Failure of R_0 near the epidemic threshold in the classical SIS model

Mozzamil Mohammed

Mathematical Modelling Department, Al Neelain University, Khartoum, Sudan

September 18, 2023

Abstract

The primary predictor of a disease outbreak and severity is the basic reproduction number R_0, which represents the average number of secondary cases produced by introducing an infected individual into an entirely susceptible population. According to the classical SIS model, a disease with R_0 less than one will eventually die out and persist if R_0 is greater than one. Using the pair-approximation method, we reconstruct the classical SIS model by explicitly accounting for neighbourhood interactions between susceptible and infected individuals. Specifically, the disease can only be transmitted, with some transmission probability, if a susceptible individual is surrounded by at least one infected individual within its direct neighborhoods. Despite the simplicity of the SIS model present here, results produced by the pair-approximation model deviates significantly from predictions by the mean-field approximation model, particularly near the epidemic threshold. Contrasting the standard SIS model based on the mean-field approach, we find scenarios where the disease dies out even if R_0 is greater than one. We suggest a crucial need for redefining the basic reproduction number on a smaller spatial scale and taking the average R_0 over a global scale, rather than applying it globally to an entire population. However, in the realm of more intricate models of infectious diseases, it remains an open question to what extent mean-field approximation predictions diverge from predictions produced by models that consider neighborhood interactions.

Keywords: Infectious disease modelling; The basic reproduction number; Pair approximation

Correspondence to: mozzamilm@gmail.com

1 Introduction

The basic reproduction number R_0 quantifies the average number of secondary infections produced by an infected individual when introduced into a completely susceptible population (Dietz, 1993). The basic reproduction number has long been considered as import predictor of a disease spread and severity, and often plays an important role for supporting decisions of public health officials and policymakers (Guerra et al., 2017; Zhao et al., 2020). Estimation of the basic reproduction number is typically complex as it depends on various biological and environmental factors (Delamater et al., 2019). However, in the simple classical susceptible-infected-susceptible (SIS) model, the basic reproduction number emerges from just two key parameters and is defined as the average disease transmission rate divided by average recovery rate per an infected individual (Allen, 1994). It is commonly accepted that a disease with R_0 less than 1 will eventually die out and persist if R_0 is greater than 1. The epidemic threshold, occurring at R_0 equal to 1, represents the boundary between the extinction and sustained transmission of the disease. The outcome of a disease transmission away from the epidemic threshold might be intuitively straightforward, depending on the transmission complexity.

However, estimation of R_0 for the same disease can vary as the model set up and complexity change. Depending on the method used to calculate R_0, many diseases with an R_0 value greater than 1 are predicted to die out and persist when R_0 is less than 1, leading to a failure of R_0 (Li et al., 2011). The basic reproduction number is primarily defined on a global scale, often disregarding the small local scale on which interactions between susceptible and infected individuals occur. In reality, an infected individual cannot infect a susceptible one if they are spatially distant. Nevertheless, it is commonly assumed by the mean-field approximation method that all individuals within a population do interact with one another, and this simplification has been widely accepted. Network and pair-approximation models of infectious diseases typically account for interactions between neighbouring pairs of individuals (Keeling, 1999; Keeling and Eames, 2005; Payne, 2019). Such models often describe the dynamics of infectious disease more accurately than the mean-field approximation models. The pair-approximation method has been employed to describe systems other than infectious diseases where neighborhood interactions are crucial determinants of the dynamics outcomes (Harada and Iwasa, 1994; Mohammed et al., 2023, 2018). This particular methodology correctly captures systems dynamics predicted lattice-based models, offering mathematical tractability.

Using the pair-approximation method, we reconstruct the classical SIS model by explicitly accounting for neighborhood interactions among susceptible and infected individuals. The pair-approximation approach assumes that a susceptible individual can only get infected if it is surrounded by at least one infected individual present in its direct neighbourhoods. That is a necessary condition for a disease transmission, but the transmission occurs with some probability, following an interaction between susceptible and infected individuals. The core object of this work is to assess whether the R_0 concept applies when neighbourhood interactions are taken into account in infectious disease models. Although the SIS model presented here is simple, but results of the pair-approximation model deviates significantly from previous predictions by the mean-field approximation model. Contrasting the SIS model that based mean-field approximation, the pair-approximation model exhibits a scenario where a disease can die out even if R_0 is greater than 1. This divergence in the predictions of the two models occurs mainly around the epidemic threshold. More complex models of infection diseases based on the mean-field approximation are thus expected to diverge more significantly from models that explicitly account for neighbourhood interactions. We suggest that estimations of the basic reproduction number should be redefined on a small neighbourhood scale, rather than on a global scale for the entire population. Estimation of the basic reproduction number might not be enough to inform public health officials about the current and future behaviour of the disease outbreak dynamics, and this fundamental concept of R_0 can be misleading by omitting some of its important properties such as spatial scales.

