assume that

(B3) $Q \in \mathfrak{A}(R)$ where R is a slight dispersal matrix and $Q = R(d_I)$ with $d_I > 0$.

1.4 The equilibrium problem

We will be interested in equilibrium solutions of (1.1) and (1.5), i.e., solutions of

$$\hat{S}_j = (1 - \Theta_j)\tilde{S}_j + \gamma_j \tilde{I}_j, \quad j \in \Omega, \tag{1.10a}$$
$$\hat{I}_j = \Theta_j \tilde{S}_j + (1 - \gamma_j)\tilde{I}_j, \quad j \in \Omega, \tag{1.10b}$$
$$\tilde{S}_j = (P\hat{S})_j, \quad j \in \Omega, \tag{1.10c}$$
$$\tilde{I}_j = (Q\hat{I})_j, \quad j \in \Omega, \tag{1.10d}$$

where $\tilde{S}_j$ and $\tilde{I}_j$ denote the number of susceptible and infected individuals in patch $j \in \Omega$ at equilibrium, respectively, and Θ_j in (1.2) depends on $\tilde{S}_j$ and $\tilde{I}_j$ rather than S_j and I_j. In view of (1.6) and (1.7), we impose the additional condition

$$\sum_{j\in\Omega} \left(\tilde{S}_j + \tilde{I}_j\right) = \sum_{j\in\Omega} \left(\hat{S}_j + \hat{I}_j\right) = N. \tag{1.10e}$$

Observe from (1.10a) and (1.10b) that

$$\hat{S}_j + \hat{I}_j = \tilde{S}_j + \tilde{I}_j, \quad j \in \Omega. \tag{1.11}$$

Let $\tilde{S} = (\tilde{S}_j)$ and $\tilde{I} = (\tilde{I}_j)$. We are only interested in nonnegative solutions $(\hat{S}, \hat{I}, \tilde{S}, \tilde{I})$ of (1.10). A *disease-free equilibrium* (DFE) is a solution in which $\tilde{I}_j = 0$ for all $j \in \Omega$. An *endemic equilibrium* (EE) is a solution in which $\tilde{I}_j > 0$ for some $j \in \Omega$. To distinguish between these two types of equilibria, we will for notational convenience denote a DFE by $(\hat{S}, 0, \bar{S}, 0)$ or just $(\bar{S}, 0)$ and an EE by $(\hat{S}, \hat{I}, \tilde{S}, \tilde{I})$ or just $(\tilde{S}, \tilde{I})$.

1.5 Statement of the main results

We consider in Sect. 2 properties of the DFE, including its existence, uniqueness, and stability. Let $M = (q_{jk}\nu_k)$ be the matrix obtained by weighting the k-th column of the infected dispersal matrix Q by the virulence of the k-th patch. We remark that M is not, however, itself a dispersal matrix. We show in Sect. 2.2 that its spectral radius $\rho(M) < 1$ (> 1) if and only if $\mathcal{R}_0 < 1$ (> 1), where $\mathcal{R}_0$ is the basic reproduction number.

Theorem 1. *Suppose that* (A1)–(A5) *hold and N is fixed. There exists a DFE, it is unique, and it is given by $(\bar{S}, 0)$ where $\bar{S}_j = N/n$ for $j \in \Omega$. If $\mathcal{R}_0 < 1$ then $I(t) \to 0$ as $t \to \infty$ for all initial conditions, but if $\mathcal{R}_0 > 1$ then $I(t) \nrightarrow 0$ as $t \to \infty$ for some initial conditions, even in arbitrarily small neighborhoods of the DFE.*

It is interesting to note that $\mathcal{R}_0$ does not depend on the susceptible dispersal matrix P.

When Q belongs to a family of dispersal matrices $\mathfrak{A}(R)$ with baseline dispersal matrix R and parameter d_I, we can characterize the persistence of the disease in terms of the habitat type and d_I rather than $\mathcal{R}_0$. In particular, we show that the disease persists in a low-risk habitat if and only if the diffusion coefficient for infected individuals lies below a certain threshold value (depending on R), but in a high-risk habitat, the disease always persists.

Theorem 2. *Suppose that* (A1)–(A5) *and* (B3) *hold and N is fixed.*

(a) *In a low-risk habitat ($\mathcal{B} < 1$), there exists a threshold value $d_I^* \in (0, \infty)$ such that $\mathcal{R}_0 > 1$ for $d_I < d_I^*$ and $\mathcal{R}_0 < 1$ for $d_I > d_I^*$;*

(b) *In a high-risk habitat ($\mathcal{B} \geq 1$), we have $\mathcal{R}_0 > 1$ for all d_I.*

In Sect. 3, we show that when the disease persists, then there exists a unique EE. Moreover, the disease persists in every patch.

Theorem 3. *Suppose that* (A1)–(A5) *and* (B1)–(B2) *hold and N is fixed. If $\mathcal{R}_0 > 1$ then an EE exists, it is unique, and $\tilde{I} > 0$.*

Observe from (1.8), (1.10), and the properties of Θ_j that in the limiting case $d_S = 0$ there also exists a family of infinitely many spatially inhomogeneous DFEs $(\bar{S}, 0)$, each of which satisfies

$$\bar{S} \geq 0 \quad \text{and} \quad \sum_{j \in \Omega} \bar{S}_j = N. \tag{1.12}$$

In Sect. 4, we show that if $\mathcal{R}_0 > 1$ then the EE approaches such a spatially inhomogeneous DFE as the mobility of susceptible individuals becomes very small. We write this limiting DFE as $(S^*, 0)$ and also consider the distribution of patches for which S^* is either positive or zero.

Theorem 4. *Suppose that* (A1)–(A5) *and* (B1)–(B2) *hold, N is fixed, and $\mathcal{R}_0 > 1$.*

(a) *$(\tilde{S}, \tilde{I}) \to (S^*, 0)$ as $d_S \to 0$ for some S^* satisfying* (1.12)*;*

(b) *$S^* > 0$ on H^- and $S_j^* = 0$ for some $j \in H^+$;*

(c) *If*

$$\nu_j\big(Q_j^- + \min_{k \in H^+} \{\nu_k\}\, Q_j^+\big) > 1 \tag{1.13}$$

for all $j \in H^-$ then $S^* \equiv 0$ on H^+;

(d) *If*

$$\nu_j\big(Q_j^+ + \max_{k \in H^-} \{\nu_k\}\, Q_j^-\big) < 1 \tag{1.14}$$

for some $j \in H^+$ then $S_j^* > 0$.

We now make several remarks concerning Theorem 4, which connects spatial heterogeneity, habitat connectivity, and rates of movement. First, Theorem 4 (c) immediately implies that if

$$\min_{j \in H^-} \{\nu_j Q_j^+\} \min_{k \in H^+} \{\nu_k\} > 1 \tag{1.15}$$

then $S^* \equiv 0$ on H^+. Although condition (1.13) is more inclusive than condition (1.15), the latter is usually easier to verify. Second, if infected individuals in low-risk patches must always disperse to high-risk patches ($Q_j^- = 0$ and $Q_j^+ = 1$ for all $j \in H^-$), and if

$$\min_{j\in H^-}\{\nu_j\}\min_{k\in H^+}\{\nu_k\} > 1$$

then $S^* \equiv 0$ on H^+.

Third, Theorem 4 (d) implies that if

$$\min_{j\in H^+}\left\{\nu_j\big(Q_j^+ + \max_{k\in H^-}\{\nu_k\}\,Q_j^-\big)\right\} < 1$$

then $S^* \not\equiv 0$ on H^+. In particular, if infected individuals in high-risk patches must always disperse to low-risk patches ($Q_j^+ = 0$ and $Q_j^- = 1$ for all $j \in H^+$), and if

$$\min_{j\in H^+}\{\nu_j\}\max_{k\in H^-}\{\nu_k\} < 1$$

then $S^* \not\equiv 0$ on H^+.

Fourth, if some low-risk patch is not directly connected to any high-risk patches ($Q_j^- = 1$ and $Q_j^+ = 0$ for some $j \in H^-$) then condition (1.13) will fail to be satisfied because low-risk patches have virulence less than one. Similarly, if some high-risk patch is not directly connected to any low-risk patches ($Q_j^+ = 1$ and $Q_j^- = 0$ for some $j \in H^+$) then condition (1.14) will also fail to be satisfied because high-risk patches have virulence greater than one. Finally, conditions (1.13) and (1.14) will both fail to be satisfied whenever d_I is sufficiently small, because in this case Q is slight, i.e., it closely approximates $\mathcal{I}_n$.

1.6 Examples

Before proving Theorems 1–4, we first illustrate the final theorem with some examples.

Example 1. If $H^- = \{1, 2, \ldots, n-1\}$ and $H^+ = \{n\}$ then Theorem 4 (b) implies that $S^* > 0$ on H^- and $S^* = 0$ on H^+. For this case, condition (1.13) in Theorem 4 (c) may or may not hold but condition (1.14) in Theorem 4 (d) cannot. That is, for this particular configuration of patches, the assumption that $\mathcal{R}_0 > 1$ in Theorem 4 excludes the possibility that condition (1.14) can be satisfied.

Example 2. Suppose that $H^- = \{1\}$, $H^+ = \{2, 3, \ldots, n\}$, and that in each time step infected individuals in the low-risk patch always disperse to high-risk patches and those in high-risk patches always disperse to the low-risk patch. In this case, Theorem 4 (b) implies that $S^* > 0$ on H^- and Theorem 4 (c, d) provides necessary and sufficient conditions (except in the case of equality) for determining whether $S^* \equiv 0$ on H^+. In particular, if $\nu_1 \min_{k\in H^+}\{\nu_k\} > 1$ then $S^* \equiv 0$ on H^+, but if $\nu_1 \min_{k\in H^+}\{\nu_k\} < 1$ then $S^* \not\equiv 0$ on H^+.

Example 3. Suppose there are $n = 100$ patches arranged in a 10×10 grid and connected as in Figure 1 (a). As illustrated in Figure 1 (b), the limiting DFE (with $d_S = 0$) is positive on all low-risk patches and is empty elsewhere, as predicted by Theorem 4 parts (b) and (c). For this example, condition (1.13) is satisfied because $\nu_j\big(Q_j^- + \min_{k\in H^+}\{\nu_k\}\,Q_j^+\big) = 1.10$ for all $j \in H^-$.

Example 4. Suppose there are $n = 100$ patches arranged and connected as in Figure 2 (a). As illustrated in Figure 2 (b), the limiting DFE (with $d_S = 0$) is positive not only on all low-risk patches but also on nearly all high-risk patches (96 in all). The DFE is empty on the four remaining high-risk patches: two in opposite corners of the habitat and two in the interior. Many of these properties are predicted by Theorem 4 parts (b) and (d). For this example, condition (1.15) is satisfied by the high-risk patch with coordinates $(6, 3)$, for which $\nu_j\big(Q_j^+ + \max_{k\in H^-}\{\nu_k\}\, Q_j^-\big) = 0.91$.

2 The Disease-Free Equilibrium

Throughout this section, we assume that (A1)–(A5) hold and that N is fixed. Theorem 1 follows from Lemmas 2.1, 2.2, and 2.3 and Theorem 2 follows from Lemma 2.5.

2.1 Persistence of the disease

First, we show that a unique disease-free solution exists, and it is spatially homogeneous.

Lemma 2.1. *A DFE $(\bar{S}, 0)$ exists, it is unique, and it is given by $\bar{S}_j = N/n$ for $j \in \Omega$.*

Proof. It is straightforward to see from (1.10) and the properties of Θ_j and P that $(\bar{S}, 0)$, with $\bar{S}_j = N/n$ for $j \in \Omega$, is a DFE. Now, let $(\tilde{S}, 0)$ be any DFE. Setting $\tilde{I} = 0$ in (1.10) leads to $\tilde{S} = P\tilde{S}$. The properties of P and the Frobenius Theorem ([19], Theorem 2, p. 53) imply that $\tilde{S} \in \langle 1_n \rangle$, as the eigenvectors $\tilde{S}$ and 1_n both belong to the same principle eigenvalue $\mu = 1$. In view of (1.10e) with $\tilde{I} = 0$, we obtain $\tilde{S}_j = N/n$ for $j \in \Omega$. □

Let

$$m_{jk} \stackrel{\text{def}}{=} q_{jk}\big(1 + \beta_k - \gamma_k\big) \tag{2.1}$$

and $M = (m_{jk})$. Since there are no sanctuary patches, the nonnegative matrix M has the same digraph as Q and is therefore irreducible. We now show that the spectral radius of M controls the behavior of the disease.

