

# The scaling of contact rates with population density for the infectious disease models

Hao Hu\*, Karima Nigmatulina, Philip Eckhoff

*Epidemiological Modeling (EMOD) Group, Intellectual Ventures Laboratory, 1555 132nd Ave. NE, Bellevue, WA 98005, USA*

## ARTICLE INFO

### Article history:

Received 31 August 2011  
Received in revised form 28 April 2013  
Accepted 30 April 2013  
Available online 9 May 2013

### Keywords:

Contact rate  
Epidemic model  
Transmission scaling  
Population density  
Crowd dynamics

## ABSTRACT

Contact rates and patterns among individuals in a geographic area drive transmission of directly-transmitted pathogens, making it essential to understand and estimate contacts for simulation of disease dynamics. Under the uniform mixing assumption, one of two mechanisms is typically used to describe the relation between contact rate and population density: density-dependent or frequency-dependent. Based on existing evidence of population threshold and human mobility patterns, we formulated a spatial contact model to describe the appropriate form of transmission with initial growth at low density and saturation at higher density. We show that the two mechanisms are extreme cases that do not capture real population movement across all scales. Empirical data of human and wildlife diseases indicate that a nonlinear function may work better when looking at the full spectrum of densities. This estimation can be applied to large areas with population mixing in general activities. For crowds with unusually large densities (e.g., transportation terminals, stadiums, or mass gatherings), the lack of organized social contact structure deviates the physical contacts towards a special case of the spatial contact model – the dynamics of kinetic gas molecule collision. In this case, an ideal gas model with van der Waals correction fits well; existing movement observation data and the contact rate between individuals is estimated using kinetic theory. A complete picture of contact rate scaling with population density may help clarify the definition of transmission rates in heterogeneous, large-scale spatial systems.

© 2013 The Authors. Published by Elsevier Inc. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

## 1. Introduction

For directly transmitted diseases, contact patterns drive the spread of infectious pathogens in both space and time. Recent data-driven, detailed, large-scale spatial models and contact experiments have improved our understanding of contact networks in different social settings, as well as the corresponding spread patterns of infectious diseases [1–15]. However, the generalized scaling rules of host contact patterns over population size or density still need refinement. In epidemic models targeting a large heterogeneous area, it becomes important to set the appropriate transmission rates to different local populations.

The population effect on transmission, e.g., threshold of pathogen invasion and persistence [16], has frequently been examined as a way to differentiate pathogen transmissibility in geographic locations. Stochastic effects [17] make the invasion threshold difficult to observe. However, the persistence threshold has successfully been identified in measles endemics during the

pre-vaccination era, when a minimum population threshold called Critical Community Size (CCS) affected disease persistence or fade-outs in the United Kingdom [18]. In turn, studies explored a number of scaling mechanisms in dynamic epidemic models to characterize the relation between transmission rate and population [19–24].

At the other end of the density spectrum, if a location has extremely high densities in a local area (for example in transportation terminals, stadiums, or mass gatherings), it might pose high risks that may facilitate disease spread, and recently a research agenda has been proposed to explore the role of mathematical modeling in assessing health risks associated with mass gatherings [25]. However, in such settings, the contact rates between individuals remain unexplored and it is unclear whether a scaling mechanism relationship exists between transmission rates and population density. If a spatial epidemic model needs to include such locations and quantitatively assess the risk of disease outbreak, it might not be appropriate to extrapolate the rules under general mixing to crowds. In this case, people have very low mean free path and do not have the ability to contact everyone else. Understanding the contact patterns in the extremely high density range is important, and appropriate scaling of contacts and transmissibility needs to be applied.

We aim to use simple physical models to explore the complete spectrum of contact rate scaling with population density. A spatial

\* Corresponding author. Tel.: +1 425 691 3336.

E-mail address: [hhu@intven.com](mailto:hhu@intven.com) (H. Hu).

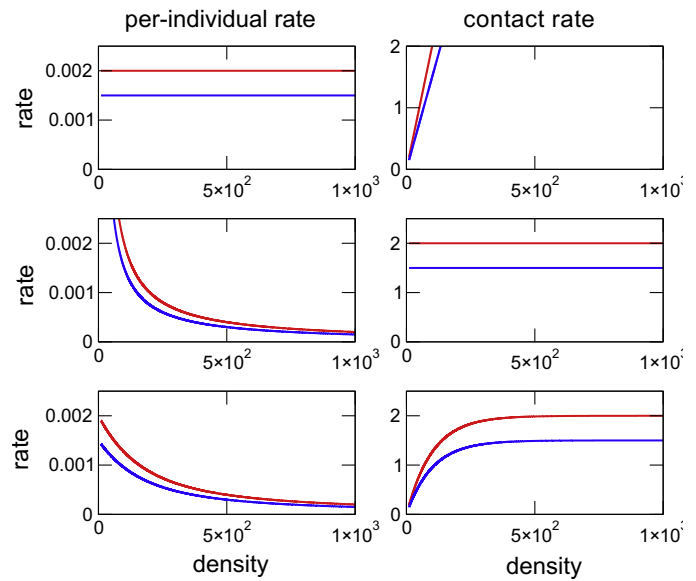
model is used to construct a function which monotonically increases with density, but saturates as density increases. We show that the acquired density scaling function is comparable to a number of physical models in different activity range limits, such as kinetic collision theory and homogeneous mixing. At the high population density limit, when looking at the mixture of crowds, it is possible to approximate crowd motion as random movements, using a corrected model from kinetic collision theory. It is a special example of the density scaling function to capture the change of contact rates at the high density limit. A number of empirical examples are presented to illustrate the use of the density scaling function: in low-density conditions, this scaling function is fitted to a number of human and wildlife diseases: the mortality data of the 1918 influenza outbreak in the United States and England & Wales, home range and transmission threshold density of red fox rabies, contact rates in brushtail possums and voles for the transmission of bovine tuberculosis and cowpox. At the high density limit, the contact rate fits well with observations from different mass gathering events.

## 2. Contact rate scaling functions

### 2.1. Density and frequency dependent mechanisms

In a homogeneously mixed, single location SIR model with population  $N$  inside area  $A$  ( $S, I$ , and  $R$  representing the number of susceptible, infected, and recovered individuals respectively), the *per capita* transmission rate  $\beta$  controls the rate of disease spread.  $\beta$  is the product of infection probability  $p$ , and individual contact rate  $C$ , therefore  $\beta = pC$ . Since the proportion of infectious contacts is  $\frac{1}{N}$ , the rate of increase for infectious individuals is  $\frac{dI}{dt} = \beta S \frac{I}{N} = \frac{pCSI}{\rho A}$ , given  $N = \rho A$ . Homogeneous mixing assumption results in a constant contact rate  $C$  within a local population, and it depends on the population density:  $C \sim C(\rho)$  [23].  $C$  is typically determined by density- or frequency-dependent mechanisms, both are widely used and are applicable to a range of infectious diseases [19,21]. Density-dependent transmission is also called ‘pseudo’ mass action [22]. Borrowing from the law of mass action in chemistry, it is analogous to the scaling between reaction rates and reactant concentrations. With this approach, the contact rate scales in a linear fashion with population density (or size);  $C(\rho) = c_0\rho$ , where  $c_0$  is a constant, and  $\frac{dI}{dt} = \frac{pC_0}{A} SI \sim cSI$  assuming  $c$  is a constant. The per-link contact rate is a constant  $c_0$ . Estimating the density-dependent contact rates from collision frequency in the ideal gas model has the same result (See Section 2.3). Frequency-dependent transmission, also called ‘true’ mass action [22] or proportionate mixing [26], assumes a constant contact rate in the local population independent of population size or density,  $C = c_1$ , so  $\frac{dI}{dt} = p c_1 S \frac{I}{N} \sim c \frac{SI}{N}$ . In this case, the per-link contact rate,  $\frac{c_1}{N}$ , decreases with larger density. Fig. 1(A) and (B) show examples of transmission scaling with population density under the above two mechanisms.