2 Methods

2.1 Pair-approximation model of disease transmission

Pair approximation is a method of constructing ordinary differential equation models, describing the dynamics of neighbouring pairs, the global and local dynamics of a given population (Harada and Iwasa, 1994). It provides an explicit description for neighborhood interactions among individuals and accurately capture the spatiotemporal dynamics predicted by lattice-based models (Mohammed et al., 2023).

Here, we use the pair-approximation method to reconstruct the classical SIS model by explicitly considering neighbourhood interactions among susceptible and infected individuals (Keeling, 1999; Payne, 2019). We denote by P_S and P_I the probabilities that a randomly chosen individual is susceptible and infected, respectively. These probabilities represent the global frequencies

(S/N) and (I/N) of susceptible and infected individuals, where S and I are the number of susceptible and infected individuals in a population of size $N = S + I$, with $P_S + P_I = 1$. The disease is transmitted locally when a susceptible individual comes into a direct contact with an infected individual present in its own neighborhood. The pair-approximation method describes the average local frequency of infected individuals present in the neighborhood of a susceptible individual. The number of nearest-neighbouring sites of a susceptible individual is quantified by the positive integer z. The average local frequency of infected individuals in the nearest neighborhoods of a susceptible individual is quantified by $q_{I|S}$, representing the conditional probability that a randomly chosen individual in the neighborhood of a susceptible individual is infected (Harada and Iwasa, 1994). The local frequency of infected individuals is the number of infected individuals in the neighborhood of a susceptible individual, averaged over the number of nearest-neighbouring sites z. A susceptible individual can be infected if at least one of neighbours in the z sites is infected.

The disease transmission rate is denoted as β. The infected individual recovers at rate γ in a period of time $1/\gamma$. In the SIS model, a susceptible individual is assumed to become susceptible again after recovery.

The differential equations governing the global disease dynamics are described by

$$\dot{P}_S = -\beta q_{I|S} P_S + \gamma P_I, \tag{1}$$

$$\dot{P}_I = \beta q_{I|S} P_S - \gamma P_I. \tag{2}$$

The first term in Eqns. (1) and (2) describes the neighborhood interaction between susceptible and infected individuals. The last term describes the overall recovery rate. The conditional probability $q_{I|S}$ needs to be specified. It is mathematically defined as $q_{I|S} = P_{SI}/P_S$ (Harada and Iwasa, 1994), where the pair $P_{SI} = P_{IS}$ represents the joint probability that a randomly chosen neighbouring pair of individuals are susceptible and infected. Notably, there are four possibilities of choosing a neighbouring pair of individuals at random: P_{II}, P_{SI}, P_{IS}, and P_{SS}, with $P_{II} + 2P_{SI} + P_{SS} = 1$.

Following (Keeling, 1999; Payne, 2019), we construct differential equations governing the dynamics of neighbouring pairs. The following transitions contribute to the dynamics of the P_{SI} pair:

1. A pair of infected individuals P_{II} contributes positively to the pair P_{SI} if one of the infected individuals in the P_{II} pair has been recovered
2. A pair of susceptible individuals P_{SS} contributes positively to the pair P_{SI} if one of the susceptible individuals in the P_{SS} pair has been infected by one of its direct neighbour from outside the pair, i.e., from the remaining nearest-neighbouring sites $(z-1)$.
3. A pair of susceptible and infected individuals P_{SI} contributes negatively to the pair P_{SI} if the susceptible individual in the P_{SI} pair has been infected by its direct neighbour in the pair
4. A pair of susceptible and infected individuals P_{SI} contributes negatively to the pair P_{SI} if the susceptible individual in the P_{SI} pair has been infected by one of its direct neighbour from outside the pair, i.e., from the remaining nearest-neighbouring sites $(z-1)$.
5. A pair of susceptible and infected individuals P_{SI} contributes negatively to the pair P_{SI} if the susceptible individual in the P_{SI} pair has been recovered