Lemma 2.2. *If $\rho(M) < 1$ then $I(t) \to 0$ as $t \to \infty$ for all initial conditions.*

Proof. Suppose that $\rho(M) < 1$, where $M = (m_{jk})$ is as in (2.1). We will use the comparison principle to show that $I(t) \to 0$ as $t \to \infty$. Observe from (1.5), (1.1), (1.2), and the properties of Θ_j that if $j \in \Omega$ then

$$\begin{aligned}
I_j' &= (Q\hat{I})_j \\
&= \sum_{k\in\Omega} q_{jk}\big[\Theta_k S_k + (1-\gamma_k)I_k\big] \\
&= \sum_{k\in\Omega} q_{jk}\left[\Theta_k\left(\frac{\beta_k I_k}{S_k + I_k}\right) S_k + (1-\gamma_k)I_k\right] \\
&\le \sum_{k\in\Omega} q_{jk}\left[\left(\frac{\beta_k I_k}{S_k + I_k}\right) S_k + (1-\gamma_k)I_k\right] \\
&\le \sum_{k\in\Omega} q_{jk}\big[\beta_k I_k + (1-\gamma_k)I_k\big] \\
&= \sum_{k\in\Omega} q_{jk}\big(1 + \beta_k - \gamma_k\big)I_k \\
&= (MI)_j.
\end{aligned}$$

It follows that $I' \leq MI$. The linear system $x' = Mx$ with $x(0) = I(0)$ satisfies $x(t) \to 0$ as $t \to \infty$ because $\rho(M) < 1$. Furthermore, the nonnegativity of M implies that if $I(t) \leq x(t)$ then $I' \leq MI \leq Mx = x'$. By induction, $I(t) \leq x(t)$ for all $t \in \mathfrak{T}$. We conclude that $I(t) \to 0$ as $t \to \infty$. □

When the disease can persist, it can do so even when a very small number of individuals in the population are initially infected.

Lemma 2.3. *If $\rho(M) > 1$ then $I(t) \nrightarrow 0$ as $t \to \infty$ for some initial conditions, even in arbitrarily small neighborhoods of the DFE.*

Proof. Linearizing (1.1) and (1.5) around the DFE $(\bar{S}, 0)$ produces the system of equations

$$\begin{bmatrix} \eta' \\ \xi' \end{bmatrix} = \begin{bmatrix} \partial F/\partial S & \partial F/\partial I \\ \partial G/\partial S & \partial G/\partial I \end{bmatrix}_* \begin{bmatrix} \eta \\ \xi \end{bmatrix}$$

where $(S, I) = (\bar{S}, 0) + (\eta, \xi)$, $F = (F_j)$, and $G = (G_j)$ all consist of vectors of length n, and

$$\begin{aligned} F_j(S, I) &= \sum_{k \in \Omega} p_{jk}\big[(1 - \Theta_k)S_k + \gamma_k I_k\big], \\ G_j(S, I) &= \sum_{k \in \Omega} q_{jk}\big[\Theta_k S_k + (1 - \gamma_k)I_k\big]. \end{aligned}$$

The $*$–notation means that the given quantity is to be evaluated at the DFE. It suffices to examine the behavior of $\xi(t)$ as $t \to \infty$. Observe that

$$\xi_j' = \left(\frac{\partial G_j}{\partial S}\right)^* \eta + \left(\frac{\partial G_j}{\partial I}\right)^* \xi.$$

We first compute

$$\begin{aligned} \left(\frac{\partial G_j}{\partial S}\right)^* \eta &= \sum_{k \in \Omega} q_{jk} \left[\left(\frac{\partial \Theta_k}{\partial S_k}\right)^* \bar{S}_k + \Theta_k^*\right] \eta_k \\ &= 0. \end{aligned}$$

In the last step, we used the fact that $\Theta_k^* = (\partial \Theta_k/\partial S_k)^* = 0$. We now compute

$$\begin{aligned} \left(\frac{\partial G_j}{\partial I}\right)^* \xi &= \sum_{k \in \Omega} q_{jk} \left[\left(\frac{\partial \Theta_k}{\partial I_k}\right)^* \bar{S}_k + (1 - \gamma_k)\right] \xi_k \\ &= \sum_{k \in \Omega} q_{jk}\big(1 + \beta_k - \gamma_k\big)\xi_k, \end{aligned}$$

where this time $(\partial \Theta_k/\partial I_k)^* = \beta_k/\bar{S}_k$. It follows from these calculations that $\xi_j' = (M\xi)_j$ and hence $\xi' = M\xi$. Since M is nonnegative and irreducible, $\rho(M)$ is an eigenvalue and it possesses a positive eigenvector ϕ^* by the Frobenius theorem [19]. Now let $\epsilon > 0$ and set $\xi(0) = \epsilon\phi^*$. Then $\|\xi(\ell\Delta t)\|_2 = \epsilon[\rho(M)]^\ell \|\phi^*\|_2 \to \infty$ as $\ell \to \infty$. □

2.2 The basic reproduction number

Here, we use the next generation matrix approach to calculate the basic reproduction number $\mathcal{R}_0$ [6, 13, 22, 23]. In particular, we will show that $\mathcal{R}_0 < 1$ (> 1) if and only if $\rho(M) < 1$ (> 1).

To calculate the next generation matrix, we separate new infections from other transitions and reorder the states. For $j \in \Omega$, let

$$\mathcal{J}_j = \sum_{k\in\Omega} q_{jk}\Theta_k S_k, \quad \mathcal{T}_j = \sum_{k\in\Omega} q_{jk}(1-\gamma_k)I_k, \quad \text{and} \quad \mathcal{F}_j = \sum_{k\in\Omega} p_{jk}[(1-\Theta_k)S_k + \gamma_k I_k].$$

With $\mathcal{J} = (\mathcal{J}_j)$, $\mathcal{T} = (\mathcal{T}_j)$, and $\mathcal{F} = (\mathcal{F}_j)$, the equations in (1.1) and (1.5) are equivalent to

$$I' = \mathcal{J} + \mathcal{T} \quad \text{and} \quad S' = \mathcal{F}.$$

As shown in Sect. 2.1, linearization of this reordered system about the DFE leads to

$$\begin{bmatrix}\xi' \\ \eta'\end{bmatrix} = \begin{bmatrix}(\partial\mathcal{J}/\partial I)^* + (\partial\mathcal{T}/\partial I)^* & O_n \\ (\partial\mathcal{F}/\partial I)^* & (\partial\mathcal{F}/\partial S)^*\end{bmatrix}\begin{bmatrix}\xi \\ \eta\end{bmatrix}.$$

Here, O_n is the $n\times n$ zero matrix. Next, let $J_{jk} = q_{jk}\beta_k$, $T_{jk} = q_{jk}(1-\gamma_k)$, and $B_{jk} = p_{jk}(\gamma_k - \beta_k)$ for $j,k \in \Omega$. Then

$$\begin{bmatrix}\xi' \\ \eta'\end{bmatrix} = \begin{bmatrix}J+T & O_n \\ B & P\end{bmatrix}\begin{bmatrix}\xi \\ \eta\end{bmatrix},$$

where $J = (J_{jk})$, $T = (T_{jk})$, and $B = (B_{jk})$. In particular, $\xi' = (J+T)\xi = M\xi$. Now, we derive conditions which are equivalent to $\rho(M) < 1$ (> 1). Note that $M = J + T$ is irreducible and $\rho(T) < 1$. The next generation matrix is well-known in discrete-time population dynamical theory [12] and has the form

$$\mathcal{M} = J(\mathcal{I}_n - T)^{-1} = J(\mathcal{I}_n + T + T^2 + \cdots).$$

Biologically, $\mathcal{M}I(0)$ represents the accumulated number of infections over all generations. The basic reproduction number is now defined to be $\mathcal{R}_0 = \rho(\mathcal{M})$, and it follows from Theorem 3.3 in Li and Schneider ([23], p. 455) that $\mathcal{R}_0 < 1$ (> 1) if and only if $\rho(M) < 1$ (> 1).

2.3 An eigenvalue problem

Recall from the proof of Lemma 2.3 that if $(\bar{S}, 0) + (\eta, \xi)$ is a perturbation around the DFE then

$$\xi'_j = (M\xi)_j = \sum_{k\in\Omega} q_{jk}\big(1 + \beta_k - \gamma_k\big)\xi_k.$$

Suppose that $\xi = \tilde{\lambda}^t\phi$, where $\tilde{\lambda} \in \mathbb{R}$ and $\phi = (\phi_j) \in \mathbb{R}^n$. Then

$$\lambda\phi_j = \sum_{k\in\Omega} q_{jk}\big(1 + \beta_k - \gamma_k\big)\phi_k, \quad j \in \Omega, \tag{2.2}$$

where $\lambda = \tilde{\lambda}^{\Delta t}$. Alternatively,

$$\lambda\varphi_j = \big(1 + \beta_j - \gamma_j\big)(Q\varphi)_j, \quad j \in \Omega, \tag{2.3}$$

where $\varphi_j \stackrel{\text{def}}{=} (1 + \beta_j - \gamma_j)\phi_j$ and $\varphi = (\varphi_j)$. As an aside, we remark that (2.3) is actually the eigenvalue problem for the related half-step problem in which infection and recovery occur *after* movement.

Observe from (2.1) and (2.2) that $M\phi = \lambda\phi$, where M is nonnegative and irreducible. By the Frobenius Theorem [19], the spectral radius of M is an eigenvalue, which we denote by λ^*. This principal eigenvalue is positive and real, and its eigenspace is given by $\langle\phi^*\rangle$ for some $\phi^* = (\phi_j^*) > 0$. The next result follows immediately from these remarks and the Frobenius Theorem.

Lemma 2.4. *Let $\varphi_j^* = (1+\beta_j-\gamma_j)\phi_j^*$ and $\varphi^* = (\varphi_j^*)$. Then (λ^*, φ^*) satisfies* (2.3)*, i.e.,*

$$\lambda^*\varphi_j^* = \big(1+\beta_j-\gamma_j\big)(Q\varphi^*)_j, \quad j \in \Omega. \tag{2.4}$$

Moreover, (λ^*, φ) satisfies (2.3) *if and only if $\varphi \in \langle\varphi^*\rangle$. Finally, if (λ, φ) satisfies* (2.3) *with $\lambda \neq \lambda^*$ then $|\lambda| \leq \lambda^*$ and $\varphi_j \leq 0$ for some $j \in \Omega$.*

2.4 Habitat type and the movement of infected individuals

The next result connects the fate of the disease to both the habitat type and the movement rate of infected individuals when Q belongs to a family $\mathfrak{A}(R)$ with slight baseline dispersal matrix R and parameter d_I. Recall that in this case, $\mathfrak{A}(R) = \{R(d_I) : d_I > 0\}$.

Lemma 2.5. *Suppose that* (B3) *holds and let $\mathcal{B}$ be as in* (1.4)*.*

(a) *λ^* is a strictly monotone decreasing function of $d_I > 0$;*

(b) *$\lambda^* \to \max_{j\in\Omega}\{1+\beta_j-\gamma_j\}$ as $d_I \to 0$;*

(c) *$\lambda^* \to \mathcal{B}$ as $d_I \to \infty$;*

(d) *If $\mathcal{B} \geq 1$ then $\lambda^* > 1$ for all $d_I > 0$;*

(e) *If $\mathcal{B} < 1$ then the equation $\lambda^*(d_I) = 0$ has a unique positive root denoted by d_I^*. Furthermore, if $d_I < d_I^*$ then $\lambda^* > 1$ and if $d_I > d_I^*$ then $\lambda^* < 1$.*

The proof of Lemma 2.5 appears in Appendix A.

3 The Endemic Equilibrium

Throughout this section, we assume that (A1)–(A5) and (B1)–(B2) hold and that N is fixed. We also assume that $\lambda^* = \rho(M) > 1$, so that the disease can persist for some initial conditions, even in arbitrarily small neighborhoods of the DFE. Theorem 3 follows from Lemma 3.10.

3.1 Existence of the EE

It will be useful to consider several alternative statements of the equilibrium problem. The first such statement is as follows.

Lemma 3.1. *$(\hat{S}, \hat{I}, \tilde{S}, \tilde{I})$ is a solution of* (1.10) *if and only if $(\hat{S}, \hat{I}, \tilde{S}, \tilde{I})$ is a solution of*

$$\kappa = d_S\hat{S}_j + d_I\hat{I}_j, \qquad j \in \Omega, \tag{3.1a}$$

$$\hat{I}_j = \Theta_j\tilde{S}_j + (1-\gamma_j)\tilde{I}_j, \quad j \in \Omega, \tag{3.1b}$$

$$\tilde{S}_j = (P\hat{S})_j, \qquad j \in \Omega, \tag{3.1c}$$

$$\tilde{I}_j = (Q\hat{I})_j, \qquad j \in \Omega, \tag{3.1d}$$

$$N = \sum_{j\in\Omega}(\tilde{S}_j + \tilde{I}_j), \tag{3.1e}$$

where κ is some positive constant independent of $j \in \Omega$.