Choice of mechanism makes no difference when investigating a single fixed size population as the parameters can be calibrated appropriately, but patterns of disease spread will differ when modeling a large heterogeneous spatial area. The reproductive number  $R_0$  for frequency-dependent transmission is fixed regardless of population density, and therefore, does not have a population invasion or persistence threshold for the given parameters because  $R_0 > 1$ . For density-dependent transmission, the actual reproductive number depends on the population size and an invasion threshold. The mechanism, however, assumes indefinite linear contact rate growth with the population size, and it is questionable whether an individual has more contacts in a population of  $10^6$  than  $10^5$  [24,27,28].



**Fig. 1.** Different scaling functions of transmission rates with population density: (A) density-dependent. (B) frequency-dependent. (C) Mechanistic nonlinear scaling function. For comparison, all plots have the same ranges. The area is assumed to be  $10 \text{ km}^2$ . The reproductive number for frequency-dependent function is chosen as  $R_0 = 2.0$  (red) and  $R_0 = 1.5$  (blue). For density-dependent function, we assume a population threshold of  $100 \text{ individuals/km}^2$  ( $10^3$  individuals), therefore the constant per-individual rate is calculated using  $R_0 = 1$  when density =  $100 \text{ individuals/km}^2$ . For nonlinear scaling function, the saturation values are chosen the same as  $R_0$  for direct comparison. (For interpretation of the references to colors in this figure caption, the reader is referred to the web version of this article.)

### 2.2. The nonlinear transmission scaling function

Based on existing evidences, it is possible that neither mechanism may adequately describe the correct scaling over the entire population density spectrum. Therefore, a nonlinear function is usually preferred [16,29,30], such as the Holling Type 2 relationship in ecological predator–prey models [31,32] or a power function with population [30,33]. These non-linear functions have the expected features on both limits, and they continuously shift with host density. A number of mathematical forms have been explored [19–21].

Here, we aim to use a spatial contact model with constant population density and limited activity range to estimate contact rates. Within a local population, individuals move around and contact with each other randomly, and a spatial dispersal kernel controls the decay behavior of contact rate. Since the decay usually depends on both density and activity range, a dispersal kernel function  $\varphi_\rho(r)$  belonging to exponential power distribution family is selected in order to calculate the contact rate within the activity range.

We assume a land area with uniform population density  $\rho$ , and each individual has a constant fraction  $k$  ( $0 < k < 1$ ) of contacts among the calculated effective population. Physical contact of each individual is confined to a maximum activity range  $r_{\max}$ . Contact rate  $C$  within the reachable area is calculated by the spatial kernel  $\varphi_\rho(r)$ , density  $\rho$  and maximum activity range  $r_{\max}$ :

$$C(\rho) = \int \varphi_\rho(r) dr d\theta = \int_0^{r_{\max}} \varphi_\rho(r) 2\pi r dr$$

A number of disease transmission models include a distance-based transmission kernel to capture population movement and the effect of spatial clustering. For instance, previous research shows examples of the functional decay based on daily travel distances, in this case a spatial kernel is included to calculate the risks of community infection [3,4]. Other models, such as that of foot-and-mouth disease (FMD) in the United Kingdom in 2001, proves that a

distance-explicit exponential decay kernel captures the progression of FMD in a heterogeneous space [34]. Furthermore, studies of contact rate heterogeneities within populations have demonstrated that these can also alter transmission dynamics [35,36]. Using data from either dollar bill circulation or cell phone locations as proxies [37–39], population-level movement studies describe the quantitative scaling of people's movement patterns with distance. All of these insights may help facilitate an appropriate form of spatial kernel  $\varphi_\rho(r)$  to describe the pattern of transmission.

We define an exponential power distribution with limited activity range to estimate how contact rates change with density and distance. In the function, contact rate decays with distance, and the relationship between home range and population density tends to be inverse. Therefore, we start with the generalized decay form of spatial kernel with density and distance dependence,

$\varphi_\rho(r) = k\rho e^{-c\left(\frac{\rho}{\rho_0}\right)^\alpha} r^\beta$ . After integrating the spatial kernel to calculate the contact rate, we have:

$$C(\rho) = \int_0^{r_{\max}} k\rho e^{-c\left(\frac{\rho}{\rho_0}\right)^\alpha} r^\beta 2\pi r dr$$

A numerical study is performed on different  $\alpha$  and  $\beta$ . Fig. 2 shows the effect of different  $\alpha$  on the contact scaling function under different pre-selected  $\beta$  values. Under all conditions, the scaling function approaches a constant value only when  $\beta = 2\alpha$  is satisfied. Therefore, in order to have a saturated contact rate at high density, the exponential powers on density and distance must satisfy:

$$\varphi_\rho(r) = k\rho e^{-c\left(\frac{\rho}{\rho_0}\right)^\alpha} r^{2\alpha}$$

$\alpha$  ( $\alpha > 0$ ) defines the scaling laws of different physical movement models. Choosing  $\alpha = 1$  corresponds to a distribution of distance in Gaussian form; it represents the physical process of random walk. For different  $\alpha$ , the integration of the spatial kernel gives similar saturation function forms as shown in Fig. 2.

When  $\alpha = 1$ , spatial kernel density becomes:  $\varphi_\rho(r) = k\rho e^{-\frac{\rho}{\rho_0}r^2}$ .  $\rho_0$  is the characteristic density constant, which is determined by the type of disease. Integrating the contact rate function, we obtain:

$$C(\rho) = \pi k \rho_0 \left(1 - e^{-\frac{r_{\max}^2 \rho}{\rho_0}}\right)$$

In the following discussions, we use the above form for kernel density function as an example. A nonlinear function approximates density and frequency dependent mechanisms at low- and high-density ends in the general population mixing range, respectively:

$$C(\rho) = \pi k \rho_0 \left(1 - e^{-\frac{r_{\max}^2 \rho}{\rho_0}}\right) \sim c_1 (1 - e^{-c_2 \rho})$$

where  $c_1$  and  $c_2$  are constants. As density increases, this functional form results has initial growth, and it is saturated when  $\rho$  is high. It

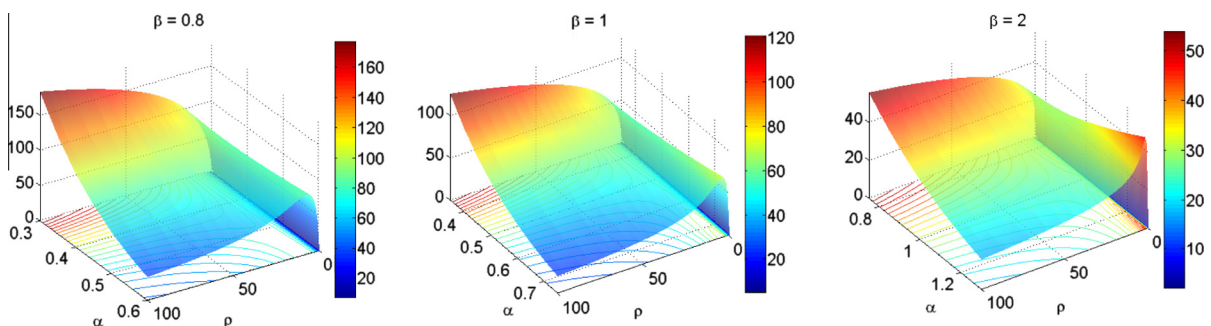
is easy to get the limits of both extreme conditions using the above function: given  $\rho_0$ , the contact rate will approach density and frequency dependent limits at  $\rho \rightarrow 0$  ( $C(\rho) \sim \pi k r_{\max}^2 \rho$ , and therefore,  $C(\rho) \sim \rho$ ) and large  $\rho$  ( $C(\rho) \rightarrow \pi k \rho_0$ ), respectively. Fig. 1(C) has an example describing the behavior of this nonlinear function form, and show the contact rate satisfies the two existing mechanisms in different density limits.