Using the pair-approximation approach (Harada and Iwasa, 1994; Keeling, 1999; Mohammed et al., 2023), the dynamics of the P_{SI} pair are thus described by

$$\dot{P}_{SI} = 2\gamma P_{II} + 2\frac{\beta}{z}(z-1)q_{I|SS}P_{SS} - 2\frac{\beta}{z}P_{SI} - 2\frac{\beta}{z}(z-1)q_{I|SI}P_{SI} - 2\gamma P_{SI}. \quad (3)$$

To reduce the model complexity, we replace the neighbouring pair P_{II} by $(1 - P_{SS} - 2P_{SI})$, and we only need to describe the P_{SS} pair dynamics to construct an equation for the average local frequency $q_{I|S}$. In Eqn. (3), the first term indicates the transition from P_{II} to either P_{SI} or P_{IS}, thus the factor 2. In the subsequent term, the quantity β/z represents disease transmission rate per infected individual present in the neighborhood of a susceptible individual, and the transition from P_{SS} to either P_{SI} or P_{IS} will require an infected individual from the remaining neighbouring sites $(z-1)$, thus the multiplication by the conditional probability $q_{I|SS}$. This conditional probability would require the computation of higher-order frequencies, and would not allow to obtain a closed system of differential equations. Therefore, pair-approximation method ignores the effect of the neighbour-of-the neighbour (it assumes correlation is weaker with non-direct neighbours, i.e., correlation between sites decreases exponentially with the distance between them, (see, e.g., Harada and Iwasa, 1994) and approximates $q_{I|SS} \approx q_{I|S}$ and $q_{I|SI} \approx q_{I|S}$. That is, correlations with non-nearest neighbours are approximately reconstructed from nearest neighbour correlations. Similar explanations apply to the other terms in Eqn. (3).

The following transitions contribute to the dynamics of the P_{SS} pair:

1. A pair of susceptible and infected individuals P_{SI} contributes positively to the pair P_{SS} if the susceptible individual in the pair P_{SI} has been recovered
2. A pair of susceptible individuals P_{SS} contributes negatively to the pair P_{SS} if any of the susceptible individuals in the pair P_{SS} has been infected by its direct neighbour from outside the pair, i.e., from the remaining nearest-neighbouring sites $(z-1)$.

The dynamics of the neighbouring pair P_{SS} are thus described by

$$\dot{P}_{SS} = 2\gamma P_{SI} - 2\frac{\beta}{z}(z-1)q_{I|SS}P_{SS}. \quad (4)$$

Using the definition of the local frequency of infected individuals $q_{I|S} = P_{SI}/P_S$, and taking the time derivative of both sides, the dynamics of the average local frequency of infected individuals in the neighborhoods of a susceptible individual are described by

$$\dot{q}_{I|S} = -P_{SI}\dot{P}_S/P_S^2 + \dot{P}_{SI}/P_S \quad (5)$$

Equations (1) – (5) represent the reconstructed SIS model under neighborhood interactions among susceptible and infected individuals present in the population. Equations (1) and (2) corresponds to the mean-field approximation model when $q_{I|S}$ is replaced by P_I.

2.2 The basic reproduction number

As defined in the traditional SIS model (Daley and Gani, 2001), the basis reproduction number is given as $R_0 = \beta/\gamma$. According to the mean-field SIS model, a disease with $R_0 < 1$ will die out and persist if $R_0 > 1$. The bifurcation point occurs at $R_0 = 1$, separating the extinction and persistence regions of the disease.

2.3 Numerical analysis

The numerical simulation of the pair-approximation and the mean-field approximation SIS models has been implemented in MATLAB. Parameter values used to simulate the model are provided in *Table 1*.

Table 1: Variable and parameter descriptions, and numerical values used in the model.

Symbol	Description	Reference value	Range
P_S	Global frequency of susceptible individuals		[0, 1]
P_I	Global frequency of infected individuals		[0, 1]
P_{SS}	Global frequency of susceptible pair		[0, 1]
P_{II}	Global frequency of infected pair		[0, 1]
P_{SI}	Global frequency of susceptible and infected pair		[0, 1]
$q_{I\|S}$	The average local frequency of infected individuals in the neighborhood of a susceptible individual		[0, 1]
β	Disease transmission rate	0.2	$[0, \infty)$
γ	Recovery rate	0.02-0.9	$[0, \infty)$
R_0	The basic reproduction number	0.2-10	$[0, \infty)$
z	The number of nearest-neighbouring sites	4	$[0, \infty)$

3 Results

We showcased the SIS model simulation for the global frequencies of susceptible and infected individuals, P_S and P_I, respectively. The dynamics of neighbouring pairs (P_{SI} and P_{SS}) and the average local dynamics of infected individuals present in the neighborhood of a susceptible individual, $q_{I|S}$. We compared results produced by the pair-approximation model with previous predictions by the classical mean-field SIS model.

