Catching a wave: on the suitability of traveling-wave solutions in epidemiological modeling

Anna M. Langmüller[a,b], Joachim Hermisson[b,d], Courtney C. Murdock[c], Philipp W. Messer[a]

[a] *Cornell University, Department of Computational Biology, 102 Tower Rd, Ithaca, 14850, NY, USA*
[b] *University of Vienna, Department of Mathematics, Oskar-Morgenstern-Platz 1, Vienna, 1090, Vienna, Austria*
[c] *Cornell University, Department of Entomology, 129 Garden Ave, Ithaca, 14850, NY, USA*
[d] *Max Perutz Labs, Vienna Biocenter Campus, Dr.-Bohr-Gasse 9 Vienna, 1030, Vienna, Austria*

Abstract

Ordinary differential equation models such as the classical SIR model are widely used in epidemiology to study and predict infectious disease dynamics. However, these models typically assume that populations are homogeneously mixed, disregarding possible variations in disease prevalence due to spatial heterogeneity. To address this issue, reaction-diffusion models have been proposed as an alternative approach for modeling spatially continuous populations in which individuals move in a diffusive manner. In this study, we explore the conditions under which such spatial structure must be explicitly considered to accurately predict disease spread, and when the assumption of homogeneous mixing remains adequate. In particular, we derive a critical threshold for the diffusion coefficient below which disease transmission dynamics display spatial heterogeneity. We validate our analytical results with individual-based simulations of disease transmission across a two-dimensional continuous landscape. Using this framework, we further explore how key epidemiological parameters such as the establishment probability of a disease, its maximal incidence, and its final epidemic size are affected by incorporating spatial structure in SI, SIS, and SIR models. We discuss the implications of our findings for epidemiological modeling and identify design considerations and limitations for spatial simulation models of disease dynamics.

Keywords: ordinary differential equations, diffusion theory, spatial disease

1. Introduction

A fundamental understanding of how infectious diseases spread across time and space is crucial for predicting the course of epidemics and guiding potential control measures. Compartmental models have provided the workhorse for infectious disease modeling for almost a century [1, 2, 3]. In these models, the population is divided into compartments based on their infection status, such as susceptible (S), infectious (I), and recovered (R) individuals. One of the simplest compartmental models is the so-called SI model, where individuals can only be susceptible or infected [1]. Once infected, an individual perpetually remains in that state. Two classic extensions of this model are the SIS and SIR models [1]: In the former, infected individuals can return to the susceptible class, whereas in the latter they can recover, often implying that they are immune to future infections.

Compartmental models are typically studied with ordinary differential equations (ODEs) describing the expected changes in the proportions of the different population compartments over time [1, 2, 4, 5, 3]. For example, consider a simple SI model where individuals come into contact with each other at a constant rate r. Let $i(t)$ specify the fraction of infectious individuals in the population. Whenever a susceptible individual meets an infected one, it contracts the pathogen with probability α (for simplicity, we will use the terms 'pathogen' and 'disease' interchangeably throughout this paper, acknowledging their conceptual distinction). If we assume homogeneous mixing among the individuals, the change in the proportion of infectious individuals over time will be given by:

$$\frac{\partial i}{\partial t} = r\alpha i(1-i) \tag{1}$$

Epidemiological models of this type can be extended to include additional compartments, such as a recovered class, in which case their dynamics may need to be described by a set of ODEs. Over the past century, such models have provided invaluable insights into the factors that shape disease transmission by providing conceptual results and threshold quantities (e.g., herd immunity, basic reproduction number, and replacement number) to anticipate

disease spread and the impact of countermeasures, thereby guiding decision-making and public health policies during disease outbreaks [4, 2, 3, 5, 6].

The limitations of ODE models often lie in the underlying assumptions they make for the transmission process. One common assumption of ODE models is homogeneous mixing among individuals [7]. Yet, this is rarely the case for real-world populations, which can be structured in multiple ways [8, 9, 10, 11, 12]. In particular, most populations are geographically structured, such that individuals that are close to each other are more likely to come into contact than individuals that are far apart. Besides spatial structure, kinship and social structure can also have a substantial influence on contact patterns [13, 14]. As a result, local disease prevalence can vary substantially between different parts of the population, which can have a considerable impact on the expected epidemiological dynamics [15, 11, 16]. While classic ODE models have clear advantages for modeling disease transmission, such as their well-understood mathematical properties, computational efficiency, and ability to provide analytical solutions [11], omitting the influence of population structure in these models can lead to inaccurate conclusions about the anticipated disease dynamics.

In this study, we will focus specifically on the potential effects of spatial population structure, which can be incorporated into disease transmission models in several ways. One approach is the use of a meta-population model that divides the population into subpopulations or "patches", each representing a homogeneous host population such as a city. The dynamics within the patches, and the coupling between them, can then be described using a system of ODEs [17, 18, 19]. Such coupling could be based on geographic distance, empirical observations such as mobile phone data [20], or geolocated social media posts, for instance [21]. Alternatively, individual-based or agent-based approaches can be used to model all individuals in the population and their spatial position explicitly [22, 16, 23]. The interactions between individuals (e.g., the rate and strength of contacts) can be regulated by distance-based kernels [5, 24], which can result in a highly variable infection probability across space [11]. Network models represent another class of approaches that allow for very fine-scaled modeling of population structure [25]. In these models, each node usually represents an individual, with edges between nodes representing contacts between individuals. A common assumption of this approach is that a susceptible individual can

only contract the disease if an edge connects it to an infectious individual [16, 26]. However, network models are complex, and their analysis can require an extensive amount of empirical data to model larger populations [16].

For populations inhabiting a continuous geographic landscape, where individuals primarily move within short distances compared to the overall range of the landscape, reaction-diffusion models offer an alternative approach for describing spatial disease dynamics [27, 28]. These models gained popularity in the 1930s, when both Fisher and Kolmogorov used them to explain the spread of a beneficial allele across a one-dimensional habitat [29, 30]. In such models, the beneficial allele can propagate through the population in the form of a "traveling wave", with a constant and well-defined velocity. Since then, diffusion models have been widely used in mathematical biology to study not only the spread of adaptive alleles but also invasive organisms and infectious diseases [28, 31, 32, 27, 33, 34, 35, 36, 37, 3, 38]. In the context of epidemiology, reaction-diffusion models enable us to describe the proportion of infected individuals as a function of space and time. The diffusion term models the random dispersal of infectious individuals or contacts [39], while the reaction term describes the local rates of disease transmission, similar to a compartmental ODE model [27].

In this paper, we use diffusion theory to study under which scenarios a simple compartmental ODE model is sufficient to capture disease dynamics and when spatial structure has to be taken into account. Within the framework of an individual-based simulation model, we derived a simple critical threshold D_c for the diffusion coefficient. This threshold marks the point below which the limited dispersal of individuals should significantly affect simulated disease dynamics. To validate our findings, we conduct individual-based simulations using SLiM [40], which allow us to smoothly transition from spatially unstructured to structured populations. Using this framework, we investigate the effect of spatial structure on disease dynamics in three classic compartmental disease transmission models: the SI, SIS, and SIR model [1, 4]. Furthermore, we highlight potential caveats of simulating continuous disease transmission models in a discrete, individual-based, and biologically realistic manner.

2. Results

2.1. ODE model

Consider a simple Susceptible-Infected (SI) model in a closed population of constant size N without birth or death events [1, 5]. There are only susceptible and infectious individuals in the population, and infectious individuals do not recover (i.e., they never leave the infectious compartment). Let $I(t)$ and $S(t) = N - I(t)$ denote the overall counts of infectious and susceptible individuals in the population at time t, and $s(t) = S(t)/N$ and $i(t) = I(t)/N$ their respective population frequencies with $i(t) + s(t) = 1$. If we assume a homogeneously mixing population in which individuals come into contact with each other at a constant contact rate r per time unit, and a disease that establishes in a susceptible individual after contact with an infectious individual with probability α, then the change in the proportion of infectious individuals is given by Equation (1). The solution to this ODE model is a logistic growth function [41]:

$$i(t) = \frac{1}{1 + [N/I(0) - 1]e^{-r\alpha t}} \tag{2}$$

Whenever $I(0) > 0$, we have $\lim_{t\to\infty} i(t) = 1$, meaning the disease will ultimately spread throughout the entire population. Figure 1A provides an illustration of this logistic spread in the ODE model. Note that the parameters r and α always appear in tandem, with the compound parameter $(r\alpha)^{-1}$ setting the timescale of the dynamics. However, since r and α regulate distinct processes in the context of our individual-based simulation framework described in section 2.4, they are not summarized into one compound variable in this manuscript.

Assuming one infected individual in the beginning ($I(0) = 1$) and a sufficiently large population ($N \gg 1$), the expected time for the disease to reach 50% frequency in the population is:

$$t_{1/2} \approx \frac{\ln\left(N/I(0)\right)}{r\alpha} \approx \frac{\ln N}{r\alpha} \tag{3}$$

Since $i(t)$ is symmetric around the inflection point at $i(t_{1/2}) = 1/2$, the expected time for the disease to spread through the whole population under the assumptions $N \gg 1$ and $I(0) = 1$, is therefore:

$$t_{\text{ODE}} \approx 2t_{1/2} \approx \frac{2\ln N}{r\alpha} \tag{4}$$

We call t_{ODE} the "fixation time" of the ODE model.