A simulation of endemic disease die-outs illustrates the impact of different choices of density scaling mechanisms, including the non-linear scaling function. This agent-based, stochastic, discrete-time simulation is performed with 100,000 individuals, but with different densities between 10 and 200/km<sup>2</sup> (area ranges from 500–10,000 km<sup>2</sup>) to represent a broad range of places from sparsely-populated areas to crowded metropolitans. For each selected density, 10 random simulations are executed using a SEIRS model, and we apply the repetitive infected case importation rates (5 per 60 days) to all scenarios. We run the simulation for 730 days and calculate the length of disease die-outs, i.e. the fraction of time when no new case is reported. Latent and infectious period is set to 3 days and to 4 days, respectively. People lose disease-specific immunity after 60 days on average, and at  $R_0 = 1.5$ , the disease will remain endemic. Saturated value of  $R_0$  for the non-linear scaling function is set to 1.5 and we set the same threshold density ( $R_0 = 1$  when  $\rho_{\text{threshold}} = 32.96/\text{km}^2$ ) to both density-dependent and non-linear scaling relationships. In Fig. 3, it is apparent to see the limits of both frequency- and density-dependent transmission, while the non-linear scaling relationship scales nicely in a wide spectrum of densities. For frequency-dependent transmission, very few die-outs are observed, because  $R_0$  remains constant. Both density-dependent and non-linear scaling transmission are able to show the effect of threshold density when  $R_0 < 1$ , however, for density-dependent transmission,  $R_0$  will increase to unrealistically high values when density is large, in this case disease dies out by itself because all susceptible individuals are depleted quickly.

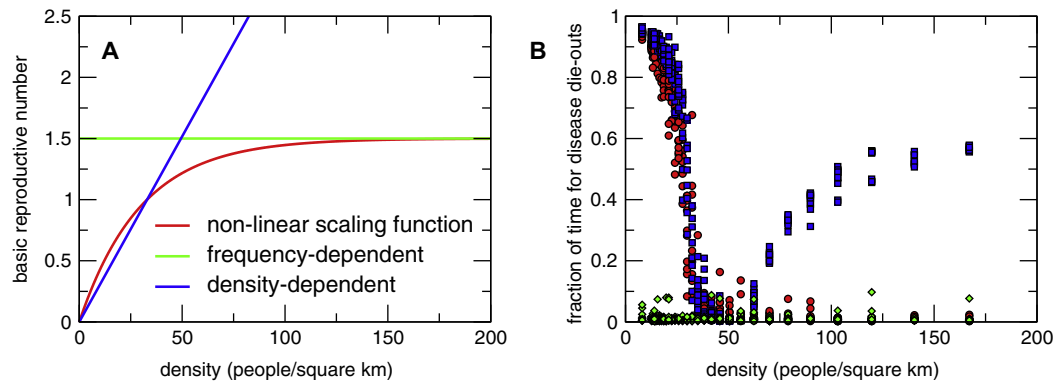
In the above discussions, we assume that spatial kernel  $\varphi$  depends on both density  $\rho$  and distance  $r$ , and  $r_{\max}$  is a constant. If  $\varphi$  only depends on  $r$ , the same scaling functions and results can be achieved by assuming a density dependent activity range  $r_{\max}(\rho)$ , where the maximum home range decreases with greater density.

### 2.3. Maximum range, width of spatial kernel and kinetic models

Both dispersal kernel function and maximum range act together to create a mechanistic model of contact dynamics for both human social networks and wildlife dynamics. In human social networks, daily life involves a certain number of human-to-human contacts which saturates at high density. People have greater rates of contacting individuals near them spatially, but still contact individuals who live further away. In lower-density areas, people will travel further on average per-contact, but there is a maximum range for



**Fig. 2.** Numerical studies on different exponents of density and distance in the spatial kernel  $\varphi_\rho(r) = k\rho e^{-c\left(\frac{\rho}{\rho_0}\right)^\alpha} r^\beta$ .  $\beta$  is pre-selected, and we show the effect on the saturation of the scaling function when selecting different  $\alpha$ . In all cases, the surface and contour plot on  $\rho$  and  $\alpha$  is reported. Only when  $\alpha = \frac{1}{2}\beta$ , the function approaches a constant at high density. The example uses  $r_{\max} = 5$ ,  $\rho_0 = 10$ , and  $k = 1$ . The results were calculated with different distance decay exponent: (A)  $\beta = 0.8$ , (B)  $\beta = 1$ , (C)  $\beta = 2$ .



**Fig. 3.** Effect of different contact rate scaling functions in a SEIRS model. (A) Three scaling functions used in the simulation: density-dependent, frequency-dependent and non-linear scaling function. Threshold density was set as  $\rho_{\text{threshold}} = 32.96/\text{km}^2$  for both density-dependent and non-linear scaling functions, and saturated  $R_0 = 1.5$  is used for both frequency-dependent and non-linear scaling functions. (B) fraction of time with disease die-outs in the simulation. The plot shows results from 10 random stochastic simulations for each density. The average waning time of immunity is 60 days, so that the disease will remain endemic at  $R_0 = 1.5$ . 5 infected cases are seeded every 60 days and we run the simulation for 730 days. For other disease-specific parameters please refer to the main text.

effective routine contacts. In wildlife populations, higher densities decrease the effective area per animal, and the density kernel shrinks in width. Animals are limited by mobility and thus at low densities, the maximum effective range limits contact rates and contact rates transition to a density-dependent regime.

The activity range  $r_{\text{max}}$  represents the proximity of practical contact distances for each individual. It is also a parameter to control the pattern of spread: whether it is well-mixed or wave-like, as these patterns have different rules of transmission [40]. When  $r_{\text{max}} \rightarrow \infty$ ,  $C(\rho) = \pi k \rho_0$  describes a well-mixed model, every individual is able to contact everyone else. In this case, transmission is controlled by the overall contact rate, regardless of population density. The interpretation of this limit is that people have a constant contact rate regardless of population density, and travel as far as necessary to achieve that contact rate. When  $r_{\text{max}} \rightarrow 0$ ,  $C(\rho) = \pi k r_{\text{max}}^2 \rho$ , the spread pattern becomes wave-like, which is controlled by the network connectivity with nearest neighbors. Therefore, it scales directly with density, assuming that connectivity scales to all neighbors within this minimal effective distance. In this case, this wave-like behavior is well demonstrated by the ideal gas kinetic model. When population density is within the general population mixing range, the volume of individual is negligible compared to the mean free path, and we assumed the container was large enough so that each individual had no restrictions. Therefore, the average velocity  $v$  remains constant. Since the relative velocity of two particles is  $\sqrt{2}v$ , the average frequency of collision  $C$ , or the average number of contacts per unit of time can be represented as:  $C = \sqrt{2}\pi d^2 \rho v$ , where  $d$  is the collision diameter and the cross-sectional area is  $\pi d^2$ . In this formulation, when density is low, the contact rate  $C$  has linear dependence with density  $\rho$ .

The kinetic model assumes a random movement with collision between particles (or individuals), rather than population movements with clustering or social organizations. This assumption does not apply when modeling human societies in a large spatial scale, since individuals form social groups and the daily trajectories are repetitive and highly predictable [39]. If physical contact behaviors were random, the kinetic model would be more applicable, for example, short and random contacts between people at public sites. However, corrections to the model needs to be made under special conditions, e.g. when population are overcrowded in limited spaces. We explore these limits and corrections in more details, and will discuss them in the next section.

#### 2.4. The contact rate scaling of crowds

Local public sites with extremely high population density, such as train stations or large social, political, or religious mass gatherings are regarded as high-risk for respiratory disease transmission due to extended close contact between hosts [41]. The Islamic pilgrimage Hajj brings more than 2 million pilgrims a year to Mecca, Saudi Arabia, where densities reach about 7 people/m<sup>2</sup> walking around the Kaaba. This raises major concerns about the risk of respiratory diseases, like the 2009 H1N1 influenza outbreak [42]. Enhanced disease surveillance and public health control measures have been performed at previous major events [43–46], but the general epidemiology at mass gatherings remains poorly understood [47]. Mathematical modeling have potential applicability in assessing effect of public health policies and assist policy planning and strategic decision making [25]. Simulating these dynamics requires better understanding of contact patterns and the appropriate adjustment of transmission rates.

The laminar movement of pedestrians can break down under high densities, resulting in self-collective behaviors, such as stop-and-go waves and turbulence [48,49]. Existing studies use a variety of experiments and models, such as Newtonian physics [50,51] or cognitive heuristics [48], to study the patterns of pedestrian flows and crowd dynamics. In the disease modeling context, we are more interested in calculating the contact rate of each individual. Sophisticated mathematical expressions require large numbers of inputs for calibration [48]. Therefore, for simplicity we assume that people are more likely to move and contact others randomly.