In the pair-approximation model, the number of nearest-neighbouring sites of a susceptible individual z reflects the potential number of susceptible and infected individuals present within the direct neighbourhood of a susceptible individual. In the mean-field approximation model, a susceptible individual has the same likelihood of interacting with all infected individuals present within a population. That is a very unlikely situation and can affect predictions of disease outbreaks. However, it is intuitive to assume that a susceptible individual can typically get infected when interacts with an infected individual present within its direct four neighbouring sites (i.e., $z = 4$).

We first restricted our model to four nearest-neighbouring sites. We found that results produced by the pair-approximation model significantly deviated from predictions by the mean-field approximation. This deviation occurred, particularly, close to the epidemic threshold, $R_0 = 2.5$ and $R_0 = 1.17$ (Fig. 1b and c). For large and small values of R_0 (10 and 0.8), corresponding to the sustained transmission and extinction of the disease, the two models produced very similar results (Fig. 1a and d).

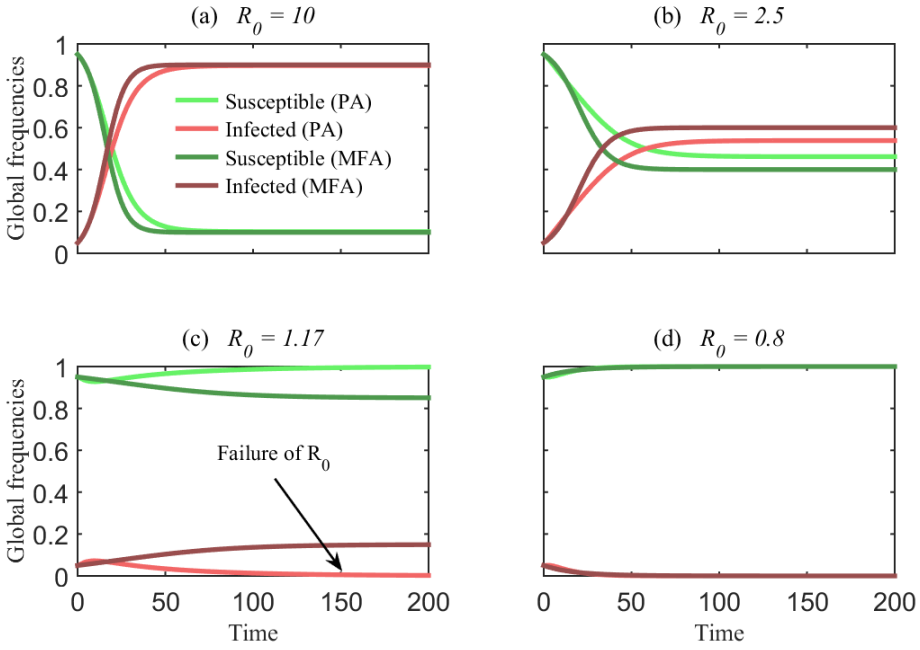

Fig. 1. The temporal dynamics of the global frequencies of susceptible (P_S) and infected (P_I) individuals under four scenarios of the R_0 value (10, 2.5, 1.17, and 0.8). The global frequencies of the pair-approximation (PA) model are represented by the bright-green and red curves, respectively, while the dark curves correspond to the mean-field approximation (MFA) model. The x-axis represents the simulation time, and the y-axis represents the global frequencies. The initial conditions: $P_S(0) = 0.95$, $P_I(0) = 1 - P_S(0)$, $P_{SI}(0) = 0.05$, $P_{SS}(0) = 0.8$, and $q_{IS}(0) = P_{SI}(0)/P_S(0)$. Parameters used: $\beta = 0.2$, $R_0 = 10, 2.5, 1.17,$ and 0.8, $\gamma = \beta/R_0$, and $z = 4$.