2.2. Reaction-diffusion model

ODE models such as the simple compartmental SI model outlined above are widely used in epidemiological modeling [42], yet their underlying assumption of homogeneous mixing makes it difficult to incorporate spatial population structure. For populations that inhabit a continuous landscape, diffusion theory provides an alternative modeling framework that explicitly includes spatial heterogeneity in disease prevalence [3, 43, 44, 29, 30, 28]. In a diffusion model, the frequency of infected individuals becomes a function not only of time t but also of spatial position x.

Let us consider an SI model in which individuals move randomly according to some dispersal kernel that decays with distance at least as quickly as an exponential function [38]. In such a case, we can model dispersal by a diffusion term [38, 29, 45, 28], while disease transmission can be modeled by a reaction term based on the local compartment frequencies. Together, this yields a second-order partial differential equation for the SI diffusion model of the following form:

$$\frac{\partial}{\partial t} i(x,t) = D \frac{\partial^2}{\partial x^2} i(x,t) + r\alpha i(x,t)[1 - i(x,t)] \tag{5}$$

The diffusion coefficient D accounts for the random dispersal of individuals. In a one-dimensional habitat, $2D = \sigma^2$, where σ^2 is the variance in the expected spatial displacement of individuals per time unit due to their random movement [29, 46]. The reaction term is analogous to the ODE model from Equation (1), except that the frequency of infected individuals now depends on the spatial position x as well. Equation 5 is also known as the Fisher-Kolmogorov or Fisher-KPP equation [29, 30, 3].

We can express Equation 5 in dimensionless form by re-scaling time, $\tilde{t} = (r\alpha)t$, and space, $\tilde{x} = x/(\sqrt{D/(r\alpha)})$ where $(r\alpha)^{-1}$ is the characteristic timescale of the system:

$$\frac{\partial}{\partial \tilde{t}} i(\tilde{x},\tilde{t}) = \frac{\partial^2}{\partial \tilde{x}^2} i(\tilde{x},\tilde{t}) + i(\tilde{x},\tilde{t})[1 - i(\tilde{x},\tilde{t})] \tag{6}$$

However, we maintain the habitat size L and the simulation step-size (i.e., "tick") Δt as the units of space and time. Thus, we retain the parameterization of the Fisher-KPP equation (Equation 5), facilitating mapping between

the continuous reaction-diffusion model and the finite individual-based simulation framework (section 2.4) while preserving parameter interpretability in both contexts.

Under the assumption that the sum of susceptible and infectious individuals are conserved at a constant carrying capacity, and that there is no change in dispersal due to infection, one solution to Equation 5 is a so-called "traveling wave" [29, 47, 48, 27, 3, 38, 33] in which the disease spreads outward from an initial introduction point. The minimum velocity c_0 of such a wave is defined as:

$$c_0 = 2\sqrt{Dr\alpha} \tag{7}$$

Figure 1B provides an illustration of these dynamics in a one-dimensional habitat, where the disease is introduced at the center and then spreads outward in the form of two symmetric traveling waves to the left and right. The "width" of each wave, i.e., the length of the region between its front (where the frequency of infected individuals is still close to zero) and top (where almost everyone is already infected), is approximately $w \approx 2\sqrt{D/(r\alpha)}$ [29].

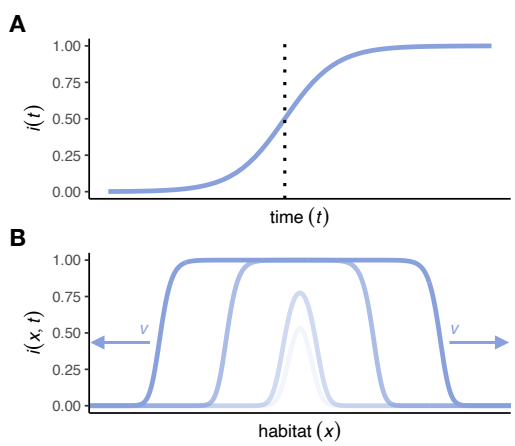

Figure 1: Schematic plots of expected disease spread under the ODE model and a traveling wave solution of the reaction-diffusion model. (A) In the ODE model, the frequency of infected individuals, $i(t)$, increases logistically according to Equation (2). The vertical dashed lines represent $t_{1/2}$ as defined in Equation (3). (B) In the diffusion model for a one-dimensional habitat, the disease spreads from an initial release at the center of the habitat via two traveling waves with velocity v (minimum velocity $c_0 = 2\sqrt{Dr\alpha}$) in either direction.

It is straightforward to extend this diffusion model to two dimensions. Specifically, if dispersal is isotropic, displacement along the x and y axes can be treated independently, with each following the one-dimensional model. The diffusion coefficient then equals half the variance in the displacement in each of the x and y coordinates: $2D = \sigma_x^2 = \sigma_y^2$. Note that the overall displacement in a two-dimensional space, $\Delta = \sqrt{\Delta_x^2 + \Delta_y^2}$, will typically be larger than the displacement in each individual dimension.

In this two-dimensional diffusion model with isotropic dispersal, the disease spreads from an initial introduction point in the form of a circle whose radius grows with velocity v. With decreasing curvature at the wavefront, v asymptotically approaches the same minimal wave speed c_0 as the traveling wave in the one-dimensional model [8, 49]. If the disease is introduced in a small number of individuals in the center of a square habitat patch with a side length L that is much larger than the wave width (which requires $L \gg 2\sqrt{D/r\alpha}$ [50]), the wave reaches the corners of the square patch when the circle grows to a radius of $L/\sqrt{2}$.

Thus, the expected fixation time of our SI diffusion model in the two-dimensional habitat patch is approximately:

$$t_{\text{DIFF}} \approx \frac{L}{\sqrt{8Dr\alpha}} \tag{8}$$

2.3. Critical threshold

The ODE model assumes a homogeneously mixing population, while the diffusion model can explicitly account for spatial heterogeneity in disease transmission due to limited dispersal. Naturally, the question arises of how the epidemiological dynamics predicted by these models differ, and under which parameters one is more accurate than the other. To answer this question, we can contrast their expected disease fixation times, t_{ODE} and t_{DIFF}. Setting both equal, we obtain a critical value for the diffusion coefficient under which both models predict the same fixation time:

$$t_{\text{ODE}} = t_{\text{DIFF}} \implies D_c = \frac{r\alpha L^2}{32(\ln(N/I(0))^2} \tag{9}$$

When $D < D_c$ the ODE model predicts a shorter fixation time than the diffusion model because the slow dispersal of individuals becomes a limiting factor for how fast a disease can spread throughout the whole habitat. In

this low-dispersal regime, we expect the diffusion model to more accurately describe epidemiological dynamics than the ODE model, which should overestimate the rate of disease spread. However, it is crucial to note that our estimation of t_{DIFF} in Equation (8) solely accounts for spatial progression, neglecting the time required for the build-up of the wavefront and the time until disease fixation after the wavefront has reached the habitat boundaries. These processes are governed by the logistic growth function of the ODE model. Thus, we use a heuristic approach that approximates $i(t)$ by summing up the expected times until a specific frequency is attained under a traveling wave model with constant speed (Equation 7), and a logistic growth model (Equation 2).

As D becomes on the order of D_c, the width of the traveling wave in the diffusion model approaches the habitat size, the disease no longer spreads as a circle with a "sharp" edge [50], and spatial progression ceases to limit disease spread. In this case, our estimate of t_{DIFF} in Equation (8) is no longer accurate and the fixation time should gradually approach the prediction of the ODE model. In the high-dispersal regime where $D \gg D_c$, diffusion is strong enough that the population is effectively well mixed over the timescale relevant for the spread of the disease and the ODE model should describe the epidemiological dynamics most accurately.

2.4. Individual-based simulation framework

To test the theoretical predictions derived above, we implemented an individual-based distributed-infectives simulation model for disease transmission in SLiM (version 4.0) [40]. By varying the dispersal rates of individuals, our simulation model can smoothly transition from spatially unstructured to highly structured populations, allowing us to investigate disease dynamics across the high and low-dispersal regimes, and to study the predicted transition around the critical diffusion coefficient D_c. We focused on an abstract, directly transmitted disease in a closed population (i.e., no birth or death processes), comprised of $N = 100,000$ individuals. These individuals move around in a continuous two-dimensional spatial domain, modeled as a square habitat patch with a side length of 1 (i.e., $L = 1$) and periodic boundaries to avoid edge effects [51, 15].

The ODE and diffusion models are both continuous-time models. Our individual-based simulation model in SLiM, on the other hand, measures time in discrete units of "ticks" (i.e., $\Delta t = 1$). Each such "tick" is associated with the opportunity for new infections and recovery of individuals. To minimize differences between the discrete-time simulations and the continuous-time mathematical models, we must ensure that Δt is sufficiently small compared to the characteristic timescale $(r\alpha)^{-1}$. However, this requirement does not limit the generality of our models. We can freely choose the time interval in the real world that a tick in our simulations corresponds to, be it a year or a millisecond. By choosing a sufficiently small period, we can always ensure that $\Delta t \ll (r\alpha)^{-1}$.

The disease is transmitted via local contacts between infectious and susceptible individuals. In each tick, we assume that an individual has contact with all individuals that are currently within its interaction distance δ, which we chose such that each individual on average has 15 contacts per tick (i.e., r=15). Note that the average number of contacts per tick is independent of the infection status (i.e., infectious individuals, on average, have as many contacts as susceptible or recovered individuals). For a square habitat with edge length normalized to $L = 1$ and toroidal boundary conditions, and with 100,000 individuals, this corresponds to an interaction radius of $\delta \approx 0.00691$. Each contact between a susceptible individual and an infectious individual results in disease transmission with a probability of α=0.001, yielding $(r\alpha)^{-1} \approx 66.67$ ticks $\gg 1$. When a susceptible individual has contracted the disease, it becomes infectious in the next tick. At the end of each tick, we simulate isotropic dispersal by sampling the x and y displacements for each individual independently from a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 2D$. The infection status of an individual has no effect on D.