As people are only able to contact others nearby, the activity range  $r_{\text{max}}$ , or the width of the spatial kernel, is close to 0, therefore contact behavior leans towards a kinetic gas model, and uniform mixing assumption is no longer appropriate for modeling crowded areas. However, when  $\rho$  is extremely high, the assumption will be different from the ideal gas kinetic model discussed in Section 2.3. Here, average speed of people is expected to be slowed down under high densities, representing the impeded pedestrians in densely packed crowds. Empirical measurements of crowd motion follow Maxwell–Boltzmann's theory [52,24]. We propose that the short, frequent pedestrian contact process is analogous to the collision process using the crowd gaseous model. The relationship between average speed and population density is obtained in this case. Besides, the van der Waals correction may need to be considered, because the volumes of molecules and inter-molecule interactions are not negligible any more. In the model, we consider individuals as molecules in a restricted container with volume  $V$ .



Similar physical properties of gas/liquids can be used with different meanings: temperature  $T$  reflects the energy mode or speed of individuals, and pressure  $P$  represents the pressure of individuals going out of the system.

Therefore, if the molecular weight is defined as a constant  $M$ , the equation of state can be written as:

$$\left(P + \frac{\rho^2 a}{M^2}\right) \left(\frac{M}{\rho} - b\right) = RT$$

$P$  is assumed to be constant (i.e., the overall momentum for individuals going out of the system does not change). Using the root mean square speed  $v = \sqrt{\frac{3RT}{M}}$  and merging all constants, we have the following relationship between  $v$  and  $\rho$ :

$$v = c_1 \sqrt{\frac{1 + c_2 \rho + c_3 \rho^2 + c_4 \rho^3}{\rho}}$$

For ideal gas, using  $P = \rho RT$ , the speed is  $v = c_1 \sqrt{\frac{1}{\rho}}$ .

Comparing between the two models, at the numerator, the first order of density is due to the size of molecules. The second and third orders of density are introduced by the interaction potential between people. In contrast to gas molecules, the inter-person potential is small enough to be ignored, even at the large density limit. Therefore, the following equation describes the quantitative relationship between people's movement and population density:

$$v = c_1 \sqrt{\frac{1 + c_2 \rho}{\rho}}$$

Note that the density needs to be smaller than  $-\frac{1}{c_2}$ , the hard limit when all molecules fill in the space.

In densely-packed crowds, the average speed of people is slower and transmission potential is greatly reduced due to the decreasing frequency of contacts. Furthermore, the spread of disease is diffusive since people only contact nearby individuals and there are no “small-world” links between contacts. Chemical reactions between molecules provide a good analogy to infectious disease transmission. The transmission rate is similar to the reaction rate, depending on collision frequency.

According to results from kinetic theory, the contact rate can be calculated as follows:

$$C = \sqrt{2} \pi d^2 \rho v = \sqrt{2} \pi d^2 c_1 \sqrt{\rho(1 + c_2 \rho)}$$

### 3. Empirical examples

#### 3.1. Human diseases

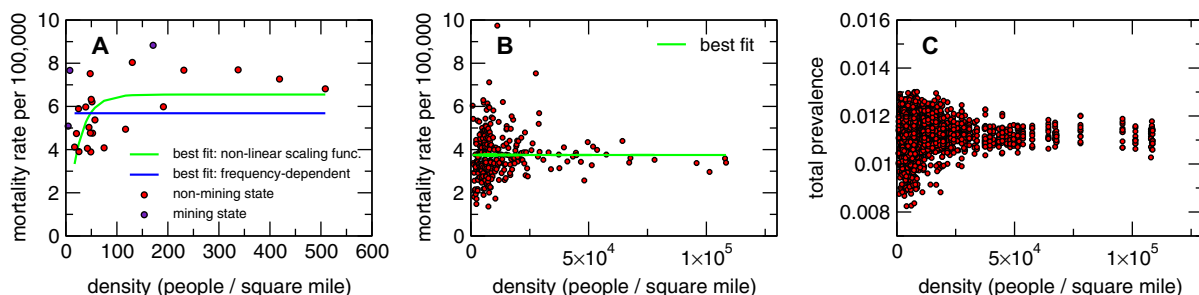
To observe if real disease outbreaks have similar scaling with population density, we assess the transmissibility of the 1918

influenza pandemic in the United States, England, and Wales. The 1918 influenza pandemic is one of the more recent, well-documented global pandemics with detailed records on cities and towns. For United States, state-level cumulative mortality rates for influenza and pneumonia, population, and area information are obtained from historical records [53] and 1920 Mortality Statistics [54]. Additional state reports on agriculture, industry, and mining come from the 1910 U.S. Census [55]. For England and Wales, large town and city-level influenza mortality data (population > 20,000) have been recorded by the UK government and made available to the public [56,57]. Demographic information for population and area is from the 1911 National Census [58]. We use mortality records during the second and third peaks, which are the two largest waves.

Invasion of an infectious disease pathogen into a local population is a two-stage process: initial importation from migration between local populations, and subsequent increases in the number of infected individuals due to contacts within the local population. After the initial seeding, secondary spread of disease is driven mainly by the force of infection among local contacts. For the 1918 influenza pandemic, we investigate the local infection stage by looking at the cumulative influenza (or influenza and pneumonia for the United States) mortality rate for each geographic location. Mortality is used as a surrogate for the final size of the outbreak, which scales with the average reproductive number. We do not calculate the exact reproductive number, but look for scaling of transmissibility due to population effects. We assume that the probability of infection, given contact with an infected case, and the duration of the infectious period are the same for all individuals, and the contact rate is only related to population density. Under these assumptions, the reported mortality value directly scales with contact rate.

Population density is calculated using the total area of the investigated region, therefore each region is expected to have uniform population density with people living in the majority of areas. For this reason, influenza mortalities in U.S. states with more than 5% of the population in mining occupations are reported separately and not included in parameter fitting. In mining states, the population is likely to form aggregated clusters rather than be spread uniformly across the state, making the actual population densities far from the estimated values. This effect is less obvious in non-mining states.

Fig. 4(A) shows the relationship between state-wide influenza mortality rates and population densities. Generally, the low-density states have lower reported mortality rates, while the rate is close to constant for higher density states. Our estimates from the spatial kernel fit the overall trend better than the frequency-dependent approach: for non-mining states, the best parameters are  $c_1 = 6.55$ ,  $c_2 = 0.04$  ( $\chi^2 = 29.6$ ), compared with the frequency-dependent fit,  $c = 5.68$  ( $\chi^2 = 39.3$ ). Of note, the mining states have a



**Fig. 4.** The scaling between contact rate and density in general activities. (A) The relationship between population density and 1918 influenza and pneumonia mortality for the United States, grouped by mining and non-mining states, with the best fit for frequency dependent and the nonlinear scaling function for non-mining states only. Mortality and demographic data are obtained from [53]. (B) The relationship between population density and 1918 influenza mortality for cities and towns in England and Wales. (C) Stochastic variation of total prevalence produced by the SEIR simulation, using the same size and density for cities and towns in England and Wales.

higher mortality rate than non-mining states with comparable density, indicating the possible effects of overcrowding and highly heterogeneous, scattered population clusters.

For England and Wales, the mortality across cities and towns appear to be frequency-dependent, as the best fit of the nonlinear scaling function gave a form that was already saturated in the density range. This outcome is in accordance with findings from prior studies of the England and Wales data [59], as well as those of large cities in the United States [60]. A prior report on the reproductive number of measles in large England and Wales cities also suggests frequency-dependent transmission, which is independent of population size [61]. Cities and towns in England and Wales have much larger population densities when compared to states in the U.S, and it might explain why the cumulative influenza mortality rate is mostly independent of population density (Fig. 4(B)). The 1918 England and Wales influenza records lack original data for small towns and rural areas. The characteristic density constant  $\rho_0$  for influenza and measles may be low: cities and towns with populations under 20,000 typically have densities of less than 500 people per square mile (the saturation region for the nonlinear scaling function obtained from the United States). Therefore, it is possible that existing data only show the saturation of transmissibility in larger towns and do not speak to the dynamics in lower density rural areas. Better data would be required for validation of the nonlinear function in the England and Wales context.