According to the classical mean-field SIS model, the disease can spread out and persist when $R_0 = \beta/\gamma > 1$, and this principle has been widely accepted for this specific model. However, results of the proposed pair-approximation model demonstrated scenarios where the disease dies out even if $R_0 > 1$ (Fig. 1c; and Fig. 3a). This divergence is likely driven by the consideration of neighborhood interactions between susceptible and infected individuals in the pair-approximation model. In fact, this is the primary difference between the mean-field and the pair-approximation models presented in this work.

Furthermore, we simulated the dynamics of the reconstructed SIS model for a large number of nearest-neighboring sites ($z = 100$) of a susceptible individual. This indicates that a susceptible individual can potentially interact with 100 individuals in its surroundings, an unlikely situation. Some of these 100 individuals can be infected and can transmit the disease to this susceptible individual at transmission rate β. By considering a very large number of individuals in the neighborhoods of a susceptible individual, the pair-approximation model corresponds to the mean-field approximation model. Therefore, our results demonstrated almost an identical behaviour of the disease outbreak dynamics, produced by the pair-approximation and the mean-field approximation models (Fig. 2; and Fig. 3b). The two models approached the same long-term attractor of the disease, with a small difference in the initial transient phase.

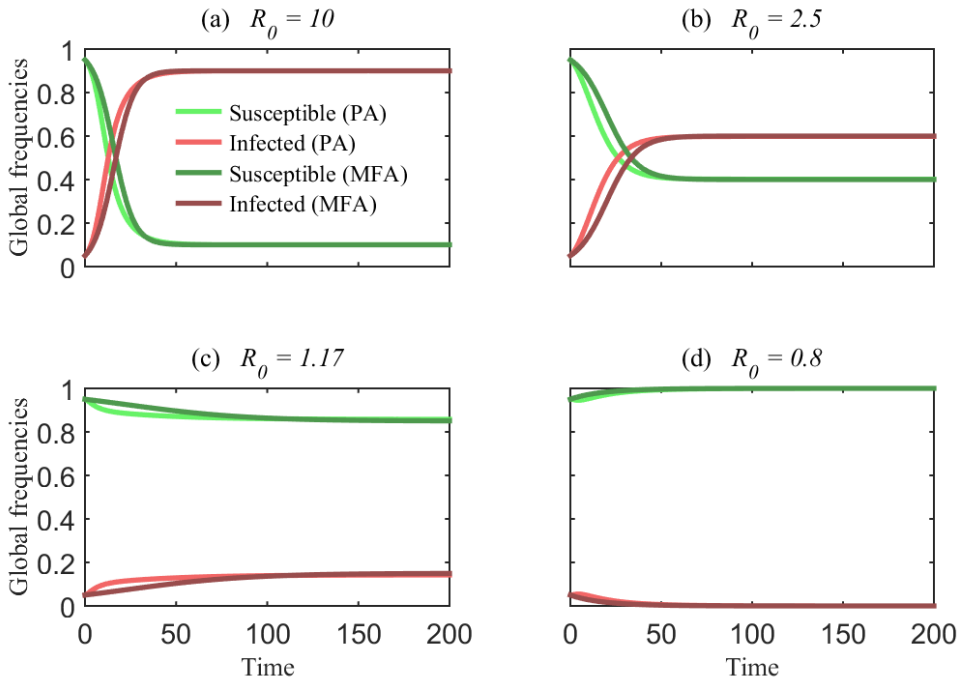

Fig. 2. The temporal dynamics of the global frequencies of susceptible (P_S) and infected (P_I) individuals under four scenarios of the R_0 value (10, 2.5, 1.17, and 0.8). The description, initial conditions, and parameters used are the same as in the caption of figure 1, and $z = 100$.

We systematically assessed the influence of the basic reproduction number on the epidemic size and compared the corresponding results of the two models. Away from the epidemic threshold ($R_0 = 1$), the outcome of the disease transmission is intuitively straightforward. Of a particular interest is to predict the disease transmission near the epidemic threshold. Our findings underscored significant differences between the results of the two models (Fig. 3a), especially around the epidemic threshold.