Proof. Suppose first that $(\hat{S},\hat{I},\tilde{S},\tilde{I})$ is a solution of (1.10). We will show that there exists some $\kappa>0$ such that $\hat{S}$ and $\hat{I}$ satisfy (3.1a). Observe from (1.8) and (1.10) that

$$\begin{aligned}\tilde{S}+\tilde{I} &= P\hat{S}+Q\hat{I}\\ &= \big[d_S R+(1-d_S)\mathcal{I}_n\big]\hat{S}+\big[d_I R+(1-d_I)\mathcal{I}_n\big]\hat{I}\\ &= R(d_S\hat{S}+d_I\hat{I})+\hat{S}+\hat{I}-(d_S\hat{S}+d_I\hat{I}).\end{aligned}$$

Eq. (1.11) implies that $R\big(d_S\hat{S}+d_I\hat{I}\big)=d_S\hat{S}+d_I\hat{I}$. The properties of R and the Frobenius Theorem [19] imply that $d_S\hat{S}+d_I\hat{I}\in\langle 1_n\rangle$, as the eigenvectors $d_S\hat{S}+d_I\hat{I}$ and 1_n both belong to the same principle eigenvalue $\mu=1$. That is, $d_S\hat{S}_j+d_I\hat{I}_j=\kappa$ for some $\kappa\in\mathbb{R}$ and every $j\in\Omega$. Since $d_S\hat{S}_j+d_I\hat{I}_j>0$ for at least one $j\in\Omega$ (because $N>0$) it must be that $\kappa>0$. The fact that $(\hat{S},\hat{I},\tilde{S},\tilde{I})$ satisfies (3.1b)–(3.1e) is clear by inspection. If $(\hat{S},\hat{I},\tilde{S},\tilde{I})$ is a solution of (3.1) for some $\kappa>0$ then it follows from a direct calculation that $(\hat{S},\hat{I},\tilde{S},\tilde{I})$ satisfies (1.10). □

For our second formulation of the equilibrium problem, let

$$S_j^\circ=\frac{\hat{S}_j}{\kappa}\quad\text{and}\quad I_j^\circ=\frac{d_I\hat{I}_j}{\kappa},\tag{3.2}$$

where κ is as in (3.1a), and for $I\in\mathbb{R}^n$ let

$$H_j(I)=\frac{d_I}{d_S}\Theta_j\left(\frac{\beta_j d_S(QI)_j}{d_I+(d_S-d_I)I_j}\right)\big[1-(PI)_j\big]+(1-\gamma_j)(QI)_j,\quad j\in\Omega.\tag{3.3}$$

Observe that $H_j(0)=0$ and $H_j(1_n)=1-\gamma_j$ for $j\in\Omega$. Finally, let $S^\circ=(S_j^\circ)$, $I^\circ=(I_j^\circ)$, and $H=(H_j)$.

Lemma 3.2. *$(\hat{S},\hat{I},\tilde{S},\tilde{I})$ is a solution of (3.1) if and only if $(S^\circ,I^\circ,\tilde{S},\tilde{I})$ is a solution of*

$$1_n=d_S S^\circ+I^\circ,\tag{3.4a}$$
$$I^\circ=H(I^\circ),\tag{3.4b}$$
$$\tilde{S}=\kappa P S^\circ,\tag{3.4c}$$
$$\tilde{I}=(\kappa/d_I)QI^\circ,\tag{3.4d}$$
$$\kappa=\frac{d_I N}{\sum_{j\in\Omega}\big[d_I(PS^\circ)_j+(QI^\circ)_j\big]}.\tag{3.4e}$$

Proof. Let $(\hat{S},\hat{I},\tilde{S},\tilde{I})$ be a solution of (3.1). It is clear from (3.1a) and (3.2) that $1=d_S S_j^\circ+I_j^\circ$ for every $j\in\Omega$. Therefore, (3.4a) is satisfied. Next, let $j\in\Omega$ and observe from (3.1) and (1.11) that

$$\hat{I}_j=\Theta_j\left(\frac{\beta_j\tilde{I}_j}{\tilde{S}_j+\tilde{I}_j}\right)\tilde{S}_j+(1-\gamma_j)\tilde{I}_j=\Theta_j\left(\frac{\beta_j(Q\hat{I})_j}{\hat{S}_j+\hat{I}_j}\right)(P\hat{S})_j+(1-\gamma_j)(Q\hat{I})_j.$$

We multiply both sides by d_I/κ and then apply (3.2), (3.4a), the properties of P, and (3.3) to get

$$\begin{aligned}I_j^\circ &= d_I\Theta_j\left(\frac{\beta_j(QI^\circ)_j}{d_I S_j^\circ+I_j^\circ}\right)(PS^\circ)_j+(1-\gamma_j)(QI^\circ)_j\\ &= \frac{d_I}{d_S}\Theta_j\left(\frac{\beta_j d_S(QI^\circ)_j}{d_I(1-I_j^\circ)+d_S I_j^\circ}\right)\big(P\big[1_n-I^\circ\big]\big)_j+(1-\gamma_j)(QI^\circ)_j\end{aligned}$$

$$= \frac{d_I}{d_S}\Theta_j\left(\frac{\beta_j d_S (QI^\circ)_j}{d_I + (d_S - d_I)I^\circ_j}\right)\left[1 - (PI^\circ)_j\right] + (1-\gamma_j)(QI^\circ)_j$$
$$= H_j(I^\circ).$$

Thus, (3.4b) is satisfied. The fact that (3.4c)–(3.4e) are satisfied is immediate from (3.2). If $(S^\circ, I^\circ, \tilde{S}, \tilde{I})$ is a solution of (3.4) for some $\kappa > 0$ then it follows from a direct calculation that $(\hat{S}, \hat{I}, \tilde{S}, \tilde{I})$ satisfies (3.1). □

The benefit of the alternative formulation in Lemma 3.2 is that (3.4b) depends on I° but not on S°, $\tilde{S}$, $\tilde{I}$, or κ. Thus, once I° is determined, then it is a simple matter to determine S° from (3.4a), κ from (3.4e), and then $\tilde{S}$ and $\tilde{I}$ from (3.4c) and (3.4d).

Lemma 3.3. *Eq.* (3.4) *has a nonnegative solution* $(S^\circ, I^\circ, \tilde{S}, \tilde{I})$ *which can be chosen to satisfy* $I^\circ \not\equiv 0$. *Furthermore, this solution (with* $I^\circ \not\equiv 0$*) is unique, it is positive, and* $I^\circ < 1_n$.

Below, we prove the existence of such a solution $(S^\circ, I^\circ, \tilde{S}, \tilde{I})$, and in the next subsection we will demonstrate that it is unique. A useful lemma, whose proof follows immediately from (1.8), is as follows.

Lemma 3.4. *If* $I \in \mathbb{R}^n$ *then* $d_S QI = d_I PI + (d_S - d_I)I$.

For $u, w \in \mathbb{R}^n$, we define $(u, w) = \{v \in \mathbb{R}^n : u_j < v_j < w_j \text{ for } j \in \Omega\}$ and $[u, w] = \{v \in \mathbb{R}^n : u_j \le v_j \le w_j \text{ for } j \in \Omega\}$. The half-open intervals $[u, w)$ and $(u, w]$ are defined similarly. The next lemma establishes important properties of the function H in (3.3). Its proof appears in Appendix B.

Lemma 3.5. *Let* φ^* *be as in Lemma* 2.4. *Then*

(a) $H(\epsilon\varphi^*) \ge \epsilon\varphi^*$ *for* $0 < \epsilon \ll 1$;

(b) $H(1_n) \le 1_n$;

(c) $H = (H_j)$ *is a monotone increasing function of* I_k *for every* $k \ne j$ *when* $I = (I_k) \in (0, 1_n]$.

We now prove the existence portion of Lemma 3.3. It is clear that $I \in [0, 1_n]$ is a solution of (3.4b) if and only if

$$G_j(I) \stackrel{\text{def}}{=} H_j(I) - I_j = 0, \quad j \in \Omega. \tag{3.5}$$

Let $G = (G_j)$. Lemma 3.5 (c) implies that (3.5) is a monotone system [28]. It follows that if $\underline{I}$ and $\overline{I}$ are ordered (i.e., $\underline{I} \le \overline{I}$), and they are a sub-solution and super-solution of (3.5), respectively, i.e., $G(\underline{I}) \ge 0 \ge G(\overline{I})$, then there must exist some $I \in [\underline{I}, \overline{I}]$ such that $G(I) = 0$. Lemma 3.5 (a, b) implies that we may take $\underline{I} = \epsilon\varphi^*$ and $\overline{I} = 1_n$ if ϵ is chosen to be positive and sufficiently small.

We conclude from the remarks above that there must be some $I^\circ \in (0, 1_n]$ such that $G(I^\circ) = 0$. That is, there exists some I° satisfying (3.4b) with $0 < I^\circ_j \le 1$ for $j \in \Omega$. We argue by contradiction to show that $I^\circ_j < 1$ for every $j \in \Omega$. First, if $I^\circ_j = 1$ for all $j \in \Omega$ then $G_j(I^\circ) = -\gamma_j < 0$ when $j \in H^-$, a contradiction. Suppose now that $I^\circ_j = 1$ and $I^\circ_m < 1$ for some $j, m \in \Omega$. Since Q is irreducible, we may assume that j and m are chosen so that $q_{jm} > 0$. It follows that $(QI^\circ)_j < 1$. Eq. (3.3), the properties of Θ_j, and Lemma 3.4 imply that

$$H_j(I^\circ) = \frac{d_I}{d_S}\Theta_j\left(\frac{\beta_j d_S (QI^\circ)_j}{d_I + (d_S - d_I)I^\circ_j}\right)\left[1 - (PI^\circ)_j\right] + (1-\gamma_j)(QI^\circ)_j$$

$$\begin{aligned}
&\leq \left[\beta_j\left(\frac{d_I - d_I(PI^\circ)_j}{d_I + (d_S - d_I)I_j^\circ}\right) + 1 - \gamma_j\right](QI^\circ)_j \\
&= \left[\beta_j\left(1 - \frac{d_S(QI^\circ)_j}{d_I + (d_S - d_I)I_j^\circ}\right) + 1 - \gamma_j\right](QI^\circ)_j \\
&= \big[1 + \beta_j - \gamma_j - \beta_j(QI^\circ)_j\big](QI^\circ)_j.
\end{aligned}$$

Let $\nu_j = 1 + \beta_j - \gamma_j$ and $q_j = (QI^\circ)_j$. Recall that $q_j \in (0,1)$. Observe that $(\nu_j - \beta_j x)x$ achieves its maximum value on $[0,1]$ when $x = \min\{\nu_j/2\beta_j, 1\}$. If $\nu_j < 2\beta_j$ then $H_j(I^\circ) \leq (\nu_j - \beta_j q_j)q_j \leq (\nu_j/2)^2/\beta_j < \beta_j \leq 1 = I_j^\circ$. Similarly, if $\nu_j \geq 2\beta_j$ then $H_j(I^\circ) \leq (\nu_j - \beta_j q_j)q_j < \nu_j - \beta_j = 1 - \gamma_j \leq 1 = I_j^\circ$. In either case, we obtain $H_j(I^\circ) < I_j^\circ$, again a contradiction. We conclude that $I^\circ < 1_n$. Therefore, $(S^\circ, I^\circ, \tilde{S}, \tilde{I})$ is a positive solution of (3.4) with $I^\circ < 1_n$. This completes the existence proof of Lemma 3.3. Next, we establish uniqueness.

3.2 Uniqueness of the EE

We will assume throughout this subsection that $(S^\circ, I^\circ, \tilde{S}, \tilde{I})$ is a nonnegative solution of (3.4). In view of (3.4a) and the argument given above, we may restrict attention to solutions I° of (3.4b) which satisfy $0 \leq I_j^\circ < 1$ for $j \in \Omega$.

Lemma 3.6. *Either $I^\circ \equiv 0$ or $I^\circ > 0$.*

Proof. We argue by contradiction. Suppose that $H(I^\circ) = I^\circ$ with $I^\circ \not\equiv 0$ and $I^\circ \not> 0$. Then there exist nonempty and disjoint subsets K^- and K^+ of Ω, whose union forms all of Ω, and with the property that $I_j^\circ = 0$ for $j \in K^-$ and $I_j^\circ > 0$ for $j \in K^+$. Suppose first that $(QI^\circ)_j = 0$ for every $j \in K^-$. Since $I_k^\circ > 0$ when $k \in K^+$, we obtain that $q_{jk} = 0$ when $j \in K^-$ and $k \in K^+$. But this contradicts the irreducibility of Q. Thus, it must be that $(QI^\circ)_j > 0$ for some $j \in K^-$. Recall from (3.3) that

$$\frac{d_I}{d_S}\Theta_j\left(\frac{\beta_j d_S(QI^\circ)_j}{d_I}\right)\big[1 - (PI^\circ)_j\big] + (1 - \gamma_j)(QI^\circ)_j = 0.$$

Observe that both terms on the left-hand side must vanish because they are nonnegative and their sum is zero. Since $(QI^\circ)_j > 0$, we conclude that $\gamma_j = 1$. As there are no sanctuary patches, we must also have that $\beta_j > 0$. Furthermore, $(PI^\circ)_j < 1$ because $I^\circ < 1_n$. But these remarks imply that the first term on the left-hand side is positive, another contradiction. We conclude that either $I^\circ \equiv 0$ or $I^\circ > 0$. □

It can be shown that there exist minimal and maximal solutions ($I^{\min}$ and $I^{\max}$) of (3.4b) such that $\epsilon\varphi^* < I^{\min} \leq I^{\max} < 1_n$ and with the property that, if ϵ is chosen small enough, then all positive solutions lie between them.