Large fluctuations at low population densities are likely due to stochastic fluctuations and the small number statistics. To illustrate this effect, we perform a agent-based, discrete, stochastic SEIR simulation with  $R_0 = 2.1$ , using the fitted contact rate

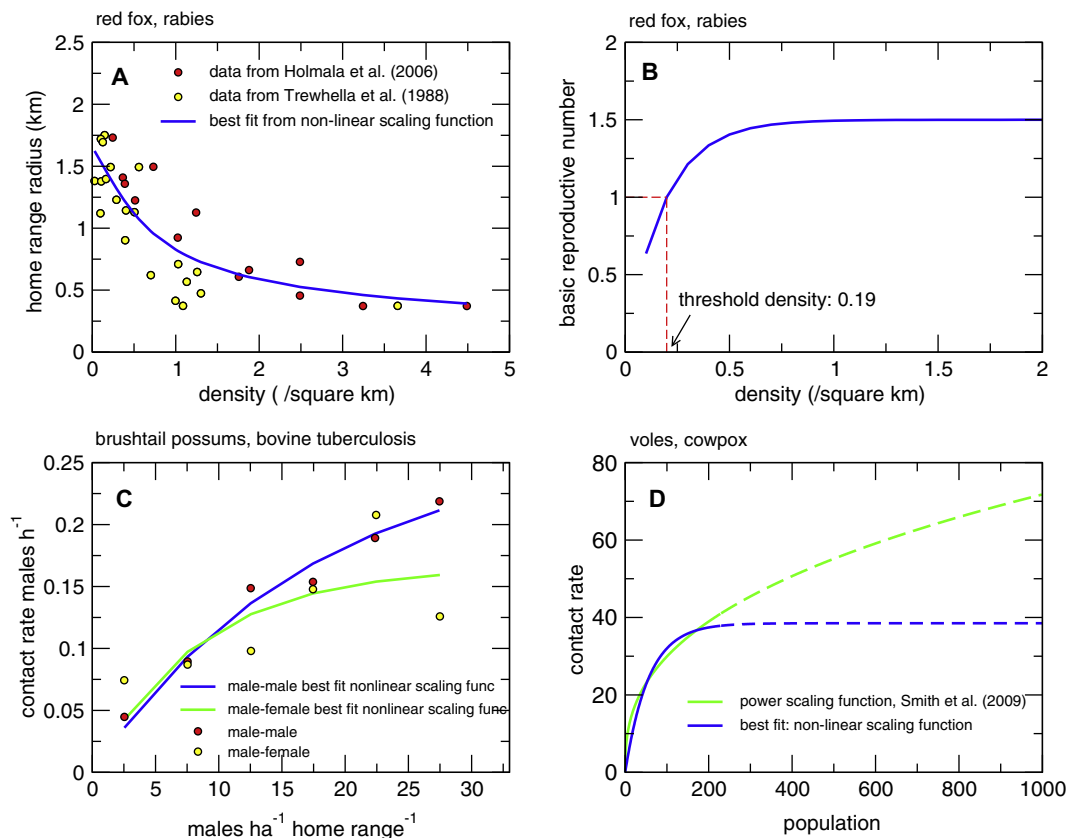
functions as shown above. For direct comparison, we use the same population sizes and densities as the England and Wales cities data in 1911. Ten stochastic simulations were performed for each population size and density. In Fig. 4(C), the results are presented, and they show a similar trend in incidence variance, which decrease with population size.

### 3.2. Wildlife diseases

For small populations, scarcity of data makes it hard to compare outcomes against other human diseases. Therefore, we apply our scaling function to certain wildlife diseases where data is relatively abundant.

Rabies epizootic spreads steadily across Europe in the last century, mostly among red foxes (*Vulpes vulpes*). Various field studies have been carried out to investigate its spatial-temporal spread [62]. There have been substantive efforts in mathematical modeling to study rabies virus transmission dynamics, in order to provide insights on spatial invasion as well as effects of interventions, such as culling or vaccination. Based on early observations of the threshold density – below which the endemic fails to sustain, most models assume a density-dependent transmission rate.

From multiple empirical studies, a negative correlation exists between fox density and home range size. Fig. 5(A) demonstrates the combined results of the two studies in [62,63]. This evidence suggests that contact rates among animals are largely – but not entirely – determined by population density [62]. When fox density is high, the contact model may be different from linear



**Fig. 5.** Density scaling function in wildlife diseases. (A) empirical data between fox home range and density relationship, and the best fit values using the spatial scaling function. Home range radius (km) are calculated from home range (hectare) assuming the area is circular. (B) the scaling of basic reproductive number and the threshold density, using best fit values in (A). (C) fitting the scaling function to mating contact rates among brushtail possums. (D) The comparison between the scaling function  $C \sim N^{\alpha}$  ( $\alpha = 0.38$ ) obtained for voles, for transmission of cowpox in [30]. The best fit for the non-linear scaling function in the same data range. The dashed lines show the extrapolated curve when population is large.

density-dependent behavior and transmission rate does not increase linearly when density is large.

The relationship between fox density and home range radius is in line with the spatial contact kernel. Therefore we fit the kernel using reported field values of home range and density from both studies and derive the density scaling function. We assume home range distance for an individual is the radius where 50% of contacts happen. The best estimates for the two constants:  $\rho_0$  and  $r_{max}$  are acquired. Fig. 5(A) shows the best fit of the spatial contact kernel ( $\rho_0 = 0.99$  and  $r_{max} = 2.34$ ), and Fig. 5(B) demonstrates the shape of non-linear scaling function which saturates at  $R_0 = 1.5$ .

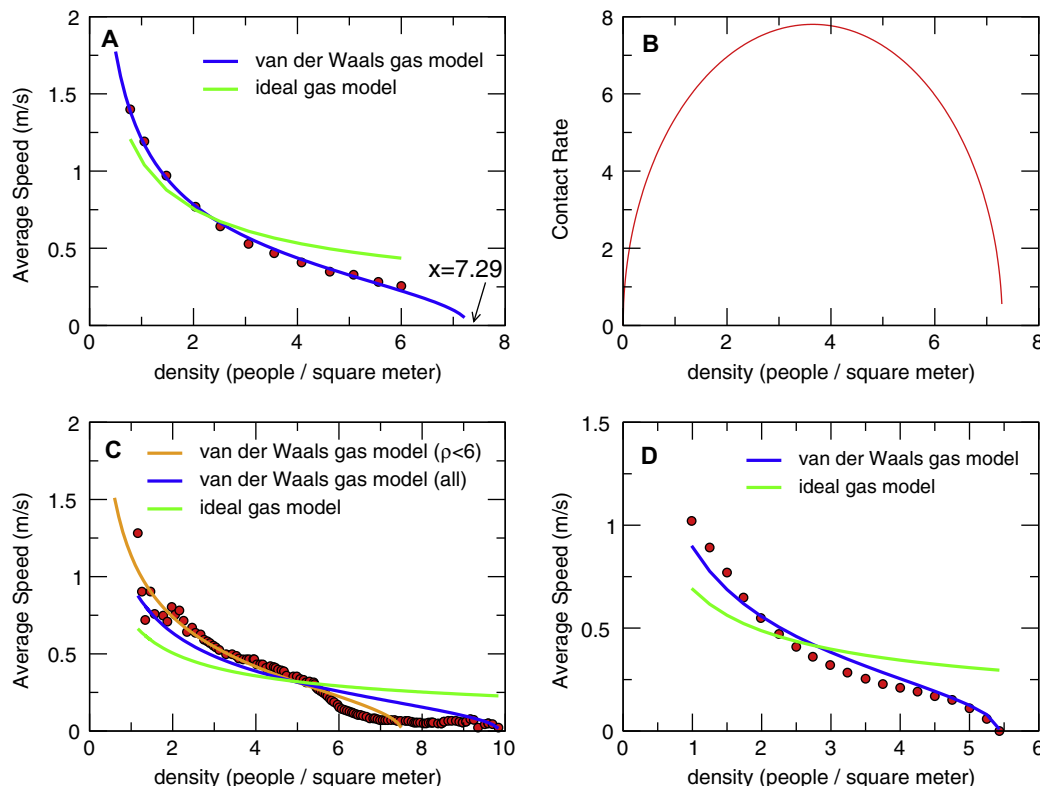
To test the validity of the obtained function, we calculate the threshold density of rabies in Europe. If the saturated  $R_0$  is assumed as 1.5, Fig. 5(B) shows the scaling function and the corresponding threshold density, which is  $0.19/\text{km}^2$ . This threshold density is consistent with estimated value of 0.2 from [64], and is close to field observations from Europe [62]: 0.63 in central Europe [65], as low as 0.25 [66], 0.3 in southern parts of western Siberia [67]. Compared with the linear relationship, because the non-linear scaling function is convex-up, under the threshold density,  $R_0$  value obtained from the non-linear scaling function decays slower than the one obtained from the density-dependent scaling function. This confirms the finding that comparing with vaccination, culling might not be effective in stopping the endemic transmission and achieve local eradication. In fact, field experiences from Europe suggest that the use of oral rabies vaccination is much more effective and there are only very few examples where culling alone have either eliminated the disease or have prevented its spread to previously uninfected areas [68]. The spatial contact kernel explains this nicely: as density decreases with culling, the width of each remaining individual's spatial contact kernel increases. This maintains a consistent level of contacts until the resulting spatial kernel width begins to increase beyond the mobility limiting  $r_{max}$ .