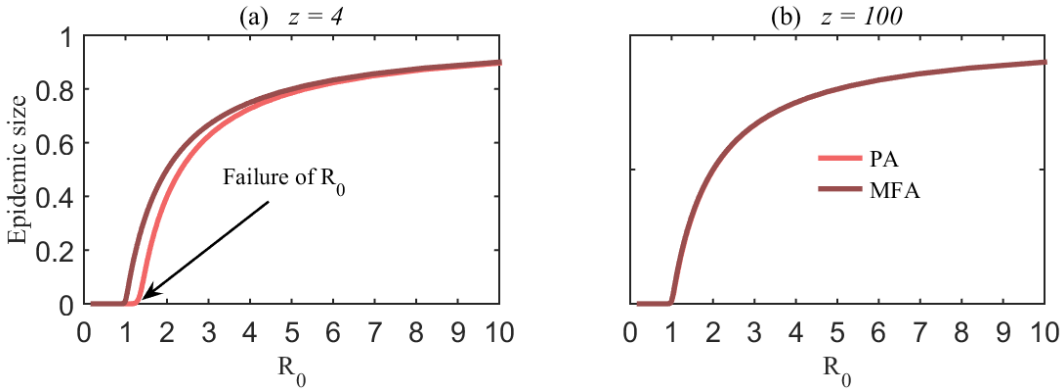

Fig. 3. The equilibrium dynamics of the epidemic size (y-axis) as a function the basic reproduction number R_0 (x-axis), predicted by the pair-approximation (PA) model (bright curve) and the mean-field approximation (MFA) model (dark curve). Initial conditions and parameter values are the same as in the caption of figure 1.

At small to moderate values of R_0, the epidemic size predicted by the pair-approximation model is lower than that predicted by the mean-field approximation model (Fig. 3a). More precisely, the epidemic size curve produced by the pair-approximation model is shifted forward, extending the disease extinction region. Notably, this forward shift emerged only under the

consideration of four nearest-neighbouring individuals ($z = 4$). The results of the two models produced the exact dependence of the epidemic size on the basic reproduction number R_0 when the number of nearest-neighbouring sites is chosen to be very large, $z = 100$ (Fig. 3b).

The temporal dynamics for the average local frequency of infected individuals within the neighbourhood a susceptible individual, and the dynamics of neighbouring pairs are shown for a number of nearest-neighbouring sites $z = 4$ (Fig. 4) and $z = 100$ (Fig. 5).

Results demonstrated that the average local frequency of infected individuals decreases as the basic reproduction number R_0 increases (Fig. 4 and 5). A small number of infected individuals are present in the neighbourhoods of each susceptible individual when the basic reproduction number is small. Thus, the probability of choosing an infected individual from the neighbourhood of a susceptible individual is low. The average local frequency of infected individuals will also reflect the probability of choosing a neighbouring pair of susceptible and infected individuals. The probability of randomly choosing a neighbouring pair of susceptible and infected P_{SI} depends on the absolute difference between the global probabilities P_S and P_I. The probability of choosing the pair P_{SI} is small when the absolute difference is high, and vice versa. The probability of randomly choosing the neighbouring pair P_{SS} typically increases as the global frequency P_S increases. Thus, a small values of R_0 will lead to higher probabilities of a neighbouring pair of susceptible individuals (Fig. 4). However, under the consideration of very large number of nearest-neighbouring sites, the probability of choosing an infected individual from the neighbourhood of a susceptible individual is relatively high when R_0 is slightly greater than one (Fig. 5c).