Lemma 3.7. *If I° is positive then $I^\circ \in [I^{\min}, I^{\max}]$.*

Proof. The proof is similar to the iteration argument presented in [1]. □

Lemma 3.8. *If $I^{\circ -}$ and $I^{\circ +}$ are positive with $I^{\circ -} \leq I^{\circ +}$, then either $I^{\circ -} < I^{\circ +}$ or $I^{\circ -} \equiv I^{\circ +}$.*

Proof. We argue by contradiction. Suppose that $I^{\circ -} = (I_j^{\circ -})$ and $I^{\circ +} = (I_j^{\circ +})$ are positive with $I^{\circ -} \leq I^{\circ +}$, and that neither $I^{\circ -} < I^{\circ +}$ nor $I^{\circ -} \equiv I^{\circ +}$. Then there exist nonempty and disjoint

subsets K^- and K^+ of Ω, whose union forms all of Ω, and with the property that $I_j^{\circ -} < I_j^{\circ +}$ for $j \in K^-$ and $I_j^{\circ -} = I_j^{\circ +}$ for $j \in K^+$. Recall from (3.3) that

$$H_j(I) = \Theta_j(X_j)\left[\frac{d_I - d_I(PI)_j}{d_S}\right] + (1-\gamma_j)(QI)_j,$$

where

$$X_j = X_j(I) = \frac{\beta_j d_S (QI)_j}{d_I + (d_S - d_I)I_j}.$$

We restrict attention to $I \in [I^{\min}, I^{\max}]$. We wish to show that $H_j(I)$ is a strictly increasing function of $(QI)_j$ when I_j is held fixed. If $\beta_j = 0$ then the result is immediate because there are no sanctuary patches (i.e., $0 \le \gamma_j < 1$). Suppose now that $\beta_j > 0$, which implies also that $X_j > 0$. We make use of Lemma 3.4 and then simplify to get

$$H_j(I) = \Theta_j(X_j)\left[\frac{d_I + (d_S - d_I)I_j}{d_S} - (QI)_j\right] + (1-\gamma_j)(QI)_j.$$

We now differentiate $H_j(I)$ with respect to $(QI)_j$ to obtain

$$H_j(I)^\dagger = \Theta_j'(X_j)X_j^\dagger\left[\frac{d_I + (d_S - d_I)I_j}{d_S} - (QI)_j\right] - \Theta_j(X_j) + 1 - \gamma_j.$$

Here, the †–notation indicates differentiation with respect to $(QI)_j$ (but keeping I_j fixed). Rewriting the bracketed expression in terms of X_j yields

$$H_j(I)^\dagger = \Theta_j'(X_j)X_j^\dagger\left[\frac{\beta_j}{X_j} - 1\right](QI)_j - \Theta_j(X_j) + 1 - \gamma_j.$$

We now use the inequalities $\Theta_j'(X_j) \ge 0$, $X_j^\dagger > 0$, and $\Theta_j(X_j) \le X_j < \beta_j$ to obtain

$$H_j(I)^\dagger > 1 - \beta_j - \gamma_j \ge 0.$$

Consequently, $H_j(I)$ is a strictly increasing function of $(QI)_j$ when I_j is held fixed. Finally, (3.4b) implies that

$$H_j(I^{\circ -}) = H_j(I^{\circ +}), \quad j \in K^+.$$

It follows from the remarks above that $(QI^{\circ -})_j = (QI^{\circ +})_j$ for $j \in K^+$. That is,

$$\sum_{k \in K^-} q_{jk}\big(I_k^{\circ -} - I_k^{\circ +}\big) = 0, \quad j \in K^+.$$

We only sum over $k \in K^-$ because $I_k^{\circ -} = I_k^{\circ +}$ for $k \in K^+$. Since $I_k^{\circ -} < I_k^{\circ +}$ for $k \in K^-$, it must be that $q_{jk} = 0$ whenever $j \in K^+$ and $k \in K^-$. But this contradicts the assumption that Q is irreducible. We conclude that either $I^{\circ -} < I^{\circ +}$ or $I^{\circ -} \equiv I^{\circ +}$. □

Lemma 3.9. *If $I^{\circ *}$ and $I^{\circ **}$ are positive then $I^{\circ *} \equiv I^{\circ **}$.*

Proof. We argue by contradiction. Suppose that $I^{\circ *}$ and $I^{\circ **}$ are positive with $I^{\circ *} \not\equiv I^{\circ **}$. Then $I^{\circ *}, I^{\circ **} \in [I^{\min}, I^{\max}]$ by Lemma 3.7. Since $I^{\circ *} \not\equiv I^{\circ **}$, it follows that $I^{\min} \not\equiv I^{\max}$. We conclude from the relation $I^{\min} \le I^{\max}$ and Lemma 3.8 that $I^{\min} < I^{\max}$. So, without loss of generality,

we may assume that $I^{\circ *} < I^{\circ **}$, for otherwise we may replace $I^{\circ *}$ with $I^{\min}$ and $I^{\circ **}$ with $I^{\max}$. Recall that

$$H_j(I^{\circ *}) = I_j^{\circ *}, \quad j \in \Omega,$$
$$H_j(I^{\circ **}) = I_j^{\circ **}, \quad j \in \Omega.$$

We multiply the first equation by $(QI^{\circ **})_j$ and the second equation by $(QI^{\circ *})_j$, subtract the resulting equations, and then sum over all $j \in \Omega$ to get

$$\sum_{j\in\Omega} \Big[H_j(I^{\circ *})(QI^{\circ **})_j - H_j(I^{\circ **})(QI^{\circ *})_j \Big] = \sum_{j\in\Omega} \Big[(QI^{\circ **})_j I_j^{\circ *} - (QI^{\circ *})_j I_j^{\circ **} \Big].$$

The symmetry of Q implies that the right-hand side vanishes. It suffices to show that the left-hand side (which we denote by Ψ) is actually positive. Observe from (3.3) that

$$\Psi = \frac{d_I}{d_S} \sum_{j\in\Omega} \big[\psi_j(I^{\circ *}) - \psi_j(I^{\circ **}) \big] (QI^{\circ *})_j (QI^{\circ **})_j,$$

where

$$\psi_j(I) = \Theta_j(X_j) \frac{1-(PI)_j}{(QI)_j} \quad \text{and} \quad X_j = X_j(I) = \frac{\beta_j d_S (QI)_j}{d_I + (d_S - d_I) I_j}.$$

Since $(QI^{\circ *})_j$ and $(QI^{\circ **})_j$ are both positive, it suffices to show that $\psi_j(I^{\circ *}) \geq \psi_j(I^{\circ **})$ for all $j \in \Omega$ and that the inequality is strict for at least one such j. If $\beta_j = 0$ then it is clear from the properties of Θ_j that $\psi_j(I^{\circ *}) = \psi_j(I^{\circ **}) = 0$. Thus, we may assume that $\beta_j > 0$. Furthermore, since H^+ is nonempty, there is at least one such $j \in \Omega$ with this property. We differentiate ψ_j with respect to I_k to get

$$\psi_j(I)^\dagger = \Theta_j'(X_j) X_j^\dagger \frac{1-(PI)_j}{(QI)_j} + \Theta_j(X_j) \left[\frac{1-(PI)_j}{(QI)_j} \right]^\dagger,$$

where the $\dagger$–notation indicates differentiation with respect to I_k and we restrict attention to $I \in [I^{\min}, I^{\max}]$. Observe that $X_j > 0$, $0 \leq \Theta_j'(X_j) \leq \Theta_j(X_j)/X_j$, and $X_j^\dagger \geq 0$. Therefore,

$$\psi_j(I)^\dagger \leq \Theta_j(X_j) \left(\frac{X_j^\dagger}{X_j} \times \frac{1-(PI)_j}{(QI)_j} + \left[\frac{1-(PI)_j}{(QI)_j} \right]^\dagger \right).$$

Since $\Theta_j(X_j) > 0$, it suffices to show that the quantity in parentheses (which we denote by Γ_{jk}) is nonpositive for all k and negative for some k. It can be shown that

$$\frac{X_j^\dagger}{X_j} = \frac{(QI)_j^\dagger}{(QI)_j} - \frac{\big[d_I + (d_S - d_I) I_j \big]^\dagger}{d_I + (d_S - d_I) I_j}$$

and

$$\left[\frac{1-(PI)_j}{(QI)_j} \right]^\dagger = \left(\frac{\big[1-(PI)_j\big]^\dagger}{1-(PI)_j} - \frac{(QI)_j^\dagger}{(QI)_j} \right) \frac{1-(PI)_j}{(QI)_j}.$$

Therefore,

$$\Gamma_{jk} = \left(\frac{\big[1-(PI)_j\big]^\dagger}{1-(PI)_j} - \frac{\big[d_I + (d_S - d_I) I_j \big]^\dagger}{d_I + (d_S - d_I) I_j} \right) \frac{1-(PI)_j}{(QI)_j}.$$

We can write this expression in turn as

$$\Gamma_{jk} = \left[\frac{1-(PI)_j}{d_I+(d_S-d_I)I_j}\right]^\dagger \frac{d_I+(d_S-d_I)I_j}{(QI)_j}.$$

Finally, if $k \neq j$ then $\Gamma_{jk} = -p_{jk}/(QI)_j$ is nonpositive, and it is negative for some k because P is irreducible. However, if $k = j$ then Lemma 3.4 implies that

$$\Gamma_{jj} = \left[\frac{d_I+(d_S-d_I)I_j-d_S(QI)_j}{d_I+(d_S-d_I)I_j}\right]^\dagger \frac{d_I+(d_S-d_I)I_j}{d_I(QI)_j}.$$

Again, we conclude that Γ_{jj} is nonpositive. Thus, if $I^{\circ *}$ and $I^{\circ **}$ are positive then $I^{\circ *} \equiv I^{\circ **}$. □

Lemmas 3.7 and 3.9 imply that (3.4b) has a unique positive solution given by $I^\circ = I^{\min} = I^{\max}$. We conclude from Lemma 3.6 that I° is the only nonnegative solution of (3.4b) satisfying $I^\circ \not\equiv 0$. We have completed the proof of Lemma 3.3. The next result follows from Lemmas 3.1–3.3 and (3.2).

Lemma 3.10. *Eq.* (1.10) *has a nonnegative solution* $(\hat{S}, \hat{I}, \tilde{S}, \tilde{I})$ *which satisfies* $\tilde{I} \not\equiv 0$. *Furthermore, this solution (with* $\tilde{I} \not\equiv 0$*) is unique, it is positive, and it is given by*

$$\hat{S} = \kappa S^\circ, \quad \hat{I} = (\kappa/d_I)I^\circ, \quad \tilde{S} = \kappa P S^\circ, \quad \textit{and} \quad \tilde{I} = (\kappa/d_I)QI^\circ$$

where κ *is as in* (3.4e).

We have shown that when $\lambda^* = \rho(M) > 1$ then an EE exists, it is unique, and it satisfies $\tilde{I} > 0$. In the next section, we consider the asymptotic behavior of the EE as $d_S \to 0$.

4 Asymptotic Behavior of the Endemic Equilibrium

Throughout this section, we still assume that (A1)–(A5) and (B1)–(B2) hold, that N is fixed, and that $\lambda^* > 1$ so that an EE exists which is both unique and positive. Theorem 4 follows from Lemmas 4.1 and 4.4–4.8.

4.1 The limiting DFE

Observe from (1.8), (1.10), and (3.1a) that $\hat{S}$, $\hat{I}$, $\tilde{S}$, $\tilde{I}$, and κ are all functions of d_S. First, we determine the asymptotic behavior of $\hat{I}$, $\tilde{I}$, and κ as $d_S \to 0$ (i.e., $P \to \mathcal{I}_n$).

Lemma 4.1. *If* $d_S \to 0$ *then* $\hat{I} \to 0$, $\tilde{I} \to 0$, *and* $\kappa \to 0$.