Spatial scaling function is also fitted to contact rates between *brush-tail possums* when looking at bovine tuberculosis transmission [69]. The best fits ( $c_0 = 0.26$ ,  $c_1 = 0.056$  for male–male contact and  $c_0 = 0.165$ ,  $c_1 = 0.116$  for male–female contact, see Fig. 5(C)) successfully replicates the non-linear convex-up relationship from the obtained field data, which implies the contact rate during the breeding season does not decrease in proportion to reductions in density [69]. The authors suggest this phenomenon is probably due to the increased male overlap of female ranges following the density reduction, which aligns well with the assumptions of the spatial kernel.

Another example involves the transmissibility of cowpox in wildlife populations. In [30] the authors report a power function  $C \sim N^\alpha$  ( $\alpha = 0.38$ ) that represents the scaling between contact rates and population. We compare this function with the best fits using our density scaling function in Fig. 5(D) ( $c_0 = 38.43$ ,  $c_1 = 0.018$ ). Although both functions provide similar scaling results when the population is small, our non-linear scaling function gives potentially more realistic results for the saturation of infectivity when extrapolating to large populations.

### 3.3. Contact rates of crowds

Using existing pedestrian movement study, the relationship between average speed of people and population density is fitted into the state equation. Fig. 6(A) demonstrates the best fit with the empirical passenger walking speed and crowd density data obtained from a Japanese railway station [70], with  $c_1 = 1.30$ ,  $c_2 = -0.137$ ,  $\chi^2 = 0.005$ . As a comparison, we also report the fitting results for the ideal gas model  $= c_1 \sqrt{\frac{1}{\rho}}$ :  $c_1 = 1.07$ ,  $\chi^2 = 0.205$ . The van der Waals state equation model fits the data more closely. In addition, the maximum possible density



**Fig. 6.** Fitting the van der Waals state equation to population movement data. (A) Best fit between the pedestrian speed and population density, data were obtained from a Japanese train station, reported in [70]. The ideal gas model fitting result is present as well. (B) The estimation of collision frequency based on population density, using the fitted van der Waals state equation and value for close-contact distance. (C) and (D) Fitting the van der Waals gas model to two other empirical crowd movement studies: (C) Hajj and (D) movements on a plane, respectively [49,71], and comparing the results to the best fits obtained for the ideal gas model.

**Table 1**

Best fit values for several empirical crowd movement studies, using van der Waals and ideal gas state equations.

Study	Van der Waals model, $v = c_1 \sqrt{\frac{1+c_2\rho}{\rho}}$	Ideal gas, $v = c_1 \sqrt{\frac{1}{\rho}}$
Train station [70]	$c_1 = 1.30, c_2 = -0.137(\chi^2 = 0.005)$	$c_1 = 1.07(\chi^2 = 0.205)$
Hajj [49]	$c_1 = 1.01, c_2 = -0.101(\chi^2 = 1.005)$	$c_1 = 0.714(\chi^2 = 3.492)$
Hajj [49] ( $\rho < 6$ )	$c_1 = 1.22, c_2 = -0.133(\chi^2 = 0.199)$	$c_1 = 0.915(\chi^2 = 0.727)$
Plane [71]	$c_1 = 0.99, c_2 = -0.184(\chi^2 = 0.063)$	$c_1 = 0.689(\chi^2 = 0.337)$

for the van der Waals state equation ends at  $7.3 / \text{m}^2$ , close to the reported maximum density value  $7.4 / \text{m}^2$  in the literature [70]. In comparison, the ideal gas model does not have density cut-offs.

Data from other movement studies in a variety of settings reveal a similar pattern. Results are reported in Table 1. Observed data from the Hajj has population densities up to  $10 / \text{m}^2$  [49]. Since reaching this density limit seem to require collective human movement patterns as opposed to random movement, we perform a separate fit considering only data points from  $\rho < 6 / \text{m}^2$ . For all scenarios, the decreasing rates and cutoff values are similar, but differ slightly due to the “pressure” related to the specific environment. The van der Waals ideal gas model is able to fit the empirical data well.

We calculate the contact rate  $C = \sqrt{2}\pi d^2 c_1 \sqrt{\rho(1+c_2\rho)}$  using results from the Japanese railway station study, assuming a collision diameter  $d = 1$  m, based on the WHO definition of a close contact as interaction within 1 m [72]. Fig. 6(B) demonstrates the density dependence of collision frequency. The contact rate keeps increasing until it reaches maximum, and then decreases due to the slower individual movement. It finally reaches zero when density is at its maximum.

#### 4. Discussion

Recent computational modeling studies use network models or social structures to represent contact heterogeneity when modeling the spread of infectious disease outbreaks, and empirical contact experiments focus on revealing the connectivity and the heterogeneous structure of the contact network. However, the scaling mechanisms are less focused. The heterogeneous contact structure seems ‘orthogonal’ to the transmission scaling mechanisms [20]. Using networks, an existing study was able to obtain a consistent scaling relationship regardless of heterogeneous network connectivity [21]. Exploring population scaling on both dimensions is essential for a full understanding of human mixing patterns.

Comparing with other contact rate functions, for example the various forms in [19], the advantage of our scaling function in both density ends is that it is constructed *ab initio* from physical contacts, it seems to follow the empirical data in both human and wildlife diseases and it is possible to formulate and parameterize a reasonable scaling relationship from independent field observations and get meaningful scaling results. For direct comparison of real data with other models, different models have different free parameters, and in many cases models are flexible enough that they will fit the data quite well. It is only easy to reject a model when it makes prediction far from reality, for example the infinite linear relationship of the density-dependent function.

Our nonlinear scaling function has several limitations. We assume an exclusive scaling relationship between reproductive number and population density. Other factors, such as age, population immunity in previously isolated communities, and seasonal variations may contribute to differences in transmissibility. The effect of important contact structures, such as household and living conditions, is not considered either. For the transmissibility of the 1918 influenza pandemic, the population data is acquired from the closest census records, and we assume those numbers do not

change significantly during the epidemic period. The density is calculated using total area rather than actual living area, which may have biased the results. In the United States, we exclude several mining states likely to have highly heterogeneous distributions of population with spatial clusters. This aligns with the observation that their dynamics more closely resemble those of more densely-populated states, rather than similarly sparsely-populated agricultural states. It is possible that other states with similar heterogeneous clusters may not be fully excluded.

The nonlinear scaling function applies mainly to diseases with direct transmission, where the density is one-dimensional. For vector-borne diseases, vector density and human vector ratio are important, and these factors complicate the scaling relationships. For instance, historical records for the flea-borne plague (Black Death) show that mortality rates in Italy and France were highest in sparsely-populated areas in the countryside rather than in cities and towns. Even the rates in the largest cities, with more than 100,000 people, weren’t comparable [73]. The pathogen was transmitted by rat fleas to human beings. In the countryside, a rat colony cohabited with a household, while in urban areas several households were crowded together within the territory of a single rat colony. Therefore, the rat/human ratio was likely to be lower in urban areas than rural ones, causing higher mortality in the countryside [73]. In rat-infested urban areas, natural selection might have also lowered the rat/human ratio by weeding out rats less able to defend themselves against the pathogen.