We implemented three compartmental disease transmission models in this simulation framework: (i) the SI model, (ii) the SIS model, and (iii) the SIR model [5]. In the SI model, once an individual is infected, it remains infectious until the end of the simulation. In the SIS model, infected individuals can recover and return to the susceptible class in each tick with probability γ. In the SIR model, they recover and gain lasting immunity with probability γ in each tick.

Each simulation run is initialized by uniformly distributing 100,000 sus-

ceptible individuals across the habitat. At tick 25, the disease is introduced into the population by infecting the individual closest to the center of the square habitat (i.e., $I(0) = 1$). For each simulation, we recorded the number of individuals in each compartment at the beginning of each tick until either the disease spread through the whole population, no infectious individuals were left, or the simulation ran for 50,000 ticks. This upper threshold of 50,000 ticks results in a very lenient timeline that is not expected to interfere with the simulated SI disease transmission dynamics. For our simulation parameters, the disease is expected to fix in the SI model between approximately 1,500–14,000 ticks (Equations 4 and 10).

2.4.1. Low-diffusion limit

A key abstraction of the diffusion approach is that it models a continuous density of individuals across the habitat. The relative frequencies of infectious and susceptible individuals are specified at any given position in the habitat, and disease transmission can occur exactly at that position. In our individual-based simulation model, however, a finite number of individuals are dispersed across a continuous habitat. It is unlikely that two of them will ever be at exactly the same position. Thus, one has to decide how close two individuals need to be for one to be able to infect the other, which we implemented by defining an interaction radius δ.

Ideally, δ should be very small (i.e., much smaller than the average dispersal distance of individuals) so that the movement of individuals remains the primary driver of disease spread. Yet, there is a lower limit on δ for any given contact rate in our individual-based simulation model. For example, if we want to ensure that each individual on average has, say, two contacts, this requires a certain minimum interaction radius so that each individual encounters enough individuals within its interaction radius. In our simulated continuous two-dimensional habitat, which is represented by a square with a side length of L and includes N uniformly distributed individuals, the interaction radius δ is defined as $\delta = L\sqrt{r/(N\pi)}$.

One important consequence is that, when D becomes sufficiently small, the spread of the disease is no longer driven primarily by the dispersal of individuals, but instead by "hopping" between neighboring individuals with overlapping interaction circles. These "hopping" dynamics can lead to the

disease fixing even when individuals do not move at all, thereby setting an upper bound for the fixation time in the limit $D \to 0$.

We can still describe disease spread under these hopping dynamics by the diffusion model, but we need to reinterpret the diffusion term. In particular, we interpret each transmission event as a "dispersal" step of the disease from the location of the infecting individual to the location of the newly infected individual. To derive a rough estimate of the diffusion coefficient (D_0) of the transmission, we consider the limit where individuals no longer move at all. Starting from the first infected individual at the center of the habitat, the disease spreads outward in a circular pattern, with the traveling wave now driven entirely by transmission events rather than by individual dispersal. The minimum speed of this traveling wave for the SI model is given by $c_0 = 2\sqrt{D_0 r\alpha}$. One complication is that, in the hopping model, D_0 varies across space because it depends on the occurrence of new transmission events. Thus, D_0 approaches zero in areas where either everyone is already infected or everyone is still susceptible. It is maximal right at the wavefront, where it determines the speed of the wave.

Let us, therefore, consider an infected individual located exactly at the wavefront. If its neighbors are uniformly distributed across its interaction circle, around half of these individuals, on average, are already infected, while the others are still susceptible. The overall rate at which transmission events from the focal infected individual occur then is $\Pr(\text{transmission}) = r\alpha/2$. If a transmission event occurs, the variance of the displacement along the x-axis is $\sigma_x^2 = \delta^2/4$. The effective diffusion coefficient of the hopping process thus results as $2D_0 \approx \Pr(\text{transmission})\sigma_x^2 = r\alpha\delta^2/8$. According to Equation (8), this yields a fixation time of:

$$t_0 \approx \frac{\sqrt{2}L}{r\alpha\delta} \tag{10}$$

This constitutes an approximate upper bound of the fixation time in the limit $D \to 0$, where the fixation time from dispersal alone, given by Equation (8), diverges to infinity. By comparing the fixation time estimates t_0 and t_{DIFF}, we determine D_l, a threshold for individual dispersal below which disease spread in our individual-based simulation model is primarily driven by transmission events:

$$t_0 = t_{\text{DIFF}} \implies D_l = \frac{r\alpha\delta^2}{16} = \frac{(rL)^2\alpha}{16N\pi}. \tag{11}$$

2.5. SI Model

To test how well our mathematical predictions describe disease spread in the individual-based simulations, we start with the SI model. Qualitatively, we expect that in the low-dispersal regime ($D < D_c = 3.536 \times 10^{-6}$), the disease spreads from its introduction at the center of the habitat as a circle with a steadily growing radius. By contrast, in the high-dispersal regime ($D \gg D_c$), the population should be sufficiently mixed so that the frequency of infected individuals increases in all areas of the habitat at a similar rate. Figure 2 confirms these qualitative predictions when comparing two simulation runs with $D = 10^{-6}$ (left column) and $D = 10^{-2}$ (right column).

We next tested our predictions for the disease fixation time under varying dispersal rates (Figure 3A). In the high-dispersal regime, the observed fixation time t_{sim} is independent of D and well approximated by the predictions from the ODE model given in Equation (4). The time-resolved proportion of infectious individuals follows the logistic growth curve (Figure 3B, bottom panel). When D becomes smaller than D_c, the fixation time starts to increase and becomes inversely proportional to D, as predicted by Equation (8). Once D becomes smaller than D_l ($D_l = 4.476 \times 10^{-8}$), the fixation time approaches the prediction under the hopping dynamics derived in Equation (10), t_{sim} levels off and again becomes independent of D.

Overall, we observe that fixation times can differ up to one order of magnitude in our simulation model depending on the dispersal rate. This highlights the potential risk of underestimating the expected epidemic duration if predictions are based solely on the ODE model without accounting for population structure.

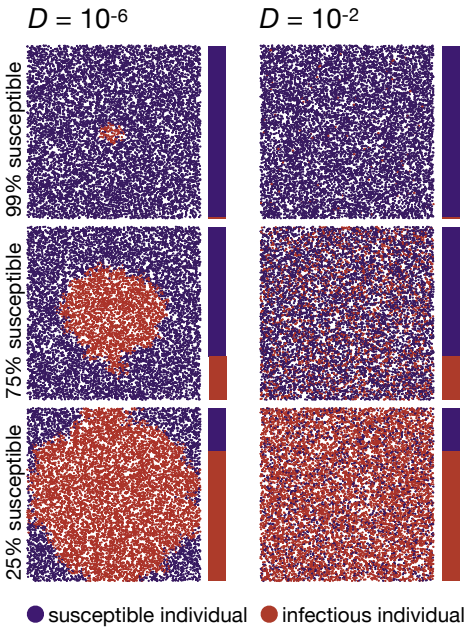

Figure 2: Two exemplary simulation runs in the SI model. The left column shows a run in the low-dispersal regime ($D = 10^{-6}$), while the right column shows a run in the high-dispersal regime ($D = 10^{-2}$). The top, middle, and bottom plots show population snapshots taken when 99%, 75%, and 25% of the population were still susceptible. As predicted, the disease progresses in a circular pattern in the low-dispersal regime. By contrast, with large dispersal, infectious individuals are homogeneously distributed across space (right column). Data were simulated using the following parameters: $N = 10,000$ (ten-times smaller than our standard model for better visualization), $\alpha = 0.01$, $r = 15$, $L = 1$. Videos of simulation runs in the SI model can be found under `https://tinyurl.com/bdddam58`.