Proof. Recall from Lemma 3.10 and (1.10e) that $\hat{S}_j, \hat{I}_j, \tilde{S}_j, \tilde{I}_j \in (0, N)$ for $j \in \Omega$. Let $\hat{S}^*$ be a limit point of $\hat{S}$ as $d_S \to 0$. Passing to a subsequence each time if necessary, let $\hat{I}^*$, $\tilde{S}^*$, and $\tilde{I}^*$ in turn be limit points of $\hat{I}$, $\tilde{S}$, and $\tilde{I}$, respectively, as $d_S \to 0$. Then $\hat{S}^*_j, \hat{I}^*_j, \tilde{S}^*_j, \tilde{I}^*_j \in [0, N]$ for $j \in \Omega$. In view of (1.8) with $d_S = 0$ and (1.10), $\tilde{S}^* = \hat{S}^*$ and $\tilde{I}^* = Q\hat{I}^*$. Observe from (1.10a), (1.2), the positivity of $\tilde{S}_j + \tilde{I}_j$, and the properties of Θ_j that if $j \in \Omega$ then

$$\begin{aligned}\hat{S}_j &= \left[1-\Theta_j\left(\frac{\beta_j\tilde{I}_j}{\tilde{S}_j+\tilde{I}_j}\right)\right]\tilde{S}_j + \gamma_j\tilde{I}_j \\ &\geq \left[1-\frac{\beta_j\tilde{I}_j}{\tilde{S}_j+\tilde{I}_j}\right]\tilde{S}_j + \gamma_j\tilde{I}_j\end{aligned}$$

$$\geq \tilde{S}_j - (\beta_j - \gamma_j)\tilde{I}_j.$$

It follows from taking limits that $\tilde{S}_j^* = \hat{S}_j^* \geq \tilde{S}_j^* - (\beta_j - \gamma_j)\tilde{I}_j^*$. We conclude that $(\beta_j - \gamma_j)\tilde{I}_j^* \geq 0$ for $j \in \Omega$, and in particular that $\tilde{I}_j^* = 0$ for $j \in H^-$. Fix $j \in H^-$. Then $(Q\hat{I}^*)_j = \tilde{I}_j^* = 0$. Consequently, $\hat{I}_k^* = 0$ for every $k \in K_j = \{k \in \Omega : q_{jk} > 0\}$. Fix $k \in K_j$. Since $\hat{S}_k, \hat{I}_k, \hat{S}_k^* \in [0, N]$ and $\hat{I}_k^* = 0$, we obtain from (3.1a) that $\kappa \to 0$ as $d_S \to 0$. Now let $j \in \Omega$ be arbitrary. Since $\hat{S}_j, \hat{I}_j, \hat{S}_j^* \in [0, N]$ and $\kappa \to 0$ as $d_S \to 0$, (3.1a) now implies that $\hat{I}_j^* = 0$. That is, $\hat{I}^* = 0$. Finally, $\tilde{I}^* = Q\hat{I}^* = 0$. □

So that we may also determine the asymptotic behavior of $\hat{S}$ and $\tilde{S}$, we first consider I° in (3.2) as a function of d_S.

Lemma 4.2. *I_j° is a monotone decreasing function of d_S for every $j \in \Omega$.*

Proof. Suppose that $0 < d_S^- < d_S^+$ and let $I^{\circ -}$ and $I^{\circ +}$ be corresponding solutions to (3.4b) with $I^{\circ -}, I^{\circ +} \in (0, 1_n)$. Then

$$H_j(I^{\circ -}; d_S^-) = I_j^{\circ -}, \quad j \in \Omega, \tag{4.1a}$$

$$H_j(I^{\circ +}; d_S^+) = I_j^{\circ +}, \quad j \in \Omega, \tag{4.1b}$$

where $H_j(I; d_S)$ is as in (3.3). We now show that $\partial H_j / \partial d_S \leq 0$ for $I \in (0, 1_n)$ and $j \in \Omega$. If $\beta_j = 0$ then the result is immediate, so we may assume for the time being that $\beta_j > 0$. Since P depends on d_S in (1.8), the product rule for three terms yields

$$\frac{\partial H_j}{\partial d_S} = \frac{d_I}{d_S}\left(-\frac{1}{d_S}\Theta_j(X)\big[1 - (PI)_j\big] + \Theta_j'(X)X^\dagger\big[1 - (PI)_j\big] + \Theta_j(X)\big[1 - (PI)_j\big]^\dagger\right),$$

where

$$X = X(d_S) = \frac{\beta_j d_S (QI)_j}{d_I + (d_S - d_I)I_j}$$

and the †–notation indicates differentiation with respect to d_S. Since $X > 0$ and $X^\dagger > 0$, it follows from the properties of Θ_j that

$$\frac{\partial H_j}{\partial d_S} \leq \frac{\Theta_j(X)}{d_S}\left(\left[\frac{X^\dagger}{X} - \frac{1}{d_S}\right]\big[d_I - d_I(PI)_j\big] + \big[d_I - d_I(PI)_j\big]^\dagger\right).$$

As $\Theta_j(X) > 0$, it suffices to show that the parenthetical quantity, which we denote by Ψ, is nonpositive. First, we use the product rule to directly compute

$$\frac{X^\dagger}{X} = \frac{1}{d_S} - \frac{I_j}{d_I + (d_S - d_I)I_j}$$

and then apply Lemma 3.4 to get

$$\Psi = -\frac{I_j}{d_I + (d_S - d_I)I_j}\big[d_I + (d_S - d_I)I_j - d_S(QI)_j\big] + I_j - (QI)_j.$$

A quick calculation shows that

$$\Psi = -\frac{d_I(1 - I_j)(QI)_j}{d_I + (d_S - d_I)I_j} \leq 0.$$

Hence, $\partial H_j/\partial d_S \le 0$ for $I \in (0, 1_n)$ and $j \in \Omega$. We now obtain from (4.1b) with d_S^- in place of d_S^+ that

$$H_j(I^{\circ +}; d_S^-) \ge I_j^{\circ +}, \quad j \in \Omega.$$

Thus, $I^{\circ +}$ is a sub-solution of (4.1a). Again, 1_n is a super-solution of (4.1a) by Lemma 3.5 (b). Also, $I^{\circ +} < 1_n$. By an argument similar to the one given in Sect. 3, $I^{\circ -} \in [I^{\circ +}, 1_n]$. We conclude that $I^{\circ -} \ge I^{\circ +}$. □

Recall from Lemma 3.3 that $0 < I_j^\circ < 1$ for each $j \in \Omega$. The above lemma implies that for every $j \in \Omega$, there exists some I_j^* such that

$$I_j^\circ \to I_j^* \text{ as } d_S \to 0 \quad \text{and} \quad 0 < I_j^* \le 1. \tag{4.2}$$

Let $I^* = (I_j^*)$. It remains to establish conditions under which $0 < I_j^* < 1$ or $I_j^* = 1$. Let

$$J^- = \{j \in \Omega : 0 < I_j^* < 1\} \quad \text{and} \quad J^+ = \{j \in \Omega : I_j^* = 1\}.$$

Observe that $J^- \cup J^+ = \Omega$. The next result characterizes the relative values of $\nu_j(QI^*)_j$ and I_j^* for different $j \in \Omega$. Recall that $\nu_j = 1 + \beta_j - \gamma_j$.

Lemma 4.3. *The following statements hold:*

(a) *If $j \in J^-$ then $\nu_j(QI^*)_j = I_j^*$;*

(b) *If $j \in J^+$ then $\nu_j(QI^*)_j \ge 1$.*

Proof.

(a) Let $j \in J^-$, so that $I_j^\circ \to I_j^* \in (0, 1)$ as $d_S \to 0$. Observe from (3.4b) and (3.3) that

$$H_j(I^\circ) = \frac{d_I}{d_S}\Theta_j(X)\big[1 - (PI^\circ)_j\big] + (1 - \gamma_j)(QI^\circ)_j = I_j^\circ, \tag{4.3}$$

where

$$X = X(d_S) = \frac{\beta_j d_S (QI^\circ)_j}{d_I + (d_S - d_I)I_j^\circ}. \tag{4.4}$$

We examine the asymptotic behavior of various terms in (4.3) and (4.4) as $d_S \to 0$, and we may assume that $\beta_j > 0$, for otherwise the result is immediate. First, since

$$\beta_j d_S (QI^\circ)_j \to 0 \quad \text{and} \quad d_I + (d_S - d_I)I_j^\circ \to d_I(1 - I_j^*) > 0,$$

we conclude that $X \to 0$ and by continuity $\Theta_j(X) \to \Theta_j(0) = 0$. Next, we see from (4.4) that

$$\frac{\Theta_j(X)}{d_S} = \frac{\Theta_j(X)}{X} \cdot \frac{X}{d_S} = \frac{\Theta_j(X)}{X} \cdot \frac{\beta_j (QI^\circ)_j}{d_I + (d_S - d_I)I_j^\circ} \to \frac{\beta_j (QI^*)_j}{d_I(1 - I_j^*)}.$$

In passing to the limit, we used L'Hospital's Rule and the relation $\Theta_j'(0) = 1$. Finally, $P \to \mathcal{I}_n$ in (1.8) implies that $(PI^\circ)_j \to I_j^*$. Taking limits in (4.3) now produces the desired result.

(b) Let $j \in J^+$, so that $I_j^\circ \to I_j^* = 1$ as $d_S \to 0$. Again, (3.4b) and (3.3) imply that (4.3) and (4.4) hold. The properties of Θ_j, (4.4), and Lemma 3.4 together imply that

$$\frac{\Theta_j(X)}{d_S} \le \frac{X}{d_S} = \frac{\beta_j(QI^\circ)_j}{d_I + (d_S - d_I)I_j^\circ} \le \frac{\beta_j(QI^\circ)_j}{d_I + (d_S - d_I)I_j^\circ - d_S(QI^\circ)_j} = \frac{\beta_j(QI^\circ)_j}{d_I\big[1 - (PI^\circ)_j\big]}.$$

Therefore,

$$\frac{d_I}{d_S}\Theta_j(X)\big[1 - (PI^\circ)_j\big] \le \beta_j(QI^\circ)_j.$$

Upon substituting this inequality into (4.3) and then letting $d_S \to 0$, we obtain the desired result. □

The set J^- is nonempty.

Lemma 4.4. $H^- \subseteq J^-$.

Proof. We argue by contradiction. Suppose that there exists some $j \in H^- \cap J^+$. Then $\nu_j < 1$ and $(QI^*)_j \le 1$ by (4.2) and the properties of Q. That is, $\nu_j(QI^*)_j < 1$. But this inequality contradicts Lemma 4.3 (b). We conclude that $H^- \cap J^+$ is empty, and therefore that $H^- \subseteq J^-$. □

We are now in a position to determine the asymptotic behavior of $\hat{S}$ and $\tilde{S}$.

Lemma 4.5. *The following statements hold:*

(a) $\kappa/d_S \to N^* \stackrel{\text{def}}{=} N/\sum_{j\in\Omega}(1 - I_j^*)$ *as* $d_S \to 0$;

(b) $\hat{S} \to S^*$ *and* $\tilde{S} \to S^*$ *as* $d_S \to 0$, *where* $S^* = (S_j^*)$ *and* $S_j^* \stackrel{\text{def}}{=} (1 - I_j^*)N^*$ *for* $j \in \Omega$;

(c) $S^* \ge 0$ *and* $\sum_{j\in\Omega} S_j^* = N$.

Proof.

(a) Eqs. (1.10e), (3.1a), and (3.2) imply that

$$N = \sum_{j\in\Omega} \frac{\kappa - d_I\hat{I}_j}{d_S} + \sum_{j\in\Omega} \hat{I}_j = \frac{\kappa}{d_S}\sum_{j\in\Omega}(1 - I_j^\circ) + \sum_{j\in\Omega} \hat{I}_j.$$

Lemma 4.1 and (4.2) imply that

$$\frac{\kappa}{d_S} \to \frac{N}{\sum_{j\in\Omega}(1 - I_j^*)} \quad \text{as} \quad d_S \to 0.$$

The denominator of the limit is positive because J^- is nonempty by Lemma 4.4.

(b) Again, (3.1a) and (3.2) imply that

$$\hat{S}_j = \frac{\kappa - d_I\hat{I}_j}{d_S} = (1 - I_j^\circ)\frac{\kappa}{d_S}.$$

Eq. (4.2) and part (a) imply that $\hat{S}_j \to (1 - I_j^*)N^*$ as $d_S \to 0$. Also, $\tilde{S} = P\hat{S} \to S^*$ as $d_S \to 0$ because $P \to \mathcal{I}_n$ in (1.8) and $\hat{S} \to S^*$.