At the high density limit, the use of collision model is constrained by its initial assumption that people move randomly. Therefore, the above scaling trend is closer to reality when people move in random directions, such as in railway stations. When people move collectively, as in the counterclockwise motion around the Kaaba at the Hajj, the scaling trend will deviate from the behavior of the estimated curve since the movement of individuals differs from the random movement of gas molecules. Besides, although detailed logs between population movement and density fit well into the kinetic model, the contact duration, which is essential to determine the transmission rate, remains unexplored. In mass gatherings most contacts may occur in a very short time, therefore the overall transmissibility of respiratory disease through crowds may not be beyond the  $R_0 = 1$  threshold, and it is reasonable to argue that transmission of influenza in crowds where frequent but short contact occurs may not be as important as household and workplace contacts, with low frequency but long durations. Nonetheless, the kinetic model provides a potential approach to model physical contacts under large density settings which could be used in modeling transmission of respiratory diseases. More empirical observations and investigations on infectious diseases spread in mass gathering are needed to prove the efficiency of direct contact transmission in a variety of density settings.

#### 5. Conclusion

Substantial evidence indicates that when population density is at the scale of general activities, the transmission mode is likely to follow an initial sub-linear density-dependent pattern until the saturation of transmission rate transitions to a frequency-dependent pattern independent of population density. For high-density clusters, such as crowds at mass gatherings, the random movement and contact of individuals initially increases the frequency of contacts. However, the contact rate decreases at extremely high density, when it becomes difficult for people to move and make contact with others. On both ends, the scaling effects can be represented using simple functional forms from spatial contact models. Existing data, though scarce, support the proposed models. We suggest that the conclusion of contact rate scaling with



population density may provide a first-order approximation inspired by mechanisms of human social interaction and wildlife ecology, when setting up transmission rates in large-scale spatial simulations. We hope that more experimental and theoretical social contact studies will reveal the complete picture of the mixing patterns between individuals.

## Acknowledgments

The authors thank Daniel J. Klein and two anonymous reviewers for their insightful discussions and suggestions. The 1918 influenza data was based on data UKDA 4350: Influenza deaths in England and Wales, 22 June 1918–10 May 1919, deposited by N. Johnson and distributed by the UK Data Archive. The data are Crown copyright. The original data creators, depositors or copyright holders, the funders of the Data Collections (if different), and the UK Data Archive bear no responsibility for the further analysis or interpretation in this publication. The authors thank Bill and Melinda Gates for their active support of the EMOD Project and their sponsorship through the Global Good Fund. This work was performed in the EMOD Program at Intellectual Ventures Laboratory, and useful discussions with colleagues in this Program are likewise greatly appreciated.

## References

- [1] I. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaworakul, D. Cummings, M. Halloran, Containing pandemic influenza at the source, *Science* 309 (2005) 1083.
- [2] T. Germann, K. Kadau, I. Longini, C. Macken, Mitigation strategies for pandemic influenza in the United States, *Proceedings of the National Academy of Sciences* 103 (2006) 5935.
- [3] N. Ferguson, D. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsrithaworn, D. Burke, Strategies for containing an emerging influenza pandemic in southeast Asia, *Nature* 437 (2005) 209.
- [4] N. Ferguson, D. Cummings, C. Fraser, J. Cajka, P. Cooley, D. Burke, Strategies for mitigating an influenza pandemic, *Nature* 442 (2006) 448.
- [5] S. Eubank, H. Guclu, V. Anil Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang, Modelling disease outbreaks in realistic urban social networks, *Nature* 429 (2004) 180.
- [6] S. Cauchemez, A. Bhattarai, T. Marchbanks, R. Fagan, S. Ostroff, N. Ferguson, D. Swardlow, S. Sodha, M. Moll, F. Angulo, Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza, *Proceedings of the National Academy of Sciences* 108 (2011) 2825.
- [7] S. Riley, Large-scale spatial-transmission models of infectious disease, *Science* 316 (2007) 1298.
- [8] R. May, Network structure and the biology of populations, *Trends in Ecology & Evolution* 21 (2006) 394.
- [9] D. Chao, M. Halloran, V. Obenchain, I. Longini Jr, Flute, a publicly available stochastic influenza epidemic simulation model, *PLoS Computational Biology* 6 (2010) e1000656.
- [10] S. Brown, J. Tai, R. Bailey, P. Cooley, W. Wheaton, M. Potter, R. Voorhees, M. Lejeune, J. Grefenstette, D. Burke, S. McGlone, B. Lee, Would school closure for the 2009 H1N1 influenza epidemic have been worth the cost? A computational simulation of Pennsylvania, *BMC Public Health* 11 (2011) 353.
- [11] B. Lee, S. Brown, P. Cooley, M. Potter, W. Wheaton, R. Voorhees, S. Stebbins, J. Grefenstette, S. Zimmer, R. Zimmerman, Simulating school closure strategies to mitigate an influenza epidemic, *Journal of Public Health Management and Practice: JPHMP* 16 (2010) 252.
- [12] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. Tomba, J. Wallinga, Social contacts and mixing patterns relevant to the spread of infectious diseases, *PLoS Medicine* 5 (2008) e74.
- [13] D. Balcan, B. Gonçalves, H. Hu, J. Ramasco, V. Colizza, A. Vespignani, Modeling the spatial spread of infectious diseases: the global epidemic and mobility computational model, *Journal of Computational Science* 1 (2010) 132.
- [14] L. Isella, M. Romano, A. Barrat, C. Cattuto, V. Colizza, W. Van den Broeck, F. Gesualdo, E. Pandolfi, L. Rav, C. Rizzo, Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors, *PLoS ONE* 6 (2011) e17144.
- [15] J. Stehle, N. Voirin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Regis, J. Pinton, N. Khanafer, W. Van den Broeck, Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees, *BMC Medicine* 9 (2011) 87.
- [16] A. Deredec, F. Courchamp, Extinction thresholds in host-parasite dynamics, *Annales Zoologici Fennici* 40 (2003) 115.
- [17] J. Lloyd-Smith, P. Cross, C. Briggs, M. Daugherty, W. Getz, J. Latto, M. Sanchez, A. Smith, A. Swei, Should we expect population thresholds for wildlife disease?, *Trends in Ecology & Evolution* 20 (2005) 511.
- [18] M. Keeling, B. Grenfell, Disease extinction and community size: modeling the persistence of measles, *Science* 275 (1997) 65.
- [19] H. McCallum, N. Barlow, J. Hone, How should pathogen transmission be modelled?, *Trends in Ecology & Evolution* 16 (2001) 295.
- [20] M. Begon, M. Bennett, R. Bowers, N. French, S. Hazel, J. Turner, A clarification of transmission terms in host-microparasite models: numbers, densities and areas, *Epidemiology and Infection* 129 (2002) 147.
- [21] M. Ferrari, S. Perkins, L. Pomeroy, O. Bjornstad, Pathogens, social networks, and the paradox of transmission scaling, *Interdisciplinary Perspectives on Infectious Diseases* 2011 (2011) 1.
- [22] M. de Jong, O. Diekmann, H. Heesterbeek, How does transmission of infection depend on population size, *Epidemic Models: Their Structure and Relation to Data* (1995) 84.
- [23] R. Anderson, R. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.
- [24] C. Rhodes, R. Anderson, Contact rate calculation for a basic epidemic model, *Mathematical Biosciences* 216 (2008) 56.
- [25] J. Tam, M. Barbeschi, N. Shapovalova, S. Briand, Z. Memish, M. Kieny, Research agenda for mass gatherings: a call to action, *The Lancet Infectious Diseases* 12 (2012) 231.
- [26] K. Dietz, D. Schenzle, Proportionate mixing models for age-dependent infection transmission, *Journal of Mathematical Biology* 22 (1985) 117.
- [27] K. Dietz, Overall population patterns in the transmission cycle of infectious disease agents, *Population Biology of Infectious Diseases* (1982) 87.
- [28] D. Schenzle, K. Dietz, Critical population sizes for endemic virus transmission, *Rumliche Persistenz und Diffusion von Krankheiten* 83 (1987) 31.
- [29] A. Fenton, J. Fairbairn, R. Norman, P. Hudson, Parasite transmission: reconciling theory and reality, *Journal of Animal Ecology* 71 (2002) 893.
- [30] M.J. Smith, S. Telfer, E.R. Kallio, S. Burthe, A.R. Cook, X. Lambin, M. Begon, Host-pathogen time series data in wildlife support a transmission function between density and frequency dependence, *Proceedings of the National Academy of Sciences* 106 (2009) 7905.
- [31] C. Holling, The components of predation as revealed by a study of small-mammal predation of the European pine sawfly, *The Canadian Entomologist* 91 (05) (1959) 293.
- [32] J. Antonovics, Y. Iwasa, M. Hassell, A generalized model of parasitoid, venereal, and vector-based transmission processes, *The American Naturalist* 145 (1995) 661.
- [33] M. Hochberg, Non-linear transmission rates and the dynamics of infectious disease, *Journal of Theoretical Biology* 153 (1991) 301.
- [34] M. Keeling, M. Woolhouse, D. Shaw, L. Matthews, M. Chase-Topping, D. Haydon, S. Cornell, J. Kappey, J. Wilesmith, B. Grenfell, Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape, *Science* 294 (2001) 813.
- [35] R. Larsson, K. Nigmatulina, Engineering responses to pandemics, in: W. Rouse, D. Cortese (Eds.), *Engineering the System of Healthcare Delivery*, IOS Press, Washington, DC, 2009, p. 311.
- [36] K. Nigmatulina, R. Larson, Living with influenza: impacts of government imposed and voluntarily selected interventions, *European Journal of Operational Research* 195 (2009) 613.
- [37] D. Brockmann, L. Hufnagel, T. Geisel, The scaling laws of human travel, *Nature* 439 (2006) 462.
- [38] M. Gonzalez, C. Hidalgo, A. Barabasi, Understanding individual human mobility patterns, *Nature* 453 (2008) 779.
- [39] C. Song, Z. Qu, N. Blumm, A. Barabasi, Limits of predictability in human mobility, *Science* 327 (2010) 1018.
- [40] D. Klein, J. Hespanha, U. Madhwa, A reaction-diffusion model for epidemic routing in sparsely connected MANETs, in: *Proceedings IEEE INFOCOM 2010*, IEEE, 2010, p. 1.
- [41] I. Abubakar, P. Gautret, G. Brunette, L. Blumberg, D. Johnson, G. Pomeroy, Z. Memish, M. Barbeschi, A. Khan, Global perspectives for prevention of infectious diseases associated with mass gatherings, *The Lancet Infectious Diseases* 12 (2012) 66.
- [42] S. Ebrahim, Z. Memish, T. Uyeki, T. Khoja, N. Marano, S. McNabb, Pandemic H1N1 and the 2009 Hajj, *Science* 326 (2009) 938.
- [43] L. Jorm, S. Thackway, T. Churches, M. Hills, Watching the games: public health surveillance for the Sydney 2000 Olympic Games, *Journal of Epidemiology and Community Health* 57 (2003) 102.
- [44] V. Demicheli, R. Raso, D. Tiberti, A. Barale, L. Ferrara, D. Lombardi, Epidemiological Consultation Team, Surveillance system in place for the 2006 Winter Olympic Games, Torino, Italy, 2006, *Euro Surveillance* 11 (2006) E060209.
- [45] J. Dvorak, A. Junge, K. Grimm, D. Kirkendall, Medical report from the 2006 FIFA World Cup Germany, *British Journal of Sports Medicine* 41 (2007) 578.
- [46] K. Schenkel, C. Williams, T. Eckmanns, G. Poggensee, J. Benzler, J. Josephsen, G. Krause, Enhanced surveillance of infectious diseases: the, 2006 FIFA World Cup experience, Germany, *Euro Surveillance* 11 (2006) 234.
- [47] H. Rashid, E. Haworth, S. Shafi, Z. Memish, R. Booy, Pandemic influenza: mass gatherings and mass infection, *Ambulatory Pediatrics* 8 (2008) 526.
- [48] M. Moussaid, D. Helbing, G. Theraulaz, How simple rules determine pedestrian behavior and crowd disasters, *Proceedings of the National Academy of Sciences* 108 (2011) 6884.
- [49] D. Helbing, A. Johansson, H. Al-Abideen, Dynamics of crowd disasters: an empirical study, *Physical Review E* 75 (2007) 046109.
- [50] D. Helbing, I. Farkas, T. Vicsek, Simulating dynamical features of escape panic, *Nature* 407 (2000) 87.