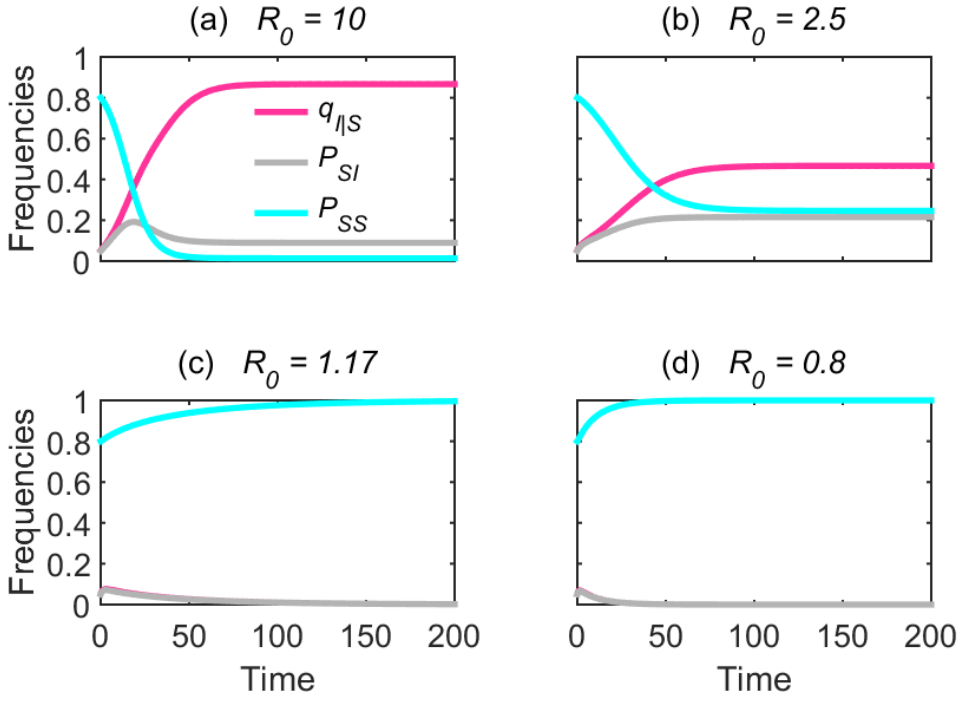

Fig. 4. The temporal dynamics of the average local frequency ($q_{I|S}$, pink curve), neighbouring pairs of susceptible and infected individuals (P_{SI}, grey curve), and neighbouring pairs of susceptible individuals (P_{SS}, cyan curve) predicted by the pair-approximation model under four scenarios of the R_0 value (10, 2.5, 1.17, and 0.8). The x-axis represents the simulation time, and the y-axis represents the frequencies. The initial conditions: $P_S(0) = 0.95$, $P_I(0) = 1 - P_S(0)$, $P_{SI}(0) = 0.05$, $P_{SS}(0) = 0.8$, and $q_{IS}(0) = P_{SI}(0)/P_S(0)$. Parameters used: $\beta = 0.2$, $R_0 = 10, 2.5, 1.17,$ and 0.8, $\gamma = \beta/R_0$, and $z = 4$.

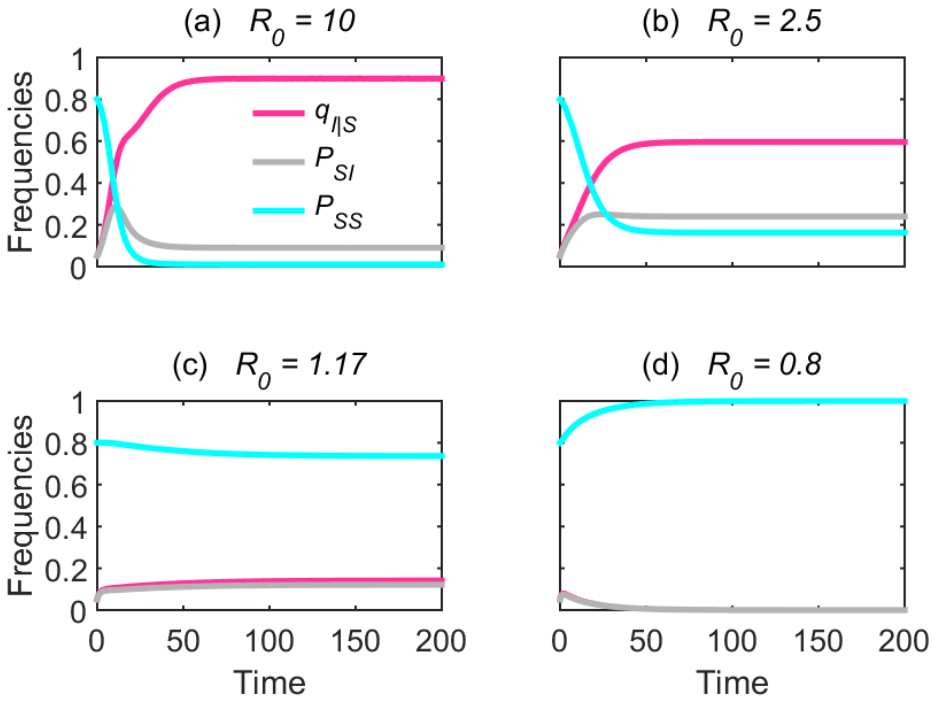

Fig. 5. The temporal dynamics of the average local frequency ($q_{I|S}$, pink curve), neighbouring pairs of susceptible and infected individuals (P_{SI}, grey curve), and neighbouring pairs of susceptible individuals (P_{SS}, cyan curve) predicted by the pair-approximation model under four scenarios of the R_0 value (10, 2.5, 1.17, and 0.8). Initial conditions and parameter values are the same as in the caption of figure 4.