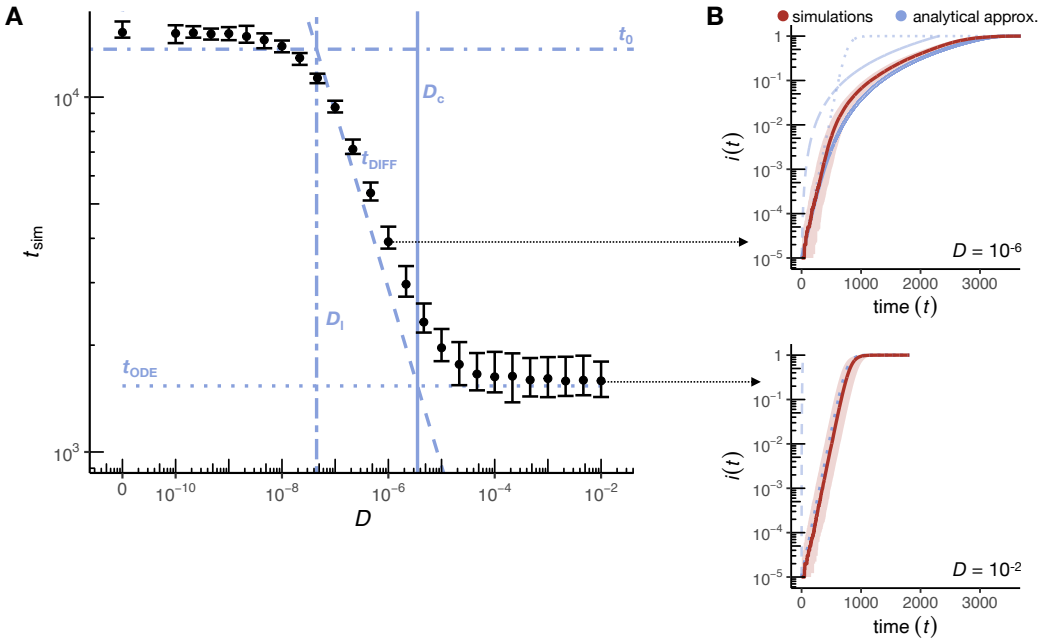

Figure 3: Disease dynamics in the SI model. (A) Time until disease fixation in the simulations (t_{sim}) for varying diffusion coefficients (D). Black dots represent median values and error bars the 2.5 and 97.5 percentiles estimated over 50 simulation runs for each D (for $D = 0$, the spread stagnated in 3 of the 50 runs due to small clusters of susceptible individuals with interaction areas devoid of infectious individuals, which we excluded from the analysis). The vertical solid line indicates the critical threshold D_c according to Equation (9). In the high-dispersal regime ($D \gg D_c$), fixation times converge to the prediction of the ODE model (t_{ODE}, dotted line), whereas in the limit of no dispersal ($D \to 0$), they are bounded by the prediction of the hopping model (t_0, dot-dashed line; D_l is indicated as vertical two-dashed line). In the low-dispersal regime where $D < D_c$, the disease advances as a traveling wave, and t_{sim} is well approximated by t_{DIFF} (dashed line). (B) The top panel shows the time-resolved disease dynamics for the low-dispersal regime ($D = 10^{-6}$). Here, the proportion of infectious individuals, $i(t)$, increases after an initial lag phase approximately as the area of a circle whose radius expands at the predicted minimum wave speed of the diffusion model (blue dashed line). The initial lag between the expected and the observed proportion of infectious individuals is likely caused by the local establishment of the disease and an initial reduction of wave speed due to the curvature of the wavefront [8, 45]. $i(t)$ of the simulations is well approximated by a heuristic approach (solid blue line) that accounts for both disease propagation via a constant traveling wave, and the local rise in $i(t)$ described by the ODE model (dotted line). The bottom panel shows an example from the high-dispersal regime ($D = 10^{-2}$). Here, the proportion of infectious individuals is well approximated by the expected logistic growth function of the ODE model (dotted blue line). Solid red lines represent the median proportion of infectious individuals across 50 simulation runs, and the shaded areas represent the range between the 2.5 and 97.5 percentiles. All data were simulated using the following parameters: $N = 100,000$, $\alpha = 0.001$, $r = 15$, $L = 1$.

2.6. SIS Model

In the SI model, infected individuals do not recover from the infection, leading to an inevitable spread and ultimate fixation of the disease as long as $r\alpha > 0$. The SIS model allows infected individuals to transition from infection back to the susceptible state at rate γ. In a deterministic ODE model of an unstructured population, such a disease is expected to reach an equilibrium frequency of $1 - 1/R_0$ as long as $R_0 > 1$, where $R_0 = r\alpha/\gamma$ represents the basic reproduction number of the disease (i.e., the average number of secondary infections caused by a single infectious individual when introduced into a completely susceptible population) [52, 5]. If $R_0 < 1$, the disease is typically incapable of establishing in the population and will quickly die out.

In addition to the critical threshold D_c established for the SI model, one can analogously determine a critical threshold D_c in the SIS model, below which we expect individual dispersal to affect the disease dynamics. For the SIS model, the frequency of infected individuals $i(t)$ follows a logistic growth function with an equilibrium frequency of $1 - 1/R_0 = 1 - \gamma/(r\alpha)$ in a homogeneously mixing population:

$$i(t) = \frac{1 - \gamma/(r\alpha)}{1 + \frac{1-\gamma/(r\alpha)-I(0)/N}{I(0)/N} e^{-(r\alpha-\gamma)t}} \tag{12}$$

Equation 12 allows the derivation of an expected "time until equilibrium" estimate for the ODE model, similar to the derivation of the "fixation time" estimate for the SI model (Equation 4). Under the assumptions $N \gg 1$ and $I(0) = 1$, this "time until equilibrium" can be estimated for the SIS model in a homogeneously mixing population as:

$$t_{\text{ODE}} \approx 2t_{1/2} \approx \frac{2\ln\left(\frac{1-\gamma/(r\alpha)-I(0)/N}{I(0)/N}\right)}{r\alpha - \gamma} \approx \frac{2\ln\left(N(1-\gamma/(r\alpha)\right)}{r\alpha - \gamma} \tag{13}$$

For a spatial SIS model, the minimal wave speed c_0 accounts for the recovery rate γ of infectious individuals back to the susceptible state, yielding $c_0 = 2\sqrt{D(r\alpha - \gamma)}$ [29]. This approximation for c_0 can then be employed to calculate the "time until equilibrium" estimate for the spatial SIS model:

$$t_{\text{DIFF}} \approx \frac{L}{\sqrt{8D(r\alpha - \gamma)}} \tag{14}$$

Following the lines discussed for the SI model above, we can contrast the expected times until equilibrium for the ODE (t_{ODE}) model and our estimation of t_{DIFF} to obtain a critical value for the diffusion coefficient above which we expect a negligible effect of spatial progression on disease dynamics.

$$t_{\mathrm{ODE}} = t_{\mathrm{DIFF}} \implies D_c = \frac{(r\alpha - \gamma)L^2}{32(\ln\left((1-\gamma/(r\alpha))N/I(0)\right))^2} \tag{15}$$

To investigate the effects of spatial structure and stochasticity on an SIS model, we modified our individual-based simulations such that infected individuals reenter the susceptible class with a probability γ per tick, yielding $R_0 = r\alpha/\gamma$ in our simulation framework. Qualitatively, we find that the disease initially still spreads in a circular fashion in the low-dispersal regime and more homogeneously in the high-dispersal regime, as in the SI model (Figure 4). However, the perpetual reentering of infectious individuals into the susceptible class breaks up this strict circular clustering more quickly in the SIS model as compared to the SI model. Over the course of the simulated time period, the frequency of infectious individuals converges to the equilibrium frequency, $1 - 1/R_0$, in both the high-dispersal and low-dispersal regimes. We note that, in our stochastic agent-based simulation model with finite population size, this equilibrium is only quasi-stable: Over much longer time scales than considered here, the infectious individuals will eventually vanish from the population.

While a deterministic ODE model predicts that any disease with $R_0 > 1$ should successfully establish after introduction, the stochastic fluctuations in our individual-based simulation model can lead to the loss of the disease in certain simulation runs, even for relatively high R_0 values (e.g., approximately 25% of simulation runs with $R_0 = 3$, as shown in Figure 5A). We did not find a strong influence of the dispersal rate on the probability of disease establishment for $R_0 > 2$. However, in the high-dispersal regime, the frequency of infected individuals approaches the expected equilibrium frequency more rapidly compared to the low-dispersal regime (Figure 5B-D). Overall, we observed that the rate of disease spread can again differ by more than an order of magnitude in our simulation model depending on the dispersal rate (Figure 5C).

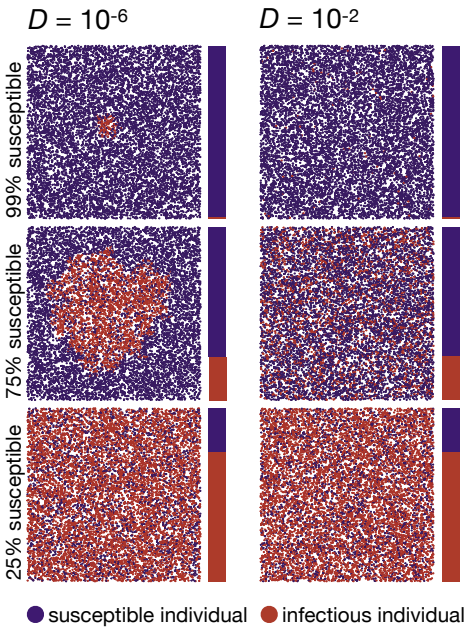

Figure 4: Two exemplary simulation runs in the SIS model. The left column shows a run in the low-dispersal regime ($D = 10^{-6} < D_c$), while the right column shows a run in the high-dispersal regime ($D = 10^{-2}$). The top, middle, and bottom plots show population snapshots taken when 99%, 75%, and 25% of the population were still susceptible. As predicted, in the low-dispersal regime, the disease initially advances by a growing circle. By contrast, in the high-dispersal regime infectious individuals are homogeneously distributed across space (right column). In contrast to the SI model, infectious individuals hardly cluster in the SIS model late in the simulations even in the low-dispersal regime due to the perpetual reentering of infectious individuals into the susceptible class with a probability γ per tick. Data were simulated using the following parameters: $D = [10^{-6}; 10^{-2}]$, $N = 10,000$, $\alpha = 0.01$, $r = 15$, $L = 1$, $\gamma = 1/30$. Videos of simulation runs in the SIS model can be found under `https://tinyurl.com/bdddam58`.