- [51] M. Moussaid, D. Helbing, S. Garnier, A. Johansson, M. Combe, G. Theraulaz, Experimental study of the behavioural mechanisms underlying self-organization in human crowds, *Proceedings of the Royal Society B: Biological Sciences* 276 (2009) 2755.
- [52] L. Henderson, The statistics of crowd fluids, *Nature* 229 (1971) 381.
- [53] T. Garrett, Pandemic economics: the 1918 influenza and its modern-day implications, *Federal Reserve Bank of St. Louis Review* 90 (2008) 75.
- [54] U.S. Census Bureau, Mortality Statistics 1920, Twenty-First Annual Report, Government Printing Office, Washington, 1922.
- [55] U.S. Census Bureau, Census of population and housing, 1910.
- [56] HM Stationery Office, Supplement to the eighty first annual report of the register general: report of the mortality from influenza in England and Wales during the epidemic 1918–1919, 1920.
- [57] N. Johnson, 1918–1919 influenza pandemic mortality in England and Wales, 2001.
- [58] Great Britain Historical GIS Project, A vision of Britain through time, 2011.
- [59] G. Chowell, L. Bettencourt, N. Johnson, W. Alonso, C. Viboud, The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact, *Proceedings of the Royal Society B: Biological Sciences* 275 (2008) 501.
- [60] C. Mills, J. Robins, M. Lipsitch, Transmissibility of 1918 pandemic influenza, *Nature* 432 (2004) 904.
- [61] O. Bjornstad, B. Finkenstadt, B. Grenfell, Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model, *Ecological Society of America* 72 (2002) 169.
- [62] K. Holmala, K. Kauhala, Ecology of wildlife rabies in Europe, *Mammal Review* 36 (2006) 17.
- [63] P. White, S. Harris, G. Smith, Fox contact behaviour and rabies spread: a model for the estimation of contact probabilities between urban foxes at different population densities and its implications for rabies control in Britain, *Journal of Applied Ecology* (1995) 693.
- [64] G. Smith, D. Wilkinson, Modeling control of rabies outbreaks in red fox populations to evaluate culling, vaccination, and vaccination combined with fertility control, *Journal of Wildlife Diseases* 39 (2003) 278.
- [65] J. David, L. Andral, M. Artois, Computer simulation model of the epi-enzootic disease of vulpine rabies, *Ecological Modelling* 15 (1982) 107.
- [66] B. Toma, L. Andral, Epidemiology of fox rabies, *Advances in Virus Research* 21 (1977) 1.
- [67] Z. Tyul'ko, I. Kuzmin, Simulation of rabies epizootic process in fox populations at a limited carrying capacity of biotopes, *Russian Journal of Ecology* 33 (2002) 331.
- [68] W.H. Organization, WHO expert consultation on rabies: first report, World Health, Organization, 2005.
- [69] D. Ramsey, N. Spencer, P. Caley, M. Efford, K. Hansen, M. Lam, D. Cooper, The effects of reducing population density on contact rates between brushtail possums: implications for transmission of bovine tuberculosis, *Journal of Applied Ecology* 39 (2002) 806.
- [70] R. Smith, Density, velocity and flow relationships for closely packed crowds, *Safety Science* 18 (1995) 321.
- [71] U. Weidmann, Transporttechnik der Fuganger, Transporttechnische Eigenschaften des Fugangerverkehrs, Literaturauswertung, Schriftenreihe des IV, ETH Zurich, 1993.
- [72] WHO, Advice on the use of masks in the community setting in Influenza A(H1N1) outbreaks, 2009.
- [73] O. Benedictow, The Black Death 1346–1353: the Complete History, Boydell Press, 2004.