Spatial scales in human movement between reservoirs of infection

Robert J. Hardwick[a,b,c,*], Carolin Vegvari[a,c], Benjamin Collyer[a,c], James E. Truscott[a,b,c], Roy M. Anderson[a,b,c]

[a] *London Centre for Neglected Tropical Disease Research (LCNTDR), Department of Infectious Disease Epidemiology, St. Marys Campus, Imperial College London, London WC2 1PG, UK*
[b] *The DeWorm3 Project, The Natural History Museum of London, London, UK*
[c] *MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, UK*

Abstract

The life cycle of parasitic organisms that are the cause of much morbidity in humans often depend on reservoirs of infection for transmission into their hosts. Understanding the daily, monthly and yearly movement patterns of individuals between reservoirs is therefore of great importance to implementers of control policies seeking to eliminate various parasitic diseases as a public health problem. This is due to the fact that the underlying spatial extent of the reservoir of infection, which drives transmission, can be strongly affected by inputs from undefined external sources. In order to study the importance of these effects, we build and examine a novel theoretical model of human movement between spatially defined reservoirs of infection. Using our model we demonstrate the potential for the reservoir of infection to vary in spatial extent and temporal stability — effects which can both strongly influence the local transmission dynamics and response to control measures, e.g., mass drug administration (MDA). Considering the helminth parasites as our main example, by varying the spatial scale at which locations are defined, we demonstrate that a critical scale exists for an evaluation unit at which the migration rate into the associated reservoir of infection can be neglected for practical purposes. This scale varies by species and geographic region. Our model is designed to be applicable to a very general pattern of infectious disease spread induced by the migration of infected individuals between clustered communities. For example, it may be readily adapted to study the spatial structure of hosts for macroparasites such as the soil-transmitted helminths, schistosomes and filarial worms.

Keywords:
Mathematical models, Spatial infection model, Control policies, Monitoring and Evaluation

1. Introduction and background

Defining the spatial scales over which transmission should be considered is a relatively recent research area in the context of neglected tropical diseases (NTD), where much of the effort has been focussed on the development of geostatistical methods for each disease in turn [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]. In particular, mathematical models of helminth transmission have only recently begun to incorporate the dynamical effects of human movement between reservoirs of infection [9, 12, 13, 14].

Helminth infections (or helminthiases) are a class of macroparasitic diseases which include, among others, the soil-transmitted helminth (STH) infections, schistosomiasis and lymphatic filariasis (LF). Following World Health Organisation (WHO) guidelines, free drugs have been donated by pharmaceutical companies since 2010 [15, 16, 17, 18, 19] to countries significantly affected by helminth morbidity. Control initiatives have been developed specifically to investigate the prospects for elimination as a public health problem in the long term [20, 21].

In Fig. 1 we have illustrated the essential role that a reservoir of infection has on driving the transmission of human helminth infections. In the case of STHs, this reservoir of infection consistutes the eggs or larval stages in the soil that are either injested or enter the body via skin penetration. Similarly, in the case of schistosomiasis transmission, this reservoir exists as water sources in which larvae, that are released by freshwater snails, penetrate the skin to infect their host. The reservoir can also be mobile, e.g., in the case of lymphatic filariasis transmission, where the larval stage of microfilariae enter the body by bites from infected mosquitoes — hence, the reservoir of infection may itself migrate with the mosquito population.

It may be possible to minimise the effect of human movement has on evaluation units (EUs) of, e.g., mass drug administration (MDA) control programmes, by specifying a critical spatial scale of the reservoir (and corresponding human community) over which one may sample and treat infections. At this scale (or larger), the number of migrants contributing to the reservoir per unit time is potentially low enough such that their effect may essentially be neglected. Note also that if such a scale exists it would also provide a natural resolution level for those generating spatial maps of the prevalence of infections in a given region. In this paper we shall demonstrate the existence of this scale, which is expected to vary between helminths and regions, by developing a spatial model of human movement between defined locations which is consistent with the past literature.

The distances travelled by human walking and foraging patterns are not purely random, but exhibit characteristics similar

*r.hardwick@imperial.ac.uk

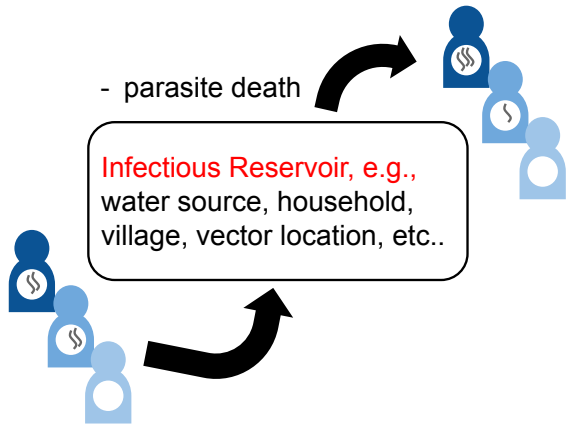

Figure 1: Illustration of the transmission dynamics of human helminth infections via a defined reservoir of infection.

to those of known random walks, e.g., the same heavy-tailed distributions as those of Lévy flights [22, 23]. Furthermore, in a variety of countries and levels of urbanisation, Ref. [24] concludes that there is evidence for some universality in the distribution of daily work-home commute distances, which appear to fit a broken power-law. Note that the distributions are similar to, e.g., those observed for radial distance in Ref. [25]. In Sec. 2, we derive a simple model with one jump per individual, which both replicates this universal behaviour and also suggests a possible mechanism for its origin. In Sec. 3 we extend this model to incorporate multiple jumps per individual and longer timescales of movement patterns — both of which are important for an accurate representation of human behaviour.