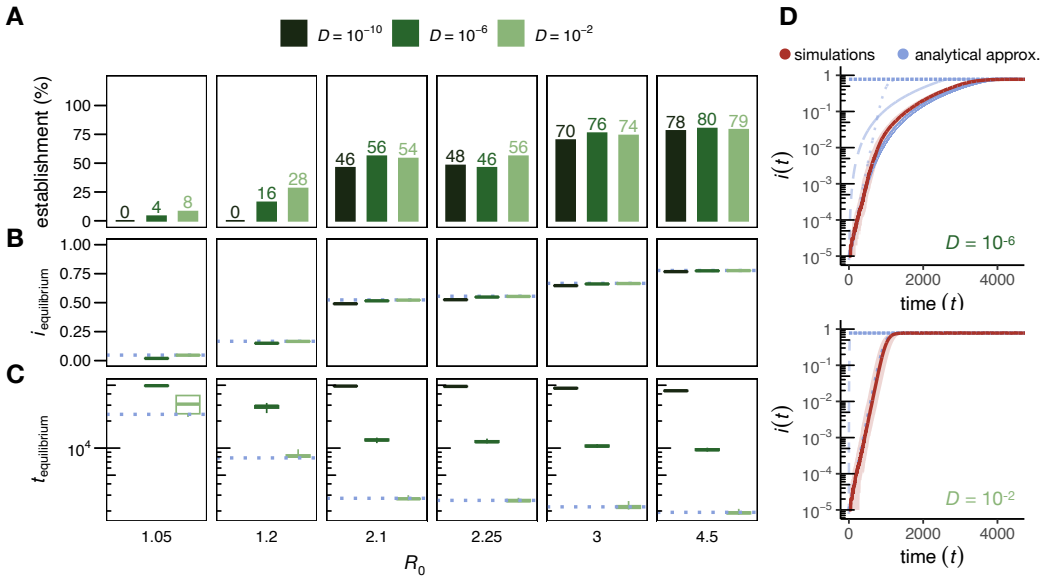

Figure 5: Disease dynamics in the SIS model. (A) Percentage of 50 simulation runs that reached at least a frequency of 1% infectious individuals for varying R_0 values (panels) and three exemplary diffusion coefficients. Only simulation runs with an $i(t)$ of at least 1% are included in (B) to (D). (B) Observed equilibrium frequencies at the end of the simulations. Boxplots are shown with whiskers indicating the most extreme values within $5 \times IQR$ of the boxes, but the variance among runs was so small that they are not visible. The expected equilibrium frequency for each R_0 value under an ODE model [5] is depicted by the dotted blue line. Observed equilibrium frequencies were estimated by fitting simulated data to a three-parameter logistic model using the drc R package [53]. (C) Time to reach the equilibrium. We defined the onset of the equilibrium as the first time the frequency of infectious individuals reaches 99.95% of the observed equilibrium frequency. The expected time of equilibrium onset under an ODE model is depicted as dotted blue line. (D) Time-resolved disease dynamics for $R_0 = 4.5$ and two exemplary diffusion coefficients (top panel: low dispersal with $D = 10^{-6} < D_c = 2.87 \times 10^{-6}$, bottom panel: high dispersal with $D = 10^{-2}$). In the low-dispersal regime, the proportion of infectious individuals, $i(t)$, increases after an initial lag phase approximately as the area of a circle whose radius expands at the predicted minimum wave speed of the diffusion model (blue dashed line). The initial lag between the expected and the observed proportion of infectious individuals is likely caused by the local establishment of the disease and an initial reduction of wave speed due to the curvature of the wavefront [8, 45]. $i(t)$ of the simulations is well approximated by a heuristic approach (solid blue line) that accounts for both disease propagation via a constant traveling wave, and the local rise in $i(t)$ described by the ODE model (dotted line). In the high-dispersal regime, the proportion of infectious individuals is well approximated by the expected logistic growth function of the ODE model with an equilibrium frequency of $1-1/R_0$ [5] (dotted blue line). Solid red lines represent the median proportion of infectious individuals across all simulations, and the shaded area represents the range between the 2.5 and 97.5 percentiles. All data were simulated using the following parameters: $N = 100,000$, $\alpha = 0.001$, $r = 15$, $\gamma = [1/70, 1/80, 1/140, 1/150, 1/200, 1/300]$. More than 99% of our simulation runs categorized as established spread across the entire habitat.

2.7. SIR Model

In the SIS model, individuals can transition between the susceptible and infectious states but do not acquire resistance to the disease. Conversely, the SIR model incorporates lasting resistance by introducing a "recovered state", to which individuals transition after being infected. Once in the recovered state, individuals cannot return to the susceptible or infectious state. According to a deterministic SIR model, a disease is expected to invade a completely susceptible and unstructured population whenever its basic reproduction number is $R_0 > 1$ [4]. In such cases, the final epidemic size (e_{final}), representing the overall proportion of individuals who have been infected during the epidemic, can be obtained by solving the transcendental equation $e^{-R_0 e_{\text{final}}} = 1 - e_{\text{final}}$ [3].

To study the SIR model, we modified our simulations of the SIS model such that infected individuals transition to the recovered class with a constant probability γ per tick, rather than returning to the susceptible class. This modification does not alter the basic reproduction number of the disease, which remains at $R_0 = r\alpha/\gamma$. Qualitatively, we observed that, similar to the SI and SIS models, the high-dispersal regime closely approximates a homogeneously mixed population (Figure 6) whereas in the low-dispersal regime, the disease spreads in a circular pattern across space, as theoretically predicted [54]. However, the traveling wave by which infections spread is now closely followed by another wave describing the emergence of recovered individuals.

The clustering of recovered individuals in proximity to the wavefront of infectious individuals results in an overall reduction in the rate of disease transmission compared to a homogeneously mixed population and has various implications for disease dynamics: First, similar to the SIS model, for $R_0 < 2$ there is a higher likelihood of the disease being eradicated before it can establish in the low-dispersal regime. This results in a lower establishment probability for smaller diffusion coefficients for $R_0 < 2$, where for extremely low diffusion coefficients ($D = 10^{-10}$) the establishment probability remains low until R_0 surpasses 3 (Figure 7A). Second, as previously demonstrated [10, 9], higher values of R_0 are required in the low-dispersal regime to approach the theoretically expected final epidemic size (Figure 7B). If dispersal is limited, the clustering of recovered individuals near the infectious wavefront can give rise to isolated 'pockets' of susceptible individuals devoid of

any contacts with infectious individuals. This phenomenon leads to a delay or even avoidance of their infection, thereby contributing significantly to the reduction in the final epidemic size. This effect is illustrated in Figure 6, where for $D = 10^{-6}$ distinct clusters of susceptible individuals surrounded by recovered ones are observable. Third, due to new infections occurring solely at the wavefront in the low-dispersal regime, the maximal disease incidence ($i_{\max}$), representing the maximum proportion of the population ever infected at a single point in time, is smaller compared to the maximal incidence in the high-dispersal regime (Figure 7C). Fourth, analogous to the SI and SIS models, reduced dispersal slows down disease spread in general. Simulations with a diffusion coefficient of $D = 10^{-6}$ require nearly an order of magnitude more time to reach their respective final epidemic size compared to simulations with $D = 10^{-2}$. Finally, the time-resolved frequency of infectious individuals, $i(t)$, resembles a bell-shaped curve in the high-dispersal regime, as predicted by the ODE model [4]. In the low-dispersal regime, on the other hand, $i(t)$ exhibits almost linear growth until late in the simulation, as is predicted from a circular wave progressing with constant speed [29, 45] (Figure 7D-E).

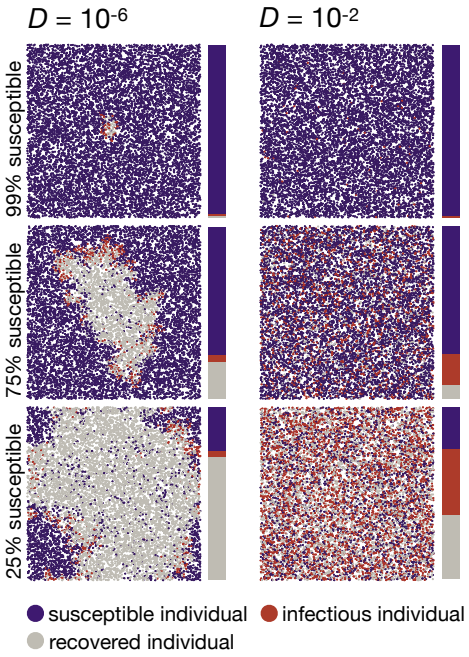

Figure 6: Two exemplary simulation runs in the SIR model. The left column shows a run in the low-dispersal regime ($D = 10^{-6}$), while the right column shows a run in the high-dispersal regime ($D = 10^{-2}$). The top, middle, and bottom plots show population snapshots taken when 99%, 75%, and 25% of the population were still susceptible. As predicted, in the low-dispersal regime the disease advances approximately by a growing circle, with infectious individuals being clustered at the wavefront, and recovered individuals being clustered around the disease origin in the middle of the area. By contrast, in the high-dispersal regime susceptible, infectious, and recovered individuals are homogeneously distributed across space (right column). Data were simulated using the following parameters: $D = [10^{-6}; 10^{-2}]$, $N = 10,000$, $\alpha = 0.01$, $r = 15$, $L = 1$, $\gamma = 1/25$. Videos of simulation runs in the SIR model can be found under `https://tinyurl.com/bdddam58`.