(c) This part follows immediately from parts (a) and (b), the positivity of N^*, and (4.2). □

4.2 The limiting DFE on high-risk patches

Observe from Lemma 4.5 (b) that $S^* > 0$ on J^- and $S^* = 0$ on J^+. We know from Lemma 4.4 that J^- is nonempty. Next we show that J^+ is also nonempty.

Lemma 4.6. *J^+ is nonempty.*

Proof. We argue by contradiction. Suppose that J^+ is empty, that is, $\Omega = J^- \cup J^+ = J^-$. Lemma 4.3 (a) implies that

$$(1 + \beta_j - \gamma_j)(QI^*)_j = I_j^*, \quad j \in \Omega.$$

Thus $(\lambda, \varphi) = (1, I^*)$ satisfies (2.3). Since $\lambda^* > 1$, Lemma 2.4 implies that $I_j^* \leq 0$ for some $j \in \Omega$. But this result contradicts (4.2). We conclude that J^+ is nonempty. □

Next, we determine a condition under which J^+ is maximal.

Lemma 4.7. *If (1.13) is satisfied for all $j \in H^-$ then $J^+ = H^+$.*

Proof. We argue by contradiction. Suppose that condition (1.13) holds on H^- but that $J^+ \neq H^+$. In view of Lemma 4.4, it must be that J^+ is a proper subset of H^+ and therefore that $H^+ \cap J^-$ is nonempty. Let $j \in H^-$ and $k \in H^+ \cap J^-$. Without loss of generality, we may assume that j and k are chosen so that $I_j^* = \min\{I_m^* : m \in H^-\}$ and $I_k^* = \min\{I_m^* : m \in H^+ \cap J^-\}$. Since $I_m^* < 1$ for $m \in H^+ \cap J^-$ and $I_m^* = 1$ for $m \in H^+ \cap J^+$, it must be that $I_k^* = \min\{I_m^* : m \in H^+\}$. Lemma 4.3 (a) implies that

$$\nu_j(QI^*)_j = I_j^*, \tag{4.5}$$

$$\nu_k(QI^*)_k = I_k^*. \tag{4.6}$$

Since $\Omega = H^- \cup H^+$, we obtain from (4.5) that

$$\nu_j \sum_{m \in H^-} q_{jm}(I_m^* - I_j^*) + \nu_j \sum_{m \in H^+} q_{jm} I_m^* = (1 - \nu_j Q_j^-) I_j^*.$$

The minimality of I_j^* over H^- and I_k^* over H^+ implies that

$$\nu_j Q_j^+ I_k^* \leq (1 - \nu_j Q_j^-) I_j^*. \tag{4.7}$$

Similarly, we obtain from (4.6) that

$$\nu_k \sum_{m \in H^-} q_{km} I_m^* + \nu_k \sum_{m \in H^+} q_{km}(I_m^* - I_k^*) = (1 - \nu_k Q_k^+) I_k^*.$$

The minimality of I_j^* over H^- and I_k^* over H^+ implies that

$$\nu_k Q_k^- I_j^* \leq (1 - \nu_k Q_k^+) I_k^*. \tag{4.8}$$

If $Q_k^- = 0$, then (4.8) with $Q_k^+ = 1$ and $I_k^* > 0$ implies that $\nu_k \leq 1$, a contradiction because $k \in H^+$. Therefore, Q_k^- must be positive. Since $\nu_k > 1$, we also see from (4.8) that

$$\nu_k Q_k^- I_j^* \leq (1 - Q_k^+) I_k^* = Q_k^- I_k^*.$$

Since $Q_k^- > 0$, it follows that

$$\nu_k I_j^* \leq I_k^*. \tag{4.9}$$

We multiply corresponding sides of (4.7) and (4.9) together and simplify to get

$$\nu_j(Q_j^- + \nu_k Q_j^+) \leq 1.$$

But this inequality contradicts (1.13). We conclude that $J^+ = H^+$. □

Finally, we determine a condition under which J^+ is a proper subset of H^+.

Lemma 4.8. *If* (1.14) *is satisfied for some* $j \in H^+$ *then* $j \in J^-$.

Proof. We argue by contradiction. Suppose that condition (1.14) holds for some $j \in H^+ \cap J^+$. Choose $k \in H^-$ such that $I_k^* = \max\{I_m^* : m \in H^-\} < 1$. Lemma 4.3 implies that

$$\nu_j (QI^*)_j \geq 1, \tag{4.10}$$

$$\nu_k (QI^*)_k = I_k^*. \tag{4.11}$$

We rewrite (4.10) as

$$\nu_j \sum_{m \in H^+} q_{jm} I_m^* + \nu_j \sum_{m \in H^-} q_{jm} I_m^* \geq 1.$$

The upper bound of 1 on I_m^* and the maximality of I_k^* over H^- implies that

$$\nu_j (Q_j^+ + I_k^* Q_j^-) \geq 1.$$

Also, since $(QI^*)_k \leq 1$, we obtain from (4.11) that $I_k^* \leq \nu_k$. It follows from these remarks that

$$\nu_j (Q_j^+ + \nu_k Q_j^-) \geq 1.$$

But this inequality contradicts (1.14). We conclude that (1.14) cannot hold for $j \in H^+ \cap J^+$. □

5 Discussion

We now comment on some of the model's assumptions, the theoretical results that were obtained, and some open problems that are connected with this research.

5.1 Model assumptions

The generality of the infection function Θ_j accommodates many forms of frequency-dependent infection, including those that arise from Poisson [3, 4, 18], logarithmic [22], and binomial processes [8]. Its concavity assumption, which we exploit repeatedly, implies that each additional infection results in fewer additional infective contacts, on average, in a time step. Since we allow the infection function Θ_j to differ between patches, our results apply to a wide range of diseases that can not only invade and persist in spatially heterogeneous habitats but can also do so in such a way that different locations engender different infection processes. For example, susceptible individuals in some patches may have m_j infective encounters each time step while individuals in other patches may experience infective encounters that are exponentially distributed in time.

We introduce the concept of patch virulence ($\nu_j = 1 + \beta_j - \gamma_j$) to aid in distinguishing between patches which are low-risk ($\beta_j < \gamma_j$) from those which are high-risk ($\beta_j > \gamma_j$). Surprisingly, the harmonic mean of these patch virulences ($\mathcal{B}$) proves itself to be a useful quantity, and so we characterize habitats to be low-risk or high-risk depending on the magnitude of $\mathcal{B}$. This definition is a marked departure from what we did for our two analogous continuous-time SI-models [1, 2], where we instead characterized habitats as being low-risk or high-risk depending on whether the spatial average of the disease transmission rates is less than or greater than the spatial average of the disease recovery rates.

The role of the baseline dispersal matrix R is to create a one-dimensional family of dispersal matrices $\{R(\alpha)\}$ whose members approach the matrix representing no dispersal ($\mathcal{I}_n$) as the

diffusion coefficient α tends to zero. The intended effect is to restrict the movement of some subpopulation (either susceptible or infected individuals) in such a way that its members experience, in all patches, correlated reductions in mobility. We suspect that restricting the movement of a subpopulation in some other way would also lead to similar results, as long as members of the resulting family of matrices are all dispersal matrices (i.e., doubly stochastic and irreducible) that tend to $\mathcal{I}_n$ in some fixed manner.

5.2 Our results

In Theorem 1 we show that a spatially homogeneous, unique DFE exists. That is, every patch supports the same number of individuals, all of whom are susceptible. Furthermore, when the basic reproduction number $\mathcal{R}_0$ is less than one, then the disease always disappears, but when $\mathcal{R}_0$ is greater than one then the disease can persist, even in arbitrarily small neighborhoods of the DFE. The proof of this theorem is given in Sect. 2, where we also outline the general procedure for computing $\mathcal{R}_0$ as the spectral radius of the next generation matrix [6, 13, 15, 16, 22, 23, 29].

Although it is usually not possible to obtain an explicit expression for $\mathcal{R}_0$, we can do so at least for some special cases. For example, it can be shown by a direct calculation that when there are $n = 2$ patches, and they both have the same disease transmission rate β and recovery rate γ, then $\mathcal{R}_0 = \beta/\gamma$. In this case, $\mathcal{R}_0$ agrees with the basic reproduction for a single patch. Suppose instead that there are n patches and that the infected dispersal matrix Q tends to the identity matrix $\mathcal{I}_n$. Then the next generation matrix $\mathcal{M} \to \text{diag}_{j\in\Omega}\,(\beta_j/\gamma_j)$. Consequently, its spectral radius $\mathcal{R}_0 = \rho(\mathcal{M}) \to \max_{j\in\Omega}\{\beta_j/\gamma_j\}$. That is, as the mobility of infected individuals tends to zero, so that they lose their ability to disperse, then the disease will persist and it will do so only on the high-risk patches. This result parallels our Lemma 2.5 (b) in which we show that the principle eigenvalue λ^* of the virulence weighted infected dispersal matrix (M) tends to the maximum of the virulences as $d_I \to 0$ and it also agrees with related results obtained in [1, 2]. Finally, suppose that every entry of the infected dispersal matrix Q tends to $1/n$, so that infected individuals show no dispersal preferences. In this case, the limiting next generation matrix $\mathcal{M}$ has a zero eigenvalue with multiplicity $n-1$ and a simple positive eigenvalue $\mathcal{R}_0 = \sum_j \beta_j / \sum_j \gamma_j$, the ratio of the average patch transmission rate to the average patch recovery rate. All of these special cases for $\mathcal{R}_0$ agree with corresponding values computed in our analogous continuous-time SI-models [1, 2]. We remark, however, that not all of these special cases conform to our assumptions in (A3) and (A5).

Our Theorem 2 implies that the disease can persist in low-risk habitats only if the mobility of infected individuals lies below a certain threshold value. Recall that in low-risk habitats, there are relatively few patches which can sustain the disease. Therefore, newly infected individuals can produce a sufficient number of secondary infections to sustain the disease only if they do not recover too quickly from the disease. This then requires them to remain in high-risk patches for as long as possible. In terms of ecological source-sink theory, the disease can persist only if carriers of the disease spend a sufficiently long period of time in high-risk patches (the sources) before moving to low-risk patches (the sinks). However, no such result exists for high-risk habitats because there are a sufficient number of high-risk patches to sustain the disease no matter how quickly infected individuals stray from them to low-risk patches.

In Theorem 3 we show that when the infection can persist ($\mathcal{R}_0 > 1$) then an EE exists, it is unique, and it supports a positive number of infected individuals on every patch. As in [1, 2], we successfully applied the methods of monotone dynamical systems theory to establish this result. For most ecological disease models, establishing uniqueness of the EE is difficult and consequently it is seldom done.

Our Theorem 4 describes the asymptotic profile of the EE as d_S tends to zero. The message here is that decreasing the mobility of susceptible individuals ($P \to \mathcal{I}$) can result in the elimination of the disease at equilibrium. Unlike in Theorem 2, this result holds for both low-risk and high-risk habitats. Although the limiting DFE is no longer spatially homogeneous, we do know that all low-risk patches are necessarily nonempty and that at least one high-risk patch is not. We then provide a sufficient condition under which all of the remaining high-risk patches are also empty, and a second sufficient condition under which some of them are not. These conditions do not, however, cover all possible cases. In certain ways, our ability to characterize the asymptotic profile of the EE crucially depends on its uniqueness. For example, in our proof of Lemma 4.2, we make use of this uniqueness to show that I_j° is monotone with respect to d_S. This result in turn allows us to prove many of the subsequent results that appear in Sect. 4.

5.3 Major open questions

When the basic reproduction number is less than one ($\mathcal{R}_0 < 1$), the proof of Lemma 2.2 shows only that the infection always disappears, i.e., $I(t) \to 0$ as $t \to \infty$. It does not, however, establish whether the DFE is globally asymptotically stable. It can be shown that if the susceptible dispersal matrix P is periodic (i.e., $P^k = P$ for some $k > 1$), then in addition to the unique DFE, there also exist disease-free periodic orbits and the periods of these orbits divide the number of patches n. Of course, the mere existence of these periodic orbits destroys any chance that the DFE can be globally asymptotically stable, however, we believe that even in this case, the DFE will still be locally asymptotically stable, and the proof of such a result may require methods from asymptotically autonomous difference equations. When they exist, it would be very interesting to determine conditions under which the disease-free periodic orbits are stable. For cases in which P is aperiodic, it is an open question as to whether the DFE is locally or globally attracting. This situation occurs, for example, when P is controlled by a baseline dispersal matrix R and the parameter d_S is very small. Recent results from [6] do not apply directly to this situation because $\rho(P) = 1$, and so determining the stability of the DFE remains an open problem.