Having developed our model and illustrated its possible extensions, in Sec. 4 we apply it to obtain a critical spatial scale at which the migration of individuals between defined infectious reservoirs declines to a negligible level. This result is shown to vary with helminth species, human movement patterns and regional geometry. The analysis also provides insight into the most important pieces of (often missing) information which are necessary to build an accurate model of human migration patterns between specified reservoirs of infection. Collecting such information will be key to the success of future helminth control programmes in reaching their targets, e.g., achieving STH elimination as a public health problem [26]. Lastly, in Sec. 5, we discuss how the spatial extent of reservoirs of infection considered here should influence the scale of EUs for control programmes and conclude with a summary of our findings and prospects for future work.

2. A one-jump model

2.1. Locations of infectious reservoirs and households

Due to the obvious important ethical implications, the availability of high-quality publicly available data on the spatial locations and movement patterns of individuals over various scales in time is scarce. This represents a challenge for those concerned with the various consequences of such movements and their methods of model validation. In this paper, we shall ground our mathematical models in reality by implementing their dynamics on the spatial patterns of, e.g., building locations, obtained from real-world datasets, e.g., the high resolution settlement layer dataset generated by the Facebook Connectivity Lab [27]. Other sources of data, e.g., call data records (from mobile phones) and data from migration questionnaires in control studies, may be available under certain circumstances. The models we introduce here should also be able to take this specific case-study information into account by appropriate parameter inference.

Let us first consider the embedding of spatial locations of infectious reservoirs, e.g., buildings, water sources, known vector population locations and households, as points in a 2-dimensional Euclidean coordinate system. Given this embedding, and the observed spatial distributions of buildings in the high resolution settlement layer dataset [27], for any chosen point we shall set the expectation of having $n_{<r}$ neighbouring points within a radial distance of r, i.e., $\mathrm{E}(n_{<r})$, as a quantity which scales according to the following power law

$$\mathrm{E}(n_{<r}) \propto r^{\alpha(r)} \,. \tag{1}$$

Note that in Eq. (1) we have given the value of the power-law index $\alpha(r)$ a radial dependence. By averaging over the observed point spread patterns Eq. (1) in the high resolution settlement layer dataset, we obtain the following cumulative neighbouring point number distributions as a function of radial distance. In order to give an indication of the effect of local spatial heterogeneity on this averaged quantity, we have plotted some example lines from individual randomly sampled initial building locations for different countries in Fig. 2. In Central Malawi, up to some scatter, below a critical radial scale $r < r_\mu$, the index follows $\alpha \simeq 1$ and above this scale $r > r_\mu$ the index $\alpha \simeq 2$. In both Northern Benin and the Central Ivory Coast, we see that this trend also appears but with less clarity — where some of the initial locations also appear to follow a trendline with $\alpha \ll 1$ for some range up to some critical scale r_μ, and then at radial distances much greater than this scale $r \gg r_\mu$ most of the power-laws tend to $\alpha \simeq 2$. For all countries and initial locations, most of the power-laws appear to converge to $\alpha \simeq 2$, which reflects the emergence of large-scale statistical spatial homogeneity.[1]

It is clear that not all points on our 2-dimensional embedding should be considered as reservoirs of infection, since, e.g, a household cannot be considered a reservoir of shistosoma larvae but it can be considered a reservoir of STH eggs — which may be injested through poor hygiene. Therefore, in order to provide a realistic spatial model for movement between reservoirs of infection, we will need to include the identification of reservoirs with particular spatial points. Note also that Eq. (1) may be generalised to include a spatial location dependence to account for substantial local heterogeneity in the distribution of

[1] This power-law signature is immediately obvious from the radial scaling of the flat measure scaling in 2-dimensional Euclidean space $r dr \propto r^2$ where heterogeneities have effectively been 'smoothed out' over large spatial scales. Deviations from this trend are also present in some locations due to the presence of substantial heterogeneity through clustering of locations. We discuss this effect further and how it may modify our results in Sec. 3.3.