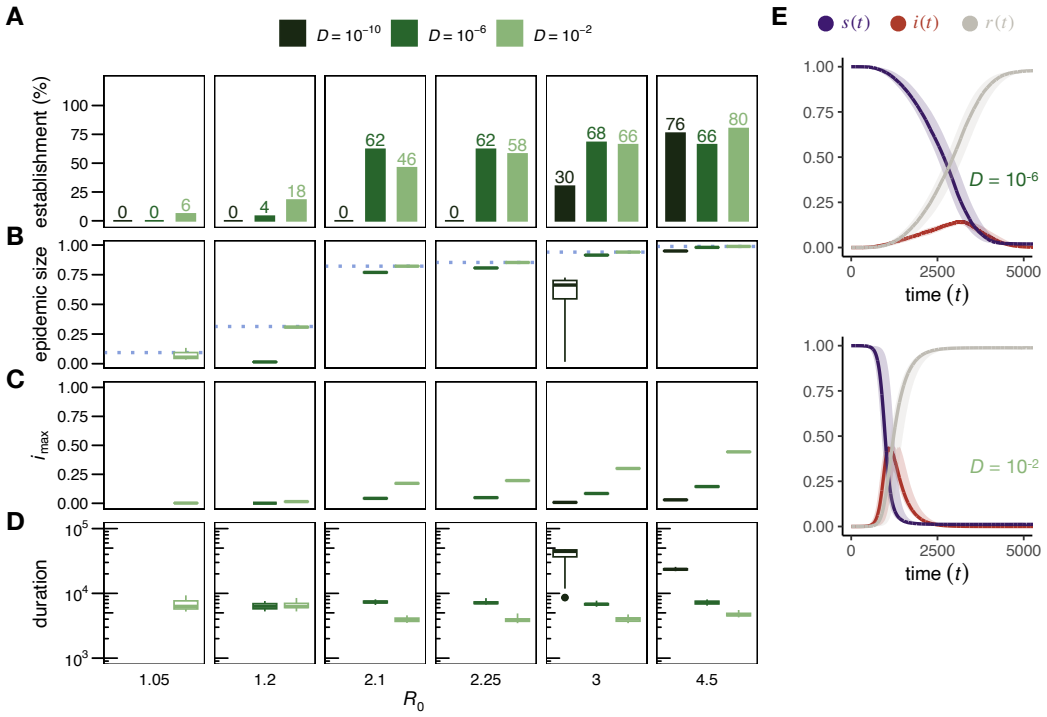

Figure 7: Disease dynamics in the SIR model. (A) Percentage of 50 simulation runs in which the disease reached an epidemic size of at least 1% for varying R_0 values (panels) and three exemplary diffusion coefficients (for $D = 10^{-10}$ and $R_0 = 3$, the spread stagnated in 3 of the 50 runs, which were excluded from the analysis). Disease establishment is less likely in the low-dispersal regime for $R_0 < 2$. Only simulation runs with an epidemic size of at least 1% are included in (B) to (E). (B) Epidemic size (boxplots are shown with whiskers indicating the most extreme values within $5 \times IQR$ of the boxes). Simulations in the low-dispersal regime require higher R_0 values to approach the theoretically predicted epidemic size. (C) Maximal incidence (the variance among runs was so small that whiskers are not visible). Lower dispersal results in smaller $i_{\max}$ values. (D) Epidemic duration. Simulations in the low-dispersal regime can take an order of magnitude longer than in the high-dispersal regime. (E) Time-resolved disease dynamics for $R_0 = 4.5$ and two exemplary diffusion coefficients (top panel: low-dispersal with $D = 10^{-6}$, bottom panel: high-dispersal with $D = 10^{-2}$). In the high-dispersal regime, the disease spreads rapidly and results in an $i_{\max}$ value of around 0.45. In the low-dispersal regime, on the other hand, infectious individuals are confined to the wavefront, resulting in an $i_{\max}$ value of less than 0.20. Solid lines represent the median proportion of susceptible (blue), infectious (red), and recovered (grey) individuals across all 50 simulations, and the shaded areas represent the ranges between the 2.5 and 97.5 percentiles. All data were simulated using the following parameters: $N = 100,000$, $\alpha = 0.001$, $r = 15$, $\gamma = [1/70, 1/80, 1/140, 1/150, 1/200, 1/300]$. More than 99% of our simulation runs categorized as established spread across the entire habitat.

3. Discussion

In this study, we explored the conditions under which it is essential to include spatial structure for precise disease spread predictions when individuals move diffusively across a habitat. We established a critical threshold D_c for the diffusion coefficient, below which disease transmission dynamics is expected to exhibit substantial spatial heterogeneity, while for $D \gg D_c$, the assumption of homogeneous mixing remains adequate. To validate our analytical results, we conducted individual-based simulations on a continuous two-dimensional landscape. Furthermore, we investigated the impact of continuous spatial structure on key epidemiological parameters, including disease establishment probability, maximal incidence, and final epidemic size.

The use of reaction-diffusion models in epidemiology is attractive due to their ability to approximate epidemic states as individuals move and interact in a spatial domain. Furthermore, they often allow calculating the expected speed of disease propagation based on measurable life-history traits, such as individual dispersal, average contact rate, and infection risk [36]. However, modeling diffusive disease transmission with a Fisher-KPP equation and its constant traveling wave solution (as we have done in this study) hinges on several assumptions. First, a traveling wave model with constant speed assumes that disease transmission is a highly localized process [25, 55], with long-range pathogen transmission being very rare. Additionally, distributed-infectives models, as we focused on in this study, assume uniform individual dispersal across the entire population, disregarding possible variations in movement patterns among individuals [32]. In reality, movement patterns can vary significantly and include processes such as long-range dispersal, migration [56], or preferential movement towards specific locations [57]. All of these processes could profoundly affect disease transmission dynamics and substantially limit the accuracy of a traveling wave model assuming a constant speed. For example, if a dispersal kernel is not radially symmetric (as assumed in this study), the advance of the wave front is going to deviate from a spherical shape [8]. If long-distance dispersal is common, the disease can spread from multiple independent locations, where it has been introduced by long-distance migrants [25]. This can ultimately result in a faster spread and accelerating wave speeds over time, as observed for wind-dispersed plant pathogens or avian influenza [45, 27, 33, 58, 25]. In such scenarios, conventional reaction-diffusion models assuming an exponentially

bounded dispersal distribution [59, 58, 38] run the risk of significantly underestimating the spread of diseases [45, 58, 60, 33]. This, in turn, can constrain the ability to identify and optimize appropriate disease interventions [61]. Unfortunately, it can be quite challenging to estimate the occurrence of rare long-distance dispersal events in empirical systems [36].

Second, conventional distributed-infectives reaction-diffusion models often assume that the contact rate and disease establishment proportion are constant for the whole population. However, it is important to recognize that these parameters can exhibit variability among individuals and in response to environmental conditions, which are not explicitly accounted for in our abstract disease model. For instance, the contact rate r can vary significantly among individuals of different age groups [36, 62, 63], the number of contacts can be impeded by complex landscape features such as mountains, rivers, or other barriers [36, 34], and vary with population density or changes in host behavior over time [64]. The disease establishment probability α is often strongly influenced by environmental factors such as temperature, relative humidity, and the level of urbanization [65]. These individual and environmental factors can play a crucial role in shaping the dynamics of disease transmission [66], yet they are not explicitly incorporated into our current model framework.

Third, the conventional Fisher-KPP equation treats space as a continuous and unbounded domain, disregarding edge effects that may exist in real-world habitats with boundaries [28, 29, 5, 51]. To mitigate such edge effects, we implemented periodic boundary conditions in our individual-based simulation framework. This approach ensured that individuals on the boundaries of the simulated population were connected to those on the opposite side, creating a continuous and seamless space. However, one consequence is that susceptible individuals can encounter disease waves closing in from several directions as the disease spreads through the population. Although this may have led to a slight acceleration of disease propagation in scenarios with low dispersal towards the end of the simulation run, the overall impact on the rate of disease fixation was relatively small compared to the total simulation time.

If a biological system meets the assumptions of a reaction-diffusion model with a constant traveling wave solution, this mathematical framework can provide an appealing approach to studying disease dynamics. For exam-

ple, Noble used a reaction-diffusion model to study the spread of the Black Death across 14th-century Europe [32], finding that the analytically predicted patterns aligned well with historical records. Disease transmission was treated as a distributed-infectives process with a value of $r\alpha \approx 1-4$ per year and $D \approx 25,900$ km^2 per year, yielding an expected wave "width" of $2\sqrt{D/(r\alpha)} \approx 161-322$ km, which is relatively small compared to the investigated habitat (the distance between Messina and Oslo is approximately 3,200 km). Furthermore, long-distance dispersal was likely very rare in medieval Europe, supporting the suitability of a constant traveling wave model for studying this particular epidemic. In another study, Murray & Brown used a reaction-diffusion model to study the potential spread of rabies among foxes in the UK [67]. Rabies is transmitted by close contacts and foxes are highly territorial, which limits their individual dispersal. Similar to Noble's work, the $r\alpha$ and D estimates resulted in a relatively small wave "width" when compared to the studied habitat (the southern part of England) [67]. Murray's predictions for a potential spread of rabies in the southern UK were in good agreement with the observed spatial spread of rabies across mainland Europe [67].

While disregarding long-distance dispersal was likely suitable in the aforementioned studies, the scenario should be markedly different for modern human populations. With the growing interconnectedness between human communities, recent epidemics require a comprehensive framework capable of accounting for both short- and long-distance dispersal [68, 58, 38]. Integrating these dispersal mechanisms will be crucial for capturing the realistic dynamics of disease spread and invasion processes [12]. For recent human epidemics, Brockmann *et al.* demonstrated how substituting geographic distances with a probabilistically inspired "effective distance" that accounts for different levels of connectivity between human communities, allows the application of relatively simple diffusion models to study complex disease dynamics [12]. However, one has to be aware that this type of network data are rarely available for non-human species. Furthermore, besides distance and connectivity, individual movement is often influenced by various factors such as behavior, social structures, and environmental conditions [38, 23, 36, 12]. These factors can significantly contribute to the complexity of human movement patterns, making it particularly challenging to robustly integrate them into disease models [69]. If movement patterns are complex, network models or individual-based simulation frameworks might be better suited for study-

ing disease dynamics, although these increasingly complex models may also limit the ability to draw general conclusions [38, 36].