As in our two previous papers [1, 2], the stability of the EE also remains a mystery. We believe that it is always locally asymptotically stable. However, as with the DFE, there arises the possibility that when P and Q are both periodic then endemic periodic orbits may also exist. The existence of these periodic orbits would preclude global asymptotic stability for the EE. However, for cases in which P or Q is aperiodic, then we believe that the EE will always be globally asymptotically stable whenever it exists. As a stepping stone to increasing our understanding of this matter, we suggest examining the stability of the EE for $n = 2$ patches when d_S is sufficiently small and positive. Since the EE closely approximates a DFE in this case, some progress may be possible. When an endemic periodic orbit does exist, it would be again interesting to determine conditions under which it could be unique or stable.

It is not clear to us why some of our results require that the activity $(\beta_j + \gamma_j)$ of every patch not exceed one. Perhaps this restriction can be lifted in future work.

As mentioned earlier, we cannot predict the structure of the limiting DFE in all cases, even when we restrict attention to one-parameter families of dispersal matrices. Since there are many other mechanisms by which susceptible dispersal can become increasingly restricted, we would like to know in general whether the results of Theorem 4 can be extended to arbitrary sequences of dispersal matrices $\{P_k\}$ which satisfy $\|P_k - \mathcal{I}_n\| \to 0$ as $k \to \infty$ for some suitable matrix norm.

Our three papers ([1, 2] and this paper) illustrate the effects of dispersal, spatial heterogeneity, and habitat connectivity on various dynamical and equilibrium properties of infectious diseases. In each case, local differences in disease transmission and recovery rates characterize whether

regions are low-risk or high-risk, and these differences collectively determine whether the spatial domain, or habitat, is low-risk or high-risk. In low-risk habitats, the disease persists only when the mobility of infected individuals lies below some threshold value, but for high-risk habitats, the disease always persists. When the disease does persist, then there exists an EE which is unique and positive everywhere. This EE tends to a spatially inhomogeneous DFE as the mobility of susceptible individuals tends to zero. The limiting DFE is nonempty on all low-risk regions and it is empty on some part of the high-risk region. Sufficient conditions for whether the limiting DFE is empty or not on the remainder of the high-risk region can be given in terms of disease transmission and recovery rates, habitat connectivity, and the infected movement rate. The results we obtain in each case closely parallel each other, yet the underlying mathematical formulations are quite distinct and they often require us to resort to considerably different mathematical techniques to analyze them.

Acknowledgements

The authors thank Carlos Castillo-Chavez for helpful comments and the Mathematical Biosciences Institute at The Ohio State University for support of this project. LJSA acknowledges support from NSF grants DMS-0201105 and DMS-0718302 and YL acknowledges support from NSF grant DMS-0615845. This material is based upon work supported by the National Science Foundation under Agreement No. 0112050.

References

[1] Allen, L.J.S., B.M. Bolker, Y. Lou, and A.L. Nevai. 2007. Asymptotic profile of the steady states for an SIS epidemic patch model. *SIAM J. Appl. Math.* **67**: 1283–1309.

[2] Allen, L.J.S., B.M. Bolker, Y. Lou, and A.L. Nevai. 2008. Asymptotic profile of the steady states for an SIS epidemic reaction-diffusion model. *Discr. Cont. Dyn. Sys. A.* In press.

[3] Allen, L.J.S. and A. Burgin. 2000. Comparison of deterministic and stochastic SIS and SIR models in discrete time. *Math. Biosci.* **163**: 1-33.

[4] Allen, L.J.S., M.A. Jones, and C.F. Martin. 1991. A discrete-time model with vaccination for a measles epidemic. *Math. Biosci.* **105**: 111–131.

[5] Allen, L.J.S., N. Kirupaharan, and S. M. Wilson. 2004. SIS epidemic models with multiple pathogen strains. *J. Difference Equ. Appl.* **10**: 53–75.

[6] Allen, L.J.S. and P. van den Driessche. The basic reproduction number in some discrete-time epidemic models. Manuscript.

[7] Arino, J., R. Jordan, and P. van den Driessche. 2007. Quarantine in a multi-species epidemic model with spatial dynamics. *Math. Biosci.* **206**: 46–60.

[8] Caraco, T., M. Duryea, G. Gardner, W. Maniatty, and B.K. Szymanski. 1998. Host spatial heterogeneity and extinction of an SIS epidemic. *J. Theor. Biol.* **192**: 351–361.

[9] Carrillo, C. and P. Fife. 2005. Spatial effects in discrete generation population models. *J. Math. Biol.* **50**: 161–188.

[10] Castillo-Chavez, C. and A.–A. Yakubu. 2001. Dispersal, disease and life-history evolution. *Math. Biosci.* **173**: 35–53.

[11] Castillo-Chavez, C. and A.–A. Yakubu. 2002. Intraspecific competition, dispersal, and disease dynamics in discrete-time patchy environments. In: C. Castillo-Chavez with S. Blower, P. van den Driessche, D. Kirschner and A.-A. Yakubu (Eds) *Mathematical Approaches for Emerging and Reemerging Infectious Diseases An Introduction*, Springer-Verlag, New York. pp. 165–181.

[12] Cushing, J.M. 1998. An Introduction to Structured Population Dynamics. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia.

[13] de Jong, M.C.M., O. Diekmann, and J.A.P. Heesterbeek. 1994. The computation of R_0 for discrete-time epidemic models with dynamic heterogeneity. *Math. Biosci.* **119**: 97–114.

[14] Dhirasakdanon, T., H. Thieme, and P. van den Driessche. 2007. A sharp threshold for disease persistence in host metapopulations. *J. Biol. Dynamics* **1**: 363–378.

[15] Diekmann, O. and J.A.P. Heesterbeek. 2000. Mathematical Epidemiology of Infectious Diseases. John Wiley and Sons Ltd, Chichester, New York.

[16] Diekmann, O., J.A.P. Heesterbeek, and J.A.J. Metz. 1990. On the definition and the computation of the basic reproduction ratio $\mathcal{R}_0$ in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**: 365–382.

[17] Doebeli, M. and G.D. Ruxton. 1998. Stabilization through pattern formation in metapopulations with long-range dispersal. *Proc. R. Soc. Lond. B.* **265**: 1325–1332.

[18] Emmert, K.E. and L.J.S. Allen. 2004. Population persistence and extinction in a discrete-time, stage-structured epidemic model. *J. Difference Equ. Appl.* **10**: 1177-1199.

[19] Gantmacher, F.R. 1960. The Theory of Matrices, Vol. II. Chelsea Pub. Co., New York, New York.

[20] Hsieh, Y.–H., P. van den Driessche, and L. Wang. 2007. Impact of travel between patches for spatial spread of disease. *Bull. Math. Biol.* **69**: 1355–1375.

[21] Lewis, M.A., J. Rencławowicz, and P. van den Driessche. 2006. Traveling waves and spread rates for a West Nile virus model. *Bull. Math. Biol.* **68**: 3–23.

[22] Lewis, M.A., J. Rencławowicz, P. van den Driessche, and M. Wonham. 2006. A comparison of continuous and discrete-time West Nile Virus models. *Bull. Math. Biol.* **68**: 491–509.

[23] Li, C.–K. and H. Schneider. 2002. Applications of Perron-Frobenius theory to population dynamics. *J. Math. Biol.* **44**: 450–462.

[24] Medlock, J. and M. Kot. 2003. Spreading disease: integro-differential equations old and new. *Math. Biosci.* **184**: 201–222.

[25] Ortega, J.M. 1987. Matrix Theory: A Second Course. Plenum Press, New York and London.

[26] Postnikov, E.B. and I.M. Sokolov. 2007. Continuum description of a contact infection spread in a SIR model. *Math. Biosci.* **208**: 205–215.

[27] Ruan, S. 2007. Spatial-temporal dynamics in nonlocal epidemiological models, in *Mathematics for Life Science and Medicine* Vol. 2. (eds. Y. Takeuchi, K. Sato and Y. Iwasa), Springer-Verlag, New York, 97–122.

[28] Smith, H. 1995. Monotone Dynamical Systems, An introduction to the theory of competitive and cooperative systems. American Mathematical Society, Mathematical Surveys and Monographs.

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

[30] Wang, M.–H., M. Kot, and M.G. Neubert. 2002. Integrodifference equations, Allee effects, and invasions. *J. Math. Biol.* **44**: 150–168.

[31] Yakubu, A.–A. and C. Castillo-Chavez. 2002. Interplay between local dynamics and dispersal in discrete-time metapopulation models. *J. Theor. Biol.* **218**: 273-288.

Appendix A

Proof of Lemma 2.5.

(a) Observe from (2.4) and (1.8) that

$$\begin{aligned}\lambda^*\varphi_j^* &= \big(1+\beta_j-\gamma_j\big)\sum_{k\in\Omega} q_{jk}\varphi_k^*, && j\in\Omega,\\ &= \big(1+\beta_j-\gamma_j\big)\left[d_I\sum_{k\in\Omega} r_{jk}\varphi_k^* + (1-d_I)\varphi_j^*\right], && j\in\Omega.\end{aligned}$$

Since R is stochastic, we obtain

$$\left[\frac{\lambda^*}{\nu_j}-1\right]\varphi_j^* = d_I\sum_{k\in\Omega} r_{jk}(\varphi_k^*-\varphi_j^*), \quad j\in\Omega, \tag{5.1}$$

where $\nu_j = 1+\beta_j-\gamma_j$. Observe that both φ^* and λ^* are functions of d_I. They are in fact differentiable functions of d_I by the implicit function theorem. We therefore differentiate both sides of (5.1) by d_I and suppress the $*$–notation to obtain

$$\frac{\lambda'}{\nu_j}\varphi_j + \left[\frac{\lambda}{\nu_j}-1\right]\varphi_j' = \sum_{k\in\Omega} r_{jk}(\varphi_k-\varphi_j) + d_I\sum_{k\in\Omega} r_{jk}(\varphi_k'-\varphi_j'), \quad j\in\Omega.$$

It suffices to show that $\lambda'<0$. We multiply both sides of the equation above by φ_j and sum over all $j\in\Omega$ to get

$$\lambda'\sum_{j\in\Omega}\frac{1}{\nu_j}\varphi_j^2 + \sum_{j\in\Omega}\left[\frac{\lambda}{\nu_j}-1\right]\varphi_j\varphi_j' = \sum_{j,k\in\Omega} r_{jk}(\varphi_k-\varphi_j)\varphi_j + d_I\sum_{j,k\in\Omega} r_{jk}(\varphi_k'-\varphi_j')\varphi_j.$$

Eq. (5.1) and the symmetry of R allow us to cancel the second sum on each side, and so obtain

$$\lambda'\sum_{j\in\Omega}\frac{1}{\nu_j}\varphi_j^2 = \sum_{j,k\in\Omega} r_{jk}(\varphi_k-\varphi_j)\varphi_j.$$

The symmetry of R again implies that

$$\lambda'\sum_{j\in\Omega}\frac{1}{\nu_j}\varphi_j^2 = -\frac{1}{2}\sum_{j,k\in\Omega} r_{jk}(\varphi_j-\varphi_k)^2. \tag{5.2}$$

Clearly, the right-hand side is nonpositive. We now show that it is in fact negative. We argue by contradiction. Suppose that

$$\sum_{j,k\in\Omega} r_{jk}(\varphi_j-\varphi_k)^2 = 0. \tag{5.3}$$

If $\varphi_j=\varphi_1$ for all $j\in\Omega$ then (5.1) and the positivity of φ imply that $\lambda=\nu_j$ for all $j\in\Omega$. But this is impossible because $\nu_j<1$ for $j\in H^-$, $\nu_j>1$ for $j\in H^+$, and H^- and H^+ are both nonempty. Therefore, it must be that $\varphi_m\neq\varphi_1$ for some $m\in\Omega$. The irreducibility of R implies that there exists a chain from 1 to m, i.e., a sequence $j_1, j_2, \ldots, j_s\in\Omega$ with $j_1=1$ and $j_s=m$ such that $r_{j_pj_{p+1}}>0$ for $1\le p\le s-1$. Eq. (5.3) implies that $\varphi_{j_p}=\varphi_{j_{p+1}}$ for $1\le p\le s-1$. Hence, $\varphi_1=\varphi_m$, another contradiction. We conclude that the right-hand side of (5.2) is negative, and thus that λ' is also negative.

(b) By (5.1) in part (a), we may write

$$\lambda^* \varphi_j^* = d_I \nu_j \sum_{k\in\Omega} r_{jk}(\varphi_k^* - \varphi_j^*) + \nu_j \varphi_j^*, \quad j \in \Omega. \tag{5.4}$$

According to Rayleigh's principle, λ^* can be characterized as

$$\lambda^* = \sup_{\sum_{j\in\Omega}\psi_j^2=1} \left\{ -\frac{d_I}{2} \sum_{j,k\in\Omega} \nu_j r_{jk}(\psi_j - \psi_k)^2 + \sum_{j\in\Omega} \nu_j \psi_j^2 \right\} \tag{5.5}$$

because R is symmetric. It follows that

$$\lim_{d_I\to 0} \lambda^* = \sup_{\sum_{j\in\Omega}\psi_j^2=1} \sum_{j\in\Omega} \nu_j \psi_j^2.$$

The supremum is achieved by setting $\psi_j = 1$ for a single $j \in \Omega$ with the property that $\nu_j = \max_{k\in\Omega}\{\nu_k\}$ and setting $\psi_j = 0$ otherwise.