4 Conclusion

The basic reproduction number R_0 has been used for several decades as an indicator of whether the disease will die out or persist. It is widely acknowledged that the disease is likely to die out if $R_0 < 1$, while persist if $R_0 > 1$. In this work, we reconstructed the classical SIS model using the pair-approximation method (Harada and Iwasa, 1994; Keeling, 1999; Payne, 2019), and presented a theoretical evidence that a disease with $R_0 > 1$ dies out in a simple SIS model. Despite this simplicity of the model, results produced by the pair-approximation model deviated pronouncedly from predictions by the mean-field approximation model. The advantage of the pair-approximation approach over the commonly-used mean-field approximation is that it explicitly accounts for neighborhood interactions among susceptible and infected individuals. However, a recent pair-approximation SIS model produced results in alignment with predictions by the classical mean-field SIS model, yet diverging from our findings in terms of the dependence of epidemic size on R_0 (Payne, 2019). That model, while informative, solely focused on dynamics of pairs and did not consider the average local frequency dynamics of infected individuals; thus, overlooked the effect of the number of nearest-neighboring sites of a susceptible individual on the disease transmission dynamics. In the realm of more intricate models of infectious diseases, it remains an open question to what extent mean-field approximation predictions deviate from predictions produced by models that consider neighborhood interactions (Keeling, 1999). We suggest a crucial need for redefining the basic reproduction number on a smaller spatial scale and taking the average R_0 over a global scale, rather than applying it globally to an entire population. This redefinition may depend on the average number of infected individuals within the neighbourhoods of a susceptible individual (Keeling, 1999). This fundamental epidemiological concept of R_0 might otherwise be misleading (Li et al., 2011).

Data availability

This study does not include or use data.

Declaration of competing interest

None.

Declaration of funding

None.

References

Allen, L.J.S., 1994. Some discrete-time SI, SIR, and SIS epidemic models. Mathematical Biosciences 124, 83–105.

Daley, D.J., Gani, J., 2001. Epidemic Modelling: An Introduction. Cambridge University Press.

Delamater, P.L., Street, E.J., Leslie, T.F., Yang, Y.T., Jacobsen, K.H., 2019. Complexity of the Basic Reproduction Number (R0). Emerg Infect Dis 25, 1–4.

Dietz, K., 1993. The estimation of the basic reproduction number for infectious diseases. Stat Methods Med Res 2, 23–41.

Guerra, F.M., Bolotin, S., Lim, G., Heffernan, J., Deeks, S.L., Li, Y., Crowcroft, N.S., 2017. The basic reproduction number (R0) of measles: a systematic review. The Lancet Infectious Diseases 17, e420–e428.

Harada, Y., Iwasa, Y., 1994. Lattice population dynamics for plants with dispersing seeds and Vegetative propagation. Res Popul Ecol 36, 237–249.

Keeling, M.J., 1999. The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society of London. Series B: Biological Sciences 266, 859–867.

Keeling, M.J., Eames, K.T.D., 2005. Networks and epidemic models. Journal of The Royal Society Interface 266, 859-867.

Li, J., Blakeley, D., Smith, R.J., 2011. The Failure of R 0. Computational and Mathematical Methods in Medicine 2011, 1–17.

Mohammed, M., Brännström, Å., Landi, P., Dieckmann, U., 2023. Fruit harvesting: A potential threat to the persistence, spatial distribution, and establishment of plants. Preprint at https://arxiv.org/abs/2309.01711.

Mohammed, M.M.A., Landi, P., Minoarivelo, H.O., Hui, C., 2018. Frugivory and seed dispersal: Extended bi-stable persistence and reduced clustering of plants. Ecological Modelling 380, 31–39.

Payne, J.L., 2019. Pair Approximations, in: Hastings, A., Gross, L. (Eds.), Encyclopedia of Theoretical Ecology. University of California Press, 531–534.

Zhao, S., Lin, Q., Ran, J., Musa, S.S., Yang, G., Wang, W., Lou, Y., Gao, D., Yang, L., He, D., Wang, M.H., 2020. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. International Journal of Infectious Diseases 92, 214–217.