The critical threshold D_c (Equation 9) allows for a back-of-the-envelope calculation of the maximum amount of individual dispersal that would be required to cause a noticeable slowdown in disease spread due to limited dispersal. As an example, consider a simple toy model of a theoretical measles outbreak in a completely susceptible population. Measles is a highly contagious disease that can cause new infections in up to 90% of close contacts ($\alpha = 0.9$) [70]. Assuming an average of four close contacts per day ($r = 4$) [71], and 979 individuals per km^2 (i.e., the average population density in urbanized areas of the US; $N = 979$) [72], we obtain $D_c = 0.002$ km^2 per day. This means that in order for our theoretical measles outbreak to be substantially slowed down due to limited dispersal in an area with $L = 1$ km, the standard deviation of the expected spatial displacement of individuals in each of the x and y coordinates needs to be less than 69 meters per day on average. This toy example demonstrates that spatial dynamics may be omitted for highly contagious diseases over relatively small habitat ranges [50], but have to be accounted for when describing disease dynamics across larger communities such as individual cities [73, 37].

In order to assess the accuracy of our mathematical predictions derived from diffusion theory, we conducted individual-based simulations. This highlighted some important differences and complexities that arise when attempting to simulate a continuous reaction-diffusion process. First, in our simulation framework, we defined contacts between individuals based on an interaction radius parameter. This approach allowed disease spread to occur by "hopping" between neighboring individuals with overlapping interaction areas even when individuals remained stationary. As a result, there is an upper bound on the fixation time in our individual-based simulations, which may not be immediately evident in other disease models. One might consider mitigating the influence of the interaction radius on the disease dynamics by implementing a k-nearest-neighbor approach to model contacts. However, caution should be exercised when using a small value of k (e.g., $k = 1$), in combination with very low dispersal, as this combination could introduce a significant delay in disease spread if the neighbor structure changes only gradually.

Second, to mitigate edge effects, we incorporated periodic boundaries into our simulation framework. Unlike real-world scenarios, where distinct boundaries can profoundly affect disease spread, simulations with periodic boundaries enable us to isolate the impact of individual dispersal on disease transmission. In contrast to non-periodic boundaries, which may introduce bias based on the initial positions of the first infected individuals relative to the edges, periodic boundaries remove this potential for bias by removing the concept of edges altogether. Reflecting boundaries, however, may be more appropriate for simulating specific real-world scenarios.

In conclusion, our study highlights the need for careful interpretation and understanding of spatial factors in epidemiological dynamics, while also demonstrating the complexities and design choices that arise in modeling such dynamics.

4. Data Accessibility

The individual-based SI, SIS, and SIR disease transmission models are implemented in the open-source software SLiM (version 4.0) [40] and are available on GitHub under `https://github.com/AnnaMariaL/SpatialDieaseSim`. Videos demonstrating disease spread in structured and unstructured populations are available on YouTube under `https://tinyurl.com/bdddam58`.

5. Acknowledgements

We thank Brandon D. Hollingsworth for valuable feedback and discussion, and Benjamin C. Haller for helpful advice on the implementation of the individual-based simulation framework. This project has received funding from the European Union's Horizon2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101025586.

References

[1] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 115 (772) (1927) 700–721. `doi:10.1098/rspa.1927.0118`.

[2] R. M. Anderson, R. M. May, Infectious diseases of humans: dynamics and control, Oxford University Press Oxford; New York, 1991.

[3] J. D. Murray, Mathematical Biology I. An Introduction, 3rd Edition, Vol. 17 of Interdisciplinary Applied Mathematics, Springer, 2002. doi:10.1007/b98868.

[4] H. W. Hethcote, Three Basic Epidemiological Models, Springer Berlin Heidelberg, Berlin, Heidelberg, 1989, Ch. 1, pp. 119–144.

[5] H. W. Hethcote, The mathematics of infectious diseases, SIAM Review 42 (4) (2000) 599–653. doi:10.1137/S0036144500371907.

[6] M. M. López-Flores, D. Marchesin, V. Matos, S. Schecter, Differential Equation Models in Epidemiology, IMPA, 2021.

[7] N. C. Grassly, C. Fraser, Mathematical models of infectious disease transmission, Nature Reviews Microbiology 6 (6) (2008) 477–487. doi:10.1038/nrmicro1845.

[8] M. Lewis, P. Kareiva, Allee dynamics and the spread of invading organisms, Theoretical Population Biology 43 (2) (1993) 141–158. doi:https://doi.org/10.1006/tpbi.1993.1007.

[9] D. Mollison, Dependence of epidemic and population velocities on basic parameters, Mathematical Biosciences 107 (2) (1991) 255–287. doi:https://doi.org/10.1016/0025-5564(91)90009-8.

[10] D. A. P. Grenfell B. T., Ecology of Infectious Diseases in Natural Populations, Cambridge University Press, 1995. doi:10.1017/CBO9780511629396.

[11] S. Riley, Large-scale spatial-transmission models of infectious disease, Science 316 (5829) (2007) 1298–1301. doi:10.1126/science.1134695.

[12] D. Brockmann, D. Helbing, The hidden geometry of complex, network-driven contagion phenomena, Science 342 (6164) (2013) 1337–1342. doi:10.1126/science.1245200.

[13] C. Rozins, M. J. Silk, D. P. Croft, R. J. Delahay, D. J. Hodgson, R. A. McDonald, N. Weber, M. Boots, Social structure contains epidemics and regulates individual roles in disease transmission in a group-living mammal, Ecology and Evolution 8 (23) (2018) 12044–12055. doi:https://doi.org/10.1002/ece3.4664.

[14] C. Buckee, A. Noor, L. Sattenspiel, Thinking clearly about social aspects of infectious disease transmission, Nature 595 (7866) (2021) 205–213. doi:10.1038/s41586-021-03694-x.

[15] V. Capasso, Mathematical structures of Epidemic Systems, Springer-Verlag, 1993.

[16] S. Riley, K. Eames, V. Isham, D. Mollison, P. Trapman, Five challenges for spatial epidemic models, Epidemics 10 (2015) 68–71, challenges in Modelling Infectious Disease Dynamics. doi:https://doi.org/10.1016/j.epidem.2014.07.001.

[17] F. Brauer, C. Castillo-Chavez, Z. Feng, Spatial structure in disease transmission models, Mathematical Models in Epidemiology 69 (2019) 457–476. doi:10.1007/978-1-4939-9828-9_14.

[18] A. Lajmanovich, J. A. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population, Mathematical Biosciences 28 (3) (1976) 221–236. doi:https://doi.org/10.1016/0025-5564(76)90125-5.

[19] V. Belik, T. Geisel, D. Brockmann, Natural human mobility patterns and spatial spread of infectious diseases, Phys. Rev. X 1 (2011) 011001. doi:10.1103/PhysRevX.1.011001.

[20] A. Wesolowski, T. Qureshi, M. F. Boni, P. R. Sundsøy, M. A. Johansson, S. B. Rasheed, K. Engø-Monsen, C. O. Buckee, Impact of human mobility on the emergence of dengue epidemics in pakistan, Proceedings of the National Academy of Sciences 112 (38) (2015) 11887–11892. doi:10.1073/pnas.1504964112.

[21] J. I. Blanford, Z. Huang, A. Savelyev, A. M. MacEachren, Geo-located tweets. enhancing mobility maps and capturing cross-border movement, PLOS ONE 10 (6) (2015) 1–16. doi:10.1371/journal.pone.0129202.

[22] T. A. Perkins, R. C. Reiner, Jr., G. España, Q. A. ten Bosch, A. Verma, K. A. Liebman, V. A. Paz-Soldan, J. P. Elder, A. C. Morrison, S. T. Stoddard, U. Kitron, G. M. Vazquez-Prokopec, T. W. Scott, D. L. Smith, An agent-based model of dengue virus transmission shows how uncertainty about breakthrough infections influences vaccination impact projections, PLOS Computational Biology 15 (3) (2019) 1–32. doi:10.1371/journal.pcbi.1006710.

[23] R. C. Reiner, S. T. Stoddard, T. W. Scott, Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal, Epidemics 6 (2014) 30–36. doi:https://doi.org/10.1016/j.epidem.2013.12.003.

[24] N. M. Ferguson, C. A. Donnelly, R. M. Anderson, Transmission intensity and impact of control policies on the foot and mouth epidemic in great britain, Nature 413 (6855) (2001) 542–548. doi:10.1038/35097116.

[25] M. J. Keeling, K. T. Eames, Networks and epidemic models, Journal of The Royal Society Interface 2 (4) (2005) 295–307. doi:10.1098/rsif.2005.0051.

[26] J. Rushmore, D. Caillaud, L. Matamba, R. M. Stumpf, S. P. Borgatti, S. Altizer, Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk, Journal of Animal Ecology 82 (5) (2013) 976–986. doi:https://doi.org/10.1111/1365-2656.12088.

[27] N. Shigesada, K. Kawasaki, Biological invasions: Theory and practice, Oxford University Press, 1997.

[28] H. Hotelling, A mathematical theory of migration, University of Washington, republished in Environment and Planning A 10 1223–1239 (1921).