(c) Parts (a) and (b) show that λ^* is a strictly monotone decreasing function of $d_I > 0$ that is bounded from above, and substituting $\psi_j = 1/\sqrt{n}$ for $j \in \Omega$ into (5.5) shows that λ^* is bounded from below by $\frac{1}{n}\sum_{j\in\Omega}\nu_j$. Therefore, λ^* has a limit $\lambda_\infty^* \in \left[\frac{1}{n}\sum_{j\in\Omega}\nu_j, \max_{j\in\Omega}\{\nu_j\}\right)$ as $d_I \to \infty$. We divide both sides of (5.4) by d_I to get

$$\frac{\lambda^* \varphi_j^*}{d_I} = \nu_j \sum_{k\in\Omega} r_{jk}(\varphi_k^* - \varphi_j^*) + \frac{\nu_j \varphi_j^*}{d_I}, \quad j \in \Omega. \tag{5.6}$$

Without loss of generality, we may assume that $\varphi^* = (\varphi_j^*)$ is a unit vector, i.e., $\sum_{j\in\Omega}(\varphi_j^*)^2 = 1$. It follows from the positivity of φ^* and compactness that $\varphi^* \to \overline{\varphi}$, where $\overline{\varphi}_j \geq 0$ for $j \in \Omega$ and $\sum_{j\in\Omega}\overline{\varphi}_j^2 = 1$, for some positive sequence of values $d_I \to \infty$. Let this sequence be denoted by $d_I^{(\ell)}$. Taking such a limit in (5.6) produces

$$\sum_{k\in\Omega} r_{jk}(\overline{\varphi}_k - \overline{\varphi}_j) = 0, \quad j \in \Omega.$$

Since R is stochastic, we can write these equations as $R\overline{\varphi} = \overline{\varphi}$. The nonnegativity and irreducibility of R imply that $\overline{\varphi}$ is proportional to 1_n, as both vectors belong to the principal eigenvalue $\mu = 1$. The fact that $\overline{\varphi}$ is a nonnegative unit vector implies that $\overline{\varphi}_j = 1/\sqrt{n}$ for $j \in \Omega$. Observe from the symmetry of R that $\sum_{j,k\in\Omega} r_{jk}(\varphi_k^* - \varphi_j^*) = 0$. Therefore, we can divide (5.4) with $d_I = d_I^{(\ell)}$ by ν_j and sum over all $j \in \Omega$ to get

$$\lambda^*(d_I^{(\ell)}) \sum_{j\in\Omega} \frac{\varphi_j^*}{\nu_j} = \sum_{j\in\Omega} \varphi_j^*.$$

We let $d_I^{(\ell)} \to \infty$ to get

$$\lambda_\infty^* \sum_{j\in\Omega} \frac{\overline{\varphi}_j}{\nu_j} = \sum_{j\in\Omega} \overline{\varphi}_j.$$

Since $\overline{\varphi}_j = 1/\sqrt{n}$ for $j \in \Omega$, we obtain

$$\lambda_\infty^* = \frac{n}{\sum_{j\in\Omega} 1/\nu_j} = \mathcal{B}.$$

Finally, parts (d) and (e) follow directly from parts (a), (b), and (c) together with the fact that H^+ is nonempty. □

Appendix B

Proof of Lemma 3.5.

(a) Observe from (3.3), (2.4), and the properties of Θ_j that

$$\begin{aligned}
H_j(\epsilon\varphi^*) &= \frac{d_I}{d_S}\Theta_j\left(\frac{\beta_j d_S(Q\epsilon\varphi^*)_j}{d_I+(d_S-d_I)\epsilon\varphi_j^*}\right)\big[1-(P\epsilon\varphi^*)_j\big]+(1-\gamma_j)(Q\epsilon\varphi^*)_j\\
&= \frac{d_I}{d_S}\Theta_j\left(\frac{\epsilon\beta_j d_S(Q\varphi^*)_j}{d_I}+\mathcal{O}(\epsilon^2)\right)\big[1+\mathcal{O}(\epsilon)\big]+(1-\gamma_j)\epsilon(Q\varphi^*)_j\\
&= \frac{d_I}{d_S}\left[\Theta_j(0)+\Theta_j'(0)\left(\frac{\epsilon\beta_j d_S(Q\varphi^*)_j}{d_I}\right)+\mathcal{O}(\epsilon^2)\right]\big[1+\mathcal{O}(\epsilon)\big]+(1-\gamma_j)\epsilon(Q\varphi^*)_j\\
&= \epsilon\Big[(1+\beta_j-\gamma_j)(Q\varphi^*)_j+\mathcal{O}(\epsilon)\Big]\\
&= \epsilon\varphi_j^*\Big[\lambda^*+\mathcal{O}(\epsilon)\Big].
\end{aligned}$$

Since $\lambda^*>1$ and $\varphi^*>0$, we conclude that $H_j(\epsilon\varphi^*)\geq\epsilon\varphi_j^*$ for $0<\epsilon\ll 1$.

(b) It is clear from (3.3) that $H_j(1_n)=1-\gamma_j\leq 1$. We conclude that $H(1_n)\leq 1_n$.

(c) We need only to show that if $I=(I_k)\in(0,1_n]$ then $H_j(I)$ is a monotone increasing function of I_k for $k\neq j$. It follows from (3.3) that

$$H_j(I)^* = \frac{d_I}{d_S}Z^*+(1-\gamma_j)q_{jk}, \tag{5.7}$$

where

$$Z=Z(X,I)=\Theta_j(X)\big[1-(PI)_j\big]\quad\text{and}\quad X=X(I)=\frac{\beta_j d_S(QI)_j}{d_I+(d_S-d_I)I_j}.$$

The $*$–notation here indicates that the total derivative with respect to I_k is to be taken. For $I\in(0,1_n]$, the numerator of X achieves its maximum value of $\beta_j d_S$ when $I=1_n$ and the denominator achieves its minimum value of d_S whenever $I_j=1$ because $0<d_S\leq d_I$. Therefore $0\leq X\leq\beta_j$. Observe also that

$$Z^*=\Theta_j'(X)X^*\big[1-(PI)_j\big]-\Theta_j(X)p_{jk}.$$

Recalling that $k\neq j$, we make several remarks: (i) $X^*=q_{jk}X/(QI)_j$, (ii) $p_{jk}=(d_S/d_I)q_{jk}$, and (iii) $\Theta_j(X)\leq X$. Consequently,

$$\begin{aligned}
Z^* &\geq q_{jk}X\left[\Theta_j'(X)\frac{1-(PI)_j}{(QI)_j}-\frac{d_S}{d_I}\right]\\
&= q_{jk}X\left[\Theta_j'(X)\frac{d_I-d_I(PI)_j}{d_S(QI)_j}-1\right]\frac{d_S}{d_I}.
\end{aligned}$$

We apply Lemma 3.4 to get

$$\begin{aligned}
Z^* &\geq q_{jk}X\left[\Theta_j'(X)\left(\frac{d_I+(d_S-d_I)I_j}{d_S(QI)_j}-1\right)-1\right]\frac{d_S}{d_I}\\
&= q_{jk}\Big[\Theta_j'(X)(\beta_j-X)-X\Big]\frac{d_S}{d_I}.
\end{aligned}$$

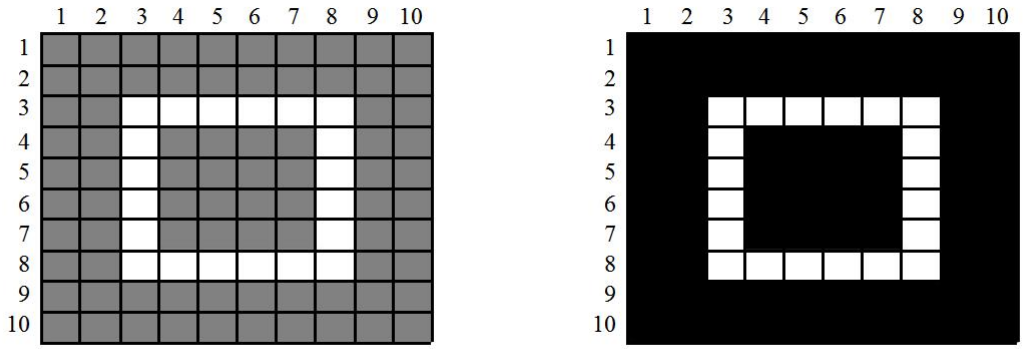
It follows from the nonnegativity of $\Theta_j'(X)$ and the fact that $0 \le X \le \beta_j$ that

$$Z^* \ge q_{jk}(-\beta_j)\frac{d_S}{d_I}.$$

Eq. (5.7) and the assumption (B1) imply that

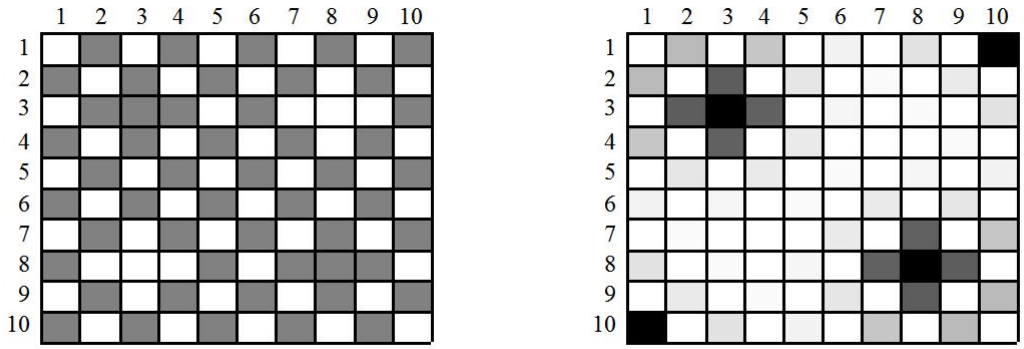
$$H_j(I)^* \ge q_{jk}(1-\beta_j-\gamma_j) \ge 0.$$

We conclude that $H = (H_j)$ is a monotone increasing function of I_k for every $k \ne j$ when $I = (I_k) \in (0, 1_n]$. □



(a) (b)

Figure 1: (a) A habitat divided into three regions: low-risk patches (unshaded) have $\beta_j = 0.2$ and $\gamma_j = 0.7$ and high-risk patches (shaded) have $\beta_j = 5.0$ and $\gamma_j = 0.2$. (b) The corresponding EE. After 8,000 time steps, the low-risk patches (unshaded) each contain more than 460 individuals (of which fewer than 0.0003 are infected) and the remaining high-risk patches (darkly-shaded) each contain fewer than 0.0008 individuals. This EE (with $d_S = 10^{-6}$) closely approximates the limiting DFE (with $d_S = 0$). In this example, the baseline dispersal matrix R has entries $r_{ij} = 0.125$ whenever distinct patches i and j share an edge and $r_{ii} \in \{0.5, 0.625, 0.75\}$. The infected diffusion coefficient $d_I = 1.0$; the infection function is $\Theta_j(x) = 1 - e^{-x}$ for $j \in \Omega$; the total population size is $N = 10,000$; and the initial data satisfy $S_j(0) + I_j(0) = 100$ for $j \in \Omega$, with $I_j(0) = 10$ in each of the four corner patches and $I_j(0) = 0$ otherwise. Thus, the habitat is high-risk ($\mathcal{B} = 1.86$) and the basic reproduction number $\mathcal{R}_0 = 21.71$.



(a) (b)

Figure 2: (a) A nearly checkerboard habitat: low-risk patches (unshaded) have $\beta_j = 0.2$ and $\gamma_j = 0.9$ and high-risk patches (shaded) have $\beta_j = 1.0$ and $\gamma_j = 0.6$. (b) The corresponding EE. After 8,000 time steps, the four darkly-shaded patches each contain fewer than 0.002 individuals, the eight medium-shaded patches (which lie adjacent to the two interior darkly-shaded patches) each contain between 16 and 19 individuals, and all other patches (unshaded and lightly-shaded) contain at least 72 individuals. Also, fewer than 0.0002 infected individuals remain in every patch. This EE (with $d_S = 10^{-6}$) closely approximates the limiting DFE (with $d_S = 0$). The baseline dispersal matrix R, the infected diffusion coefficient d_I, the infection function $\Theta_j(x)$, the total population size N, and the initial data $S_j(0)$ and $I_j(0)$ are the same as in Figure 1. Here, the habitat is low-risk ($\mathcal{B} = 0.51$) and the basic reproduction number $\mathcal{R}_0 = 1.10$.