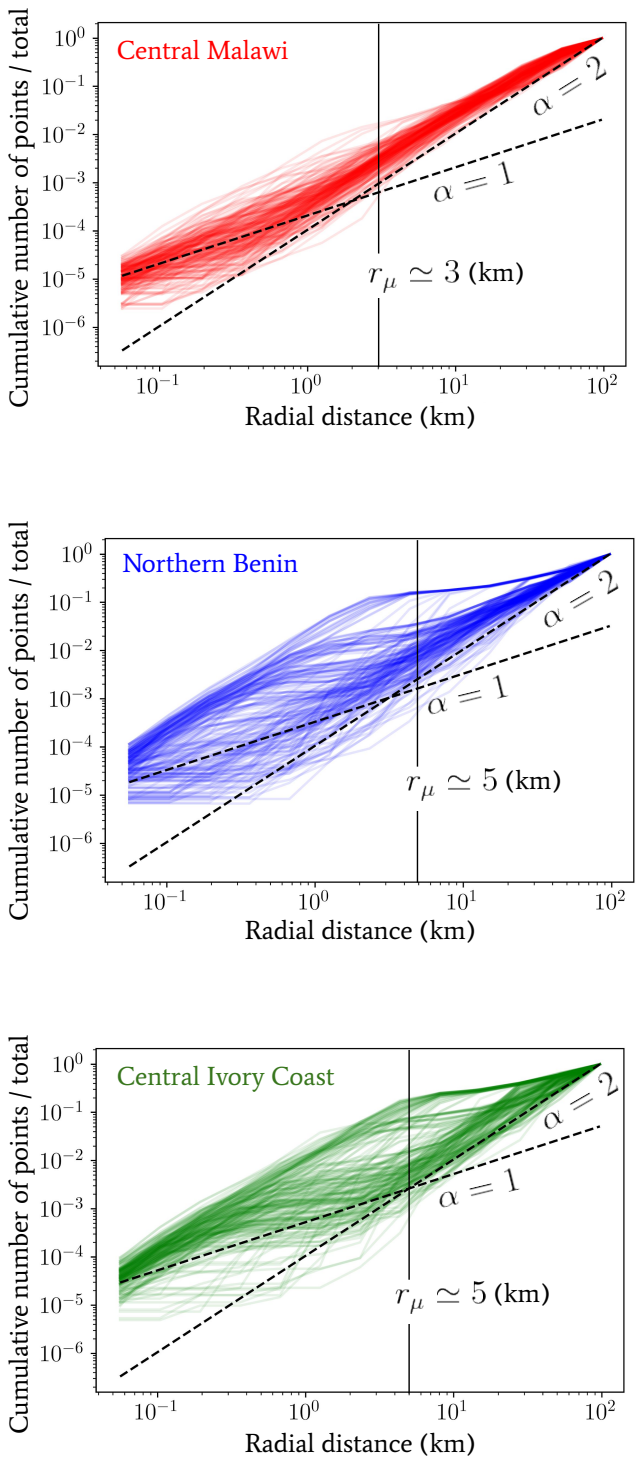

Figure 2: The cumumlative number of buildings (vertical axis) below a given radial distance (horizontal axis) from sample locations in Malawi, Benin and the Ivory Coast obtained from the high resolution settlement layer dataset generated by the Facebook Connectivity Lab [27]. In each case, the collection of sample locations was randomly drawn from: the central region of Malawi, including Lilongwe; the wide northern region of Benin, spanning between Kandi and Djougou; and a wide central region of the Ivory Coast, including Yamoussoukro.

points and/or a time dependence if, as is the case for LF and its mosquito vector, the reservoir location is capable of moving between spatial points over time. We shall discuss the effect of heterogeneity in the proceeding sections but will leave the latter extension to future work on spatial LF models.

2.2. *Obtaining a model using maximum entropy arguments*

In this section we provide a simple argument to generate the broken power-law frequency distribution of work-home commute distances (as observed in Ref. [24]) using maximum entropy arguments. For simplicity of the expressions we derive, we shall assume that the power-law distribution of nearest neighbouring points below a given distance — as given by Eq. (1) — has a constant index $\alpha(r) = \alpha$. Note, however, the results we obtain may be easily generalised to take into account transitions between different power-law behaviour.

Assuming that each individual jumps to a single new location and then returns back to their starting position over the course of each day, we may follow a simple argument that incidentally generates the observed distribution over travel distances in Ref. [24]. Let us take the position of neighbouring points to be statistically isotropic. Defining $p(r)$ as the probability density function (PDF) of an individual jumping to a given location as a function of its radial distance, one may then transform the flat radial probability measure $p(r)\mathrm{d}r$ by the following Jacobian

$$p(r)\mathrm{d}r \to p(r)\nu(r)\mathrm{d}r \equiv p(r)\frac{\mathrm{dE}(n_{<r})}{\mathrm{d}r}\mathrm{d}r \propto p(r)r^{\alpha-1}\mathrm{d}r\,, \qquad (2)$$

which accounts for the power-law density of points that may depart from linear scaling according to Eq. (1).

Note that if we additionally assume that the PDF of the jump distribution for each individual has a mean σ then the maximum entropy PDF is that of an exponential distribution $p(r) = p(r;\sigma) = \mathrm{Exp}(r;1/\sigma)$, then by using Eq. (2) we can trivially demonstrate that the resulting normalised PDF is a Gamma distribution

$$\begin{aligned} p(r;\alpha,\sigma) &= \frac{\nu(r)\mathrm{Exp}(r;1/\sigma)}{\int_0^\infty \nu(r)\mathrm{Exp}(r;1/\sigma)\,\mathrm{d}r} \\ &= \mathrm{Gamma}(r;\alpha,1/\sigma)\,. \end{aligned} \qquad (3)$$

Assuming that predispositions for $1/\sigma$ values may vary across the population of individuals with a distribution of known mean $\mathrm{E}(1/\sigma)$ and expected scaling $\mathrm{E}(\ln\sigma)$, the maximum entropy distribution for values of $1/\sigma$ is another gamma distribution. Choosing specific parameter values, we therefore have

$$1/\sigma \sim \mathrm{Gamma}(1/\sigma;\beta,r_\mu)\,, \qquad (4)$$

where the resulting marginal jump PDF is obtained by integration in the following way

$$\begin{aligned} p(r;\alpha,\beta,r_\mu) &= \int_0^\infty \mathrm{d}(1/\sigma)\,\mathrm{Gamma}(r;\alpha,1/\sigma)\,\mathrm{Gamma}(1/\sigma;\beta,r_\mu) \\ &= \int_0^\infty \mathrm{d}(1/\sigma)\frac{r^{\alpha-1}e^{-\frac{r}{\sigma}}}{\Gamma(\alpha)\,\sigma^\alpha}\frac{r_\mu^\beta(1/\sigma)^{\beta-1}e^{-r_\mu(1/\sigma)}}{\Gamma(\beta)} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{r_\mu^\beta r^{\alpha-1}}{\left(r_\mu+r\right)^{\alpha+\beta}}\,, \end{aligned} \qquad (5)$$

which takes the form similar to that of a beta distribution.[2]

In Fig. 3 we have plotted the marginal jump PDF given by Eq. (5) for a range of parameter values α, β and r_μ. By inspection of Eq. (5), in the limit where $r \ll r_\mu$ the distribution scales as $\sim r^{\alpha-1}$, whereas in the opposite limit $r \gg r_\mu$ the distribution exhibits a scaling $\sim r^{-\beta-1}$. The scale of r_μ therefore separates two regimes in the distribution of jump lengths for a given individual. For radial distances below r_μ, the geometric distribution of available spatial locations to jump to (which is encoded in the α parameter) dominates the behaviour, whereas, for radial distances above r_μ, the predisposition of individuals to travel a given distance combines with the geometry of points to give the distribution behaviour and this is encoded in the β parameter of the second gamma distribution (with PDF Gamma$(1/\sigma; \beta, r_\mu)$) that was used to obtain Eq. (5).