[29] R. A. Fisher, The wave of advance of advantageous genes, Annals of Eugenics 7 (4) (1937) 355–369. doi:10.1111/j.1469-1809.1937.tb02153.x.

[30] A. Kolmogorov, I. Petrovskii, N. Piscunov, A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem, Byul. Moskovskogo Gos. Univ. 1 (6) (1937) 1–25.

[31] T. Wang, Dynamics of an epidemic model with spatial diffusion, Physica A 409 (2014) 119–129. doi:10.1016/j.physa.2014.04.028.

[32] J. V. Noble, Geographic and temporal development of plagues, Nature 250 (5469) (1974) 726–729. doi:10.1038/250726a0.

[33] C. C. Mundt, K. E. Sackett, L. D. Wallace, C. Cowger, J. P. Dudley, Long-distance dispersal and accelerating waves of disease: Empirical relationships., The American Naturalist 173 (4) (2009) 456–466. doi:10.1086/597220.

[34] D. L. Smith, B. Lucey, L. A. Waller, J. E. Childs, L. A. Real, Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes, Proceedings of the National Academy of Sciences 99 (6) (2002) 3668–3672. doi:10.1073/pnas.042400799.

[35] J. Novembre, A. P. Galvani, M. Slatkin, The geographic spread of the ccr5 Δ32 hiv-resistance allele, PLOS Biology 3 (11) (10 2005). doi:10.1371/journal.pbio.0030339.

[36] A. Hastings, K. Cuddington, K. F. Davies, C. J. Dugaw, S. Elmendorf, A. Freestone, S. Harrison, M. Holland, J. Lambrinos, U. Malvadkar, B. A. Melbourne, K. Moore, C. Taylor, D. Thomson, The spatial spread of invasions: new developments in theory and evidence, Ecology Letters 8 (1) (2005) 91–101. doi:https://doi.org/10.1111/j.1461-0248.2004.00687.x.

[37] D. Chen, Modeling the Spread of Infectious Diseases: A Review, John Wiley and Sons, Ltd, 2014, Ch. 2, pp. 19–42. doi:https://doi.org/10.1002/9781118630013.ch2.

[38] M. C. Steiner, J. Novembre, Population genetic models for the spatial spread of adaptive variants: A review in light of SARS-CoV-2 evolution, PLOS Genetics 18 (9) (2022) 1–22. doi:10.1371/journal.pgen.1010391.

[39] D. Mollison, The rate of spatial propagation of simple epidemics, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp.

579–614.
URL http://projecteuclid.org/euclid.bsmsp/1200514358

[40] B. C. Haller, P. W. Messer, Slim 4: Multispecies eco-evolutionary modeling, The American Naturalist 201 (5) (2023) E127–E139. doi:10.1086/723601.

[41] A. Tsoularis, J. Wallace, Analysis of logistic growth models, Mathematical Biosciences 179 (1) (2002) 21–55. doi:10.1016/S0025-5564(02)00096-2.

[42] E. Hunter, J. D. Kelleher, Understanding the assumptions of an SEIR compartmental model using agentization and a complexity hierarchy, Journal of Computational Mathematics and Data Science 4 (2022) 100056. doi:10.1016/j.jcmds.2022.100056.

[43] H. Scherm, On the velocity of epidemic waves in model plant disease epidemics, Ecological Modelling 87 (1) (1996) 217–222. doi:10.1016/0304-3800(95)00030-5.

[44] A. Hastings, Models of spatial spread: A synthesis, Biological Conservation 78 (1) (1996) 143–148. doi:10.1016/0006-3207(96)00023-7.

[45] N. Shigesada, K. Kawasaki, Y. Takeda, Modeling stratified diffusion in biological invasions, The American Naturalist 146 (08 1995). doi:10.1086/285796.

[46] N. Tkachenko, J. D. Weissmann, W. P. Petersen, G. Lake, C. P. E. Zollikofer, S. Callegari, Individual-based modelling of population growth and diffusion in discrete time, PLOS ONE 12 (4) (2017) 1–22. doi:10.1371/journal.pone.0176101.

[47] P. C. Fife, Mathematical aspects of reacting and diffusing systems, Springer, 1979.

[48] M. Bramson, Convergence of solutions of Kolmogorov equation to travelling waves, American Mathematical Society, 1983.

[49] Q. Zhuang, J. Wang, A spatial epidemic model with a moving boundary, Infectious Disease Modelling 6 (2021) 1046–1060. doi:https://doi.org/10.1016/j.idm.2021.08.005.

[50] M. Kot, Elements of Mathematical Ecology, Cambridge University Press, West Nyack, 2001.

[51] R. Mazzucco, M. Doebeli, U. Dieckmann, The influence of habitat boundaries on evolutionary branching along environmental gradients, Evolutionary Ecology 32 (12 2018). doi:10.1007/s10682-018-9956-1.

[52] K. Dietz, The estimation of the basic reproduction number for infectious diseases, Statistical Methods in Medical Research 2 (1) (1993) 23–41. doi:10.1177/096228029300200103.

[53] C. Ritz, F. Baty, J. C. Streibig, D. Gerhard, Dose-response analysis using r, PLOS ONE 10 (12) (2016) 1–13. doi:10.1371/journal.pone.0146021.

[54] J. Cox, R. Durrett, Limit theorems for the spread of epidemics and forest fires, Stochastic Processes and their Applications 30 (2) (1988) 171–191. doi:https://doi.org/10.1016/0304-4149(88)90083-X.

[55] J. Medlock, M. Kot, Spreading disease: integro-differential equations old and new, Mathematical Biosciences 184 (2) (2003) 201–222. doi:https://doi.org/10.1016/S0025-5564(03)00041-5.

[56] S. N. Busenberg, C. C. Travis, Epidemic models with spatial spread due to population migration, Journal of Mathematical Biology 16 (2) (1983) 181–198. doi:10.1007/BF00276056.
URL https://doi.org/10.1007/BF00276056

[57] T. C. Reluga, J. Medlock, A. P. Galvani, A model of spatial epidemic spread when individuals move within overlapping home ranges, Bulletin of Mathematical Biology 68 (2) (2006) 401–416. doi:10.1007/s11538-005-9027-y.
URL https://doi.org/10.1007/s11538-005-9027-y

[58] M. Kot, M. A. Lewis, P. van den Driessche, Dispersal data and the spread of invading organisms, Ecology 77 (7) (1996) 2027–2042. doi:https://doi.org/10.2307/2265698.

[59] J. G. Skellam, Random dispersal in theoretical populations, Biometrika 38 (1-2) (1951) 196–218. doi:10.1093/biomet/38.1-2.196.

[60] M. Turelli, A. A. Hoffmann, Rapid spread of an inherited incompatibility factor in California Drosophila, Nature 353 (6343) (1991) 440–442. doi:10.1038/353440a0.

[61] M. J. Tildesley, P. R. Bessell, M. J. Keeling, M. E. J. Woolhouse, The role of pre-emptive culling in the control of foot-and-mouth disease, Proceedings of the Royal Society B: Biological Sciences 276 (1671) (2009) 3239–3248. doi:10.1098/rspb.2009.0427.

[62] P. Rohani, X. Zhong, A. A. King, Contact network structure explains the changing epidemiology of pertussis., Science 330 (6006) (2010) 982–985. doi:10.1126/science.1194134.

[63] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, W. J. Edmunds, Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases, PLOS Medicine 5 (3) (03 2008). doi:10.1371/journal.pmed.0050074.

[64] E. Volz, L. A. Meyers, Susceptible–infected–recovered epidemics in dynamic contact networks, Proceedings of the Royal Society B: Biological Sciences 274 (1628) (2007) 2925–2934. doi:10.1098/rspb.2007.1159.

[65] J. N. S. Eisenberg, M. A. Desai, K. Levy, S. J. Bates, S. Liang, K. Naumoff, J. C. Scott, Environmental determinants of infectious disease: a framework for tracking causal links and guiding public health research, Environmental health perspectives 115 (8) (2007) 1216–1223. doi:10.1289/ehp.9806.

[66] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence, Nature 438 (7066) (2005) 355–359. doi:10.1038/nature04153.

[67] J. D. Murray, E. A. Stanley, D. L. Brown, On the spatial spread of rabies among foxes, Proceedings of the Royal Society of London. Series B. Biological Sciences 229 (1255) (1986) 111–150. doi:10.1098/rspb.1986.0078.

[68] J. Paulose, J. Hermisson, O. Hallatschek, Spatial soft sweeps: Patterns of adaptation in populations with long-range dispersal, PLOS Genetics 15 (2) (2019) 1–29. doi:10.1371/journal.pgen.1007936.

[69] S. A. Lee, C. I. Jarvis, W. J. Edmunds, T. Economou, R. Lowe, Spatial connectivity in mosquito-borne disease models: a systematic review of methods and assumptions, Journal of The Royal Society Interface 18 (178) (2021) 20210096. doi:10.1098/rsif.2021.0096.

[70] Cdc.gov: Global measles outbreaks [online] (11 2022) [cited 2023-06-04].

[71] D. M. Feehan, A. S. Mahmud, Quantifying population contact patterns in the United States during the COVID-19 pandemic, Nature Communications 12 (1) (2021) 893. doi:10.1038/s41467-021-20990-2.

[72] Census.gov: Urban areas facts [online] (10 2021) [cited 2023-06-04].

[73] Grenfell, B. T. and Bjørnstad, O. N. and Kappey, J., Travelling waves and spatial hierarchies in measles epidemics, Nature 414 (6865) (2001) 716–723. doi:10.1038/414716a.