So far we have assumed that the parameter r_μ in the distribution over jump scales $1/\sigma$ given by Eq. (4) is essentially geometrically determined by the observed break in the power laws of Fig. 2. This assumption, in fact, is consistent with comparing our observed power law for the Ivory Coast with the observed work-home commuter distance distributions in Ref. [24], so we shall continue to apply it here without additional data available. Note, however, that this scale may also arise, e.g., due to the availability of easy access to vehicles that allow for longer-distance travel for either work or school, or water sources near a given settlement, which would render our current assumption about the value of the r_μ scale in Eq. (4) less valid.

Despite statistical isotropy and homogeneity on large spatial scales, we have observed that the variation between the power-laws for cumulative building numbers observed in Fig. 2, and hence the α index of Eq. (1), exhibits a specific location dependence between regions. In Fig. 4, we see the effect of this local heterogeneity is relatively small for our one-jump model. We plot a comparison between the binned frequency of 10^3 individuals (with exponential jump distributions each with a scale drawn from Eq. (4)) travelling a given radial distance from their initial position (solid lines) on real-world map data, corresponding to the same regions of each country as in Fig. 2, and the marginalised jump probability densities derived from Eq. (5) for choices of $\alpha = 1, 2$ (and in all cases we have fixed $\beta = 2$ which is approximately the same as the universal tails observed in the distance distributions of Ref. [24]). The agreement between the $\alpha = 1$ curves and the distances generated from single jumps between buildings on the real-world map data appears to be good and hence our one-jump model appears to be relatively unaffected by local spatial heterogeneity. We shall find, however, that the latter will have an important role to play in multi-jump processes as we will discuss in more detail in Sec. 3.3.

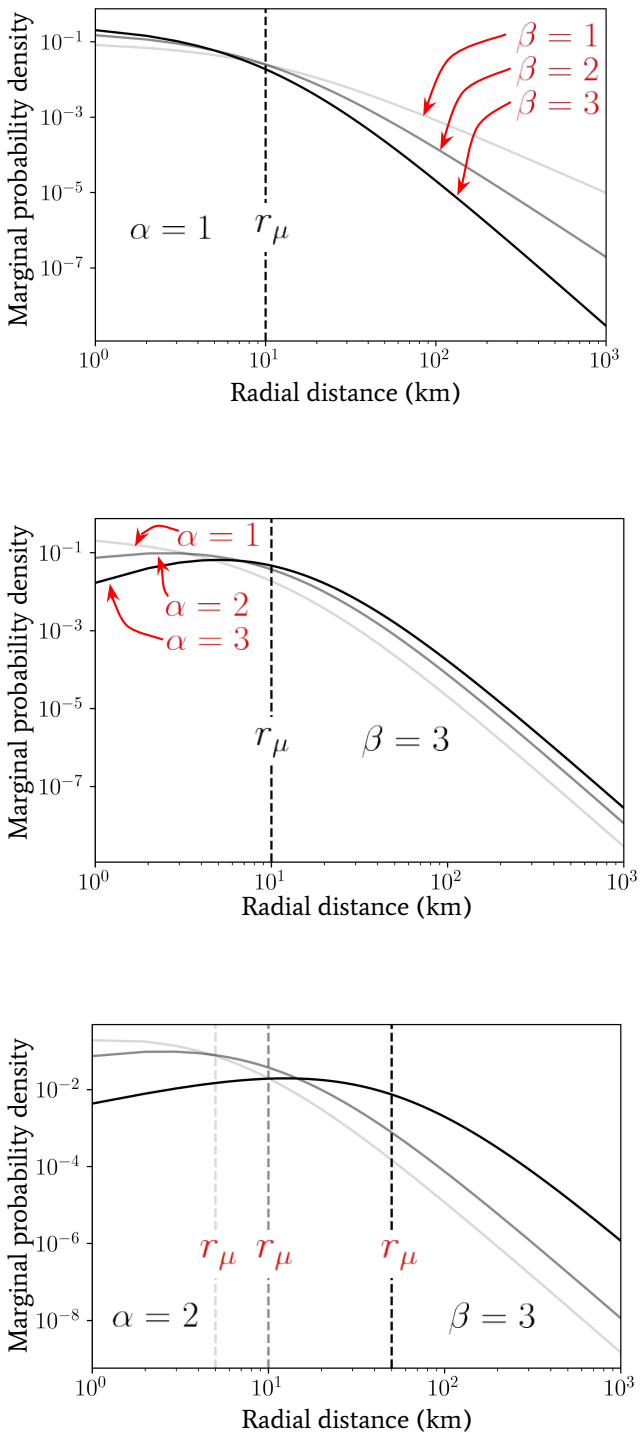

Figure 3: Plots of the marginal probability density $p(r; \alpha, \beta, r_\mu)$ as a function of radial distance (in units of km) generated using Eq. (5) for a range of α, β and r_μ values for comparison.

3. Multi-jump processes

In this section, we shall extend the basic model of human movement presented in Sec. 2 to include multiple successive jumps by individuals between reservoirs of infection over

[2] To avoid a divergent integral one must specify that $\alpha, \beta, r_\mu \in \mathbb{R}^+$.

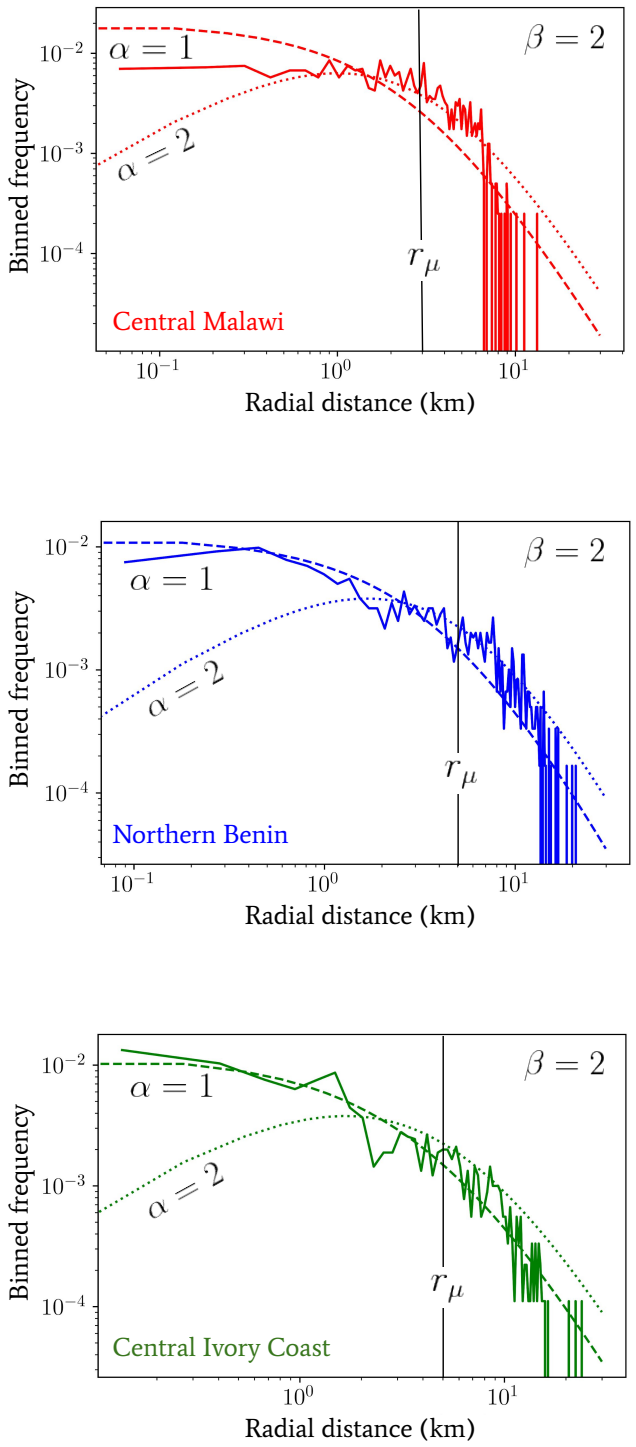

Figure 4: The binned frequency of 10^3 individuals travelling a given radial distance from their initial position (solid lines) corresponding to the same regions of each country as in Fig. 2. This corresponds to the same marginalised jump probability density as Eq. (5) but instead of using a geometric power-law given by Eq. (1) we have implemented our single jumps on real-world map data (with exponential jump distributions each with a scale drawn from Eq. (4)). For comparison, we have also plotted some jump probabilities calculated using Eq. (5) with $\alpha = 1$ (dashed lines) and $\alpha = 2$ (dotted lines). In all cases we have fixed $\beta = 2$.

longer time periods. The type of human movement we aim to capture is not just that of daily commuting, but also longer distance travel across multiple days/weeks. The latter form of movement can contribute to the variability in migrant seasonal labour or family visits which can increase/decrease the population numbers in a given region on a yearly timescale.

3.1. Multiple randomly-directed jumps

Following along similar lines to Sec. 2.2, let us now consider an individual who completes multiple successive exponentially distributed jump lengths r_i (where we have indexed each successive jump event with an i) with a fixed mean scale σ within a day. For a fixed jump rate $\mathcal{J}$ in time, consider the following compound Poisson process for a two-dimensional vector $\boldsymbol{x}(t)$ which encodes the 2-dimensional Euclidean coordinate position of an individual over time

$$\boldsymbol{x}(t) = \sum_{i=1}^{\infty} \begin{bmatrix} r_i \cos(\theta_i) \\ r_i \sin(\theta_i) \end{bmatrix} \mathbb{1}_{[t_i,\infty)}(t) \tag{6}$$

$$t_i - t_{i-1} = \Delta t_i \sim \mathrm{Exp}(\Delta t_i; \mathcal{J}) = \mathcal{J} e^{-\mathcal{J}\Delta t_i} \tag{7}$$

$$\theta_i - \theta_{i-1} = \Delta\theta_i \sim \mathrm{U}(\Delta\theta_i; -\pi, \pi) = \frac{1}{2\pi}\Theta(\pi - \theta_i)\Theta(\pi + \theta_i)\,, \tag{8}$$

where $\mathrm{U}(\Delta\theta_i; -\pi, \pi)$ is the uniform distribution PDF (assuming isotropy) over the change in angle and $\Theta(\cdot)$ is a Heaviside function.

The process specified by Eq. (6) assumes isotropy of both the geometry of the points and a uniform-random direction choice of the individual. The distribution $p[\boldsymbol{x}(t)]$ is not known as a closed-form expression, however, it clearly has both a vanishing first moment

$$\begin{aligned} &\mathrm{E}[\boldsymbol{x}(t)] = \\ &\int_{-\pi}^{\pi} \mathrm{d}\theta_i \int_0^{\infty} \mathrm{d}r_i \sum_{j=1}^{\infty} \mathrm{Pois}[j; \mathcal{J}(t - t_0)] \sum_{i=1}^{j} \begin{bmatrix} r_i \cos(\theta_i) \\ r_i \sin(\theta_i) \end{bmatrix} p(r_i; \alpha, \sigma) \\ &= \begin{bmatrix} \mathcal{J}(t - t_0)\mathrm{E}(r_i)\mathrm{E}[\cos(\theta_i)] \\ \mathcal{J}(t - t_0)\mathrm{E}(r_i)\mathrm{E}[\sin(\theta_i)] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned} \tag{9}$$

and a second moment which scales according to σ^2

$$\begin{aligned} \mathrm{E}[\boldsymbol{x}(t) \cdot \boldsymbol{x}(t)] &= 2\pi\mathcal{J}(t - t_0)\mathrm{E}(r_i^2) \\ &= 2\pi\mathcal{J}(t - t_0)(\alpha + \alpha^2)\sigma^2\,. \end{aligned} \tag{10}$$

Furthermore, with a scale distribution over jump lengths applied — as in Eq. (5) — the second moment given in Eq. (10) becomes

$$\mathrm{E}[\boldsymbol{x}(t) \cdot \boldsymbol{x}(t)] = 2\pi\mathcal{J}(t - t_0)\frac{(\alpha + \alpha^2)r_\mu^2}{(\beta - 1)(\beta - 2)}\,. \tag{11}$$

In the left panel of Fig. 5 we have plotted the binned frequency of total daily distances travelled by a population of 5×10^4 individuals following the process defined by Eq. (6) and drawing each individual's jump scale $1/\sigma$ from Eq. (4). We have also fixed $\mathcal{J} = 1, 20$ per day (in red and black lines, respectively), $\alpha = 1$ and $\beta = 1, 2, 3$ as indicated by the increasing

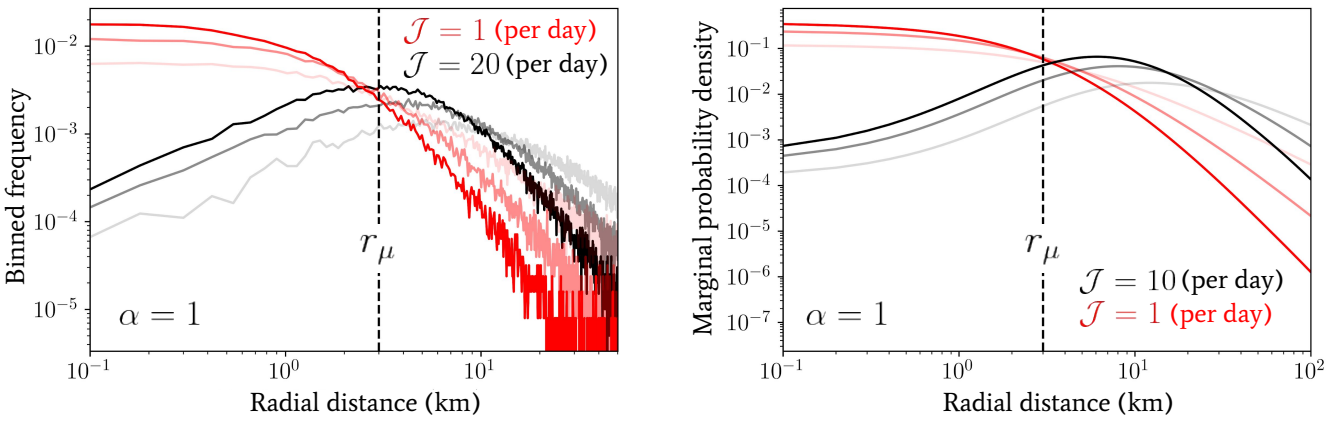

Figure 5: Numerical plots of the distance distributions generated by the multi-jump processes introduced in Sec. 3.1 with randomly-directed (left panel) and in Sec. 3.2 with unidirectional jumps (right panel). In the left panel, a population of 5×10^4 individuals following the process defined by Eq. (6) have been drawn (drawing from Eq. (4) for their jump scale predispositions) and the binned frequency of the total distance evaluated at the end of the day with jump rates of $\mathcal{J} = 1, 20$ per day (in red and black lines, respectively), $\alpha = 1$ and $\beta = 1, 2, 3$ as indicated by the increasing opacity within each triplet of lines. In the right panel, the marginal probability density given by Eq. (15) is depicted with jump rates of $\mathcal{J} = 1, 10$ per day (in red and black lines, respectively), $\alpha = 1$ and $\beta = 1, 2, 3$ as indicated by the increasing opacity within each triplet of lines.

opacity within each triplet of lines. Contrasting these distributions, it is immediately clear that by increasing the jump rate $\mathcal{J}$, the effect on the distribution of daily distances travelled is effectively the same as increasing the value of α — see the middle panel of Fig. 3 for comparison. As one increases either α or $\mathcal{J}$, the distribution tends to monotonically increase with an increasingly steep gradient up to the value of the r_μ scale.

3.2. Unidirectional jumps

Human movement patterns are not truly random, and instead, one might anticipate a strong directional dependence, e.g., to long distance travel of individuals over the course of multiple days to an intended destination. Due to this fact, let us assume that the angular dependence of an individual's long-distance travel is negligible. Following multiple successive jumps while assuming no variation in angle, Eq. (6) then becomes a compound Poisson process for the total jump distance $x(t) \equiv |\boldsymbol{x}(t)|$ in a single dimension. This process has a known characteristic function

$$\varphi_x(s) = \exp\left\{\mathcal{J}(t - t_0)\left[\varphi_r(s) - 1\right]\right\}, \tag{12}$$

where we have set $v(t_0) = 0$ and $\varphi_r(s)$ denotes the characteristic function of the stationary increments given by Eq. (3). Hence, we may write

$$\varphi_x(s) = \exp\left\{\mathcal{J}(t - t_0)\left[(1 - \sigma i s)^{-\alpha} - 1\right]\right\}. \tag{13}$$

which has no closed-form inverse transformation, but is useful as an expression to calculate the moments. Alternatively, in order to find the distribution over $x(t)$, we may marginalise over the number of jumps performed (which is Poisson-distributed by construction), which are themselves drawn from the jump distance distribution in the following way

$$p[r = x(t) \,|\, \theta_i = \theta_{i-1}] = \sum_{j=0}^{\infty} \mathrm{Pois}[j; \mathcal{J}(t - t_0)]\, \mathrm{Gamma}[r; j\alpha, 1/\sigma]\,. \tag{14}$$

Therefore, if $\alpha = 1$, Eq. (14) yields

$$p[r = x(t) \,|\, \theta_i = \theta_{i-1}, \alpha = 1] = \sqrt{\frac{\mathcal{J}(t - t_0)}{\sigma r}}\, e^{-\mathcal{J}(t-t_0) - \frac{r}{\sigma}}\, \mathcal{I}_1\left[2\sqrt{\frac{\mathcal{J}(t - t_0) r}{\sigma}}\right], \tag{15}$$

where $\mathcal{I}_n(\cdot)$ is the modified Bessel function of the first kind.[3]

In the right panel of Fig. 5 we have plotted the marginal probability density given by integrating Eqs. (15) and (4) over $1/\sigma$ for direct comparison with the one-jump model of Eq. (5). We have used values of $\mathcal{J} = 1, 10$ per day (in red and black lines, respectively), $\alpha = 1$ and $\beta = 1, 2, 3$ as indicated by the increasing opacity within each triplet of lines. Comparing this distribution with the one observed for randomly-directed multiple jumps (see Sec. 3.1), a similar, but even more pronounced, effect on the distribution of total daily distances travelled is observed, since only a value of $\mathcal{J} = 10$ is required to achieve roughly the same change. This is to be expected, as the unidirectional jumpers will always travel at least the same (or more often) a greater distance in total in comparison with the

[3] Additionally, if $\alpha = 2$, one obtains

$$p[r = x(t) \,|\, \theta_i = \theta_{i-1}, \alpha = 2] = \frac{\mathcal{J}(t - t_0) r}{\sigma^2}\, e^{-\mathcal{J}(t-t_0) - \frac{r}{\sigma}}\, {}_0F_2\left[; \frac{3}{2}, 2; \frac{\mathcal{J}(t - t_0) r^2}{4\sigma^2}\right]$$

where ${}_0F_2(;\cdot,\cdot;\cdot)$ is a hypergeometric function.

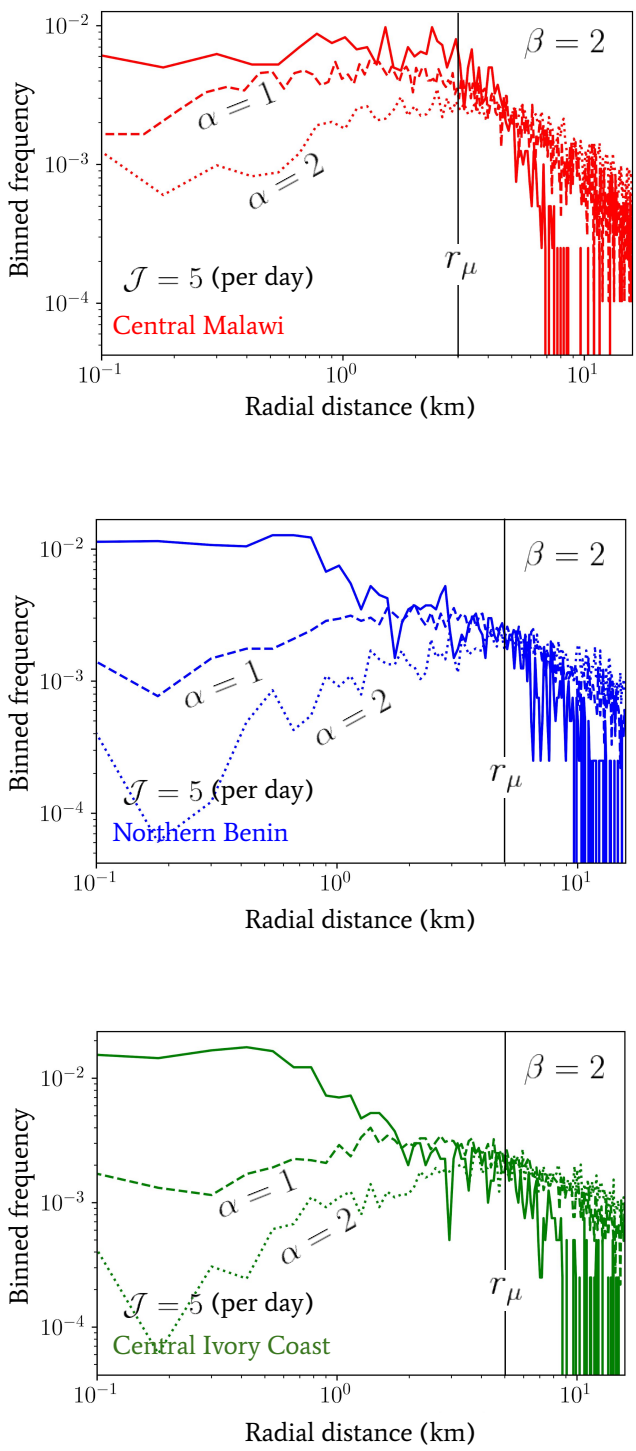

Figure 6: The multi-jump analogue of Fig. 4 with randomly-directed jumps chosen (as discussed in Sec. 3.1). The binned frequency of 10^3 individuals travelling a given radial distance from their initial position (solid lines) corresponding to the same regions of each country as in Fig. 2 with $\mathcal{J} = 5$ per day. For comparison, we have also plotted the binned frequency of the total distances achieved by 10^3 individuals computed through the multi-jump process model of Eq. (6) (and drawing jump scale predispositions from Eq. (4)) with settings of $\alpha = 1$ (dashed lines) and $\alpha = 2$ (dotted lines). In all cases we have fixed $\beta = 2$.

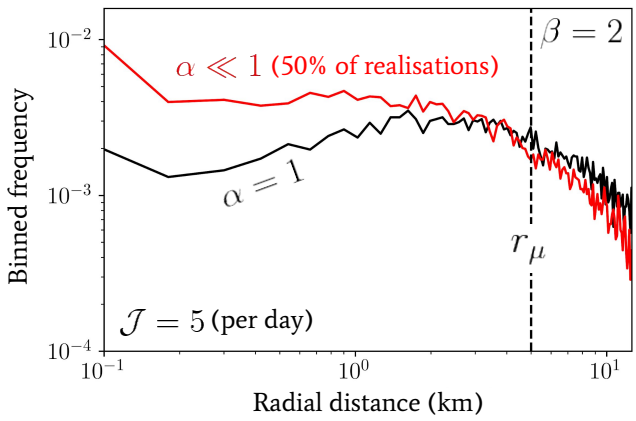

Figure 7: Binned frequency of total distances achieved by individuals carrying out the multi-jump process with randomly-directed jumps chosen (given by Eq. (6) and as discussed in Sec. 3.1). With a black line we depict the standard process (with $\alpha = 1$, $\beta = 2$ and $\mathcal{J} = 5$ per day) for comparison with the red line, which depicts a modification where a random 50% of the realisations are assigned $\alpha \ll 1$ ($\alpha = 0.05$ in this case) to mimic the effect of local heterogeneity between sites. In this case we have used 10^4 individuals for clarity of presentation.

randomly-directed ones, which will likely induce a more severe deformation of the overall distribution when using the former.

Let us note there that such a degeneracy between effects, as described above and found in Sec. 3.1, on the observed distributions between increasing either the α of the one-jump model or $\mathcal{J}$ in the multi-jump processes, suggests a test can be performed to either validate a model or discard it in favour of a modified version for a given situation with real data. By combining collected data on the total distance travelled per day and the local geometry of sites travelled to in a given region, the jump rate for each individual (if the same) could be statistically inferred. Comparing this inferred value to real data would provide a test of the movement models we have suggested in this work or potentially provide insight into where they may be improved to better reflect real human daily movement.

3.3. *Including local spatial heterogeneity*

In Fig. 6 we plot the multi-jump analogue of Fig. 4 with jump scales drawn from Eq. (4) undergoing randomly-directed jumps over the real map data (solid lines) and the generated model jumps from the process described in Sec. 3.1 with $\alpha = 1, 2$ (dashed and dotted lines, respectively). As in Fig. 4, red lines (top panel) denote central Malawi, blue lines (middle panel) denote northern Benin and green lines (bottom panel) denote central Ivory Coast. In all three cases the same region is studied in Fig. 2. In contrast to the level of agreement found in Fig. 4, we find that there is a significant difference between the distributions from the real world map data and those of the model — particularly in the latter two locations of northern Benin and central Ivory Coast — and this difference is generated due to the cumulative effect of local spatial heterogeneity.

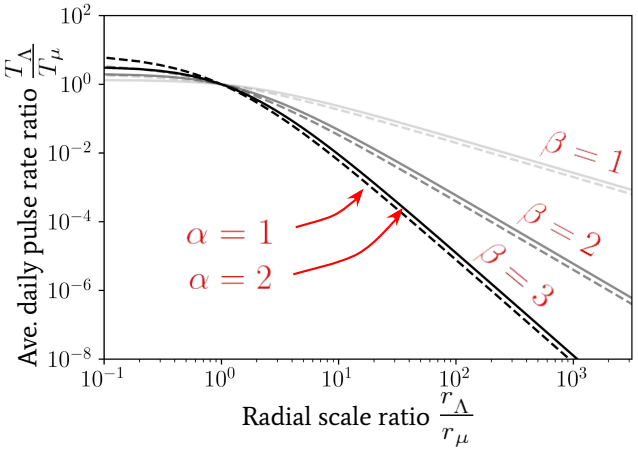

Figure 8: The renormalised average daily pulse rate T_Λ into a reservoir of infection as a fraction of its value T_μ from individuals arriving a distance of r_μ or greater away. This value is plotted as a function of the radial coarse-graining scale ratio r_Λ/r_μ used. The relationship is given by Eq. (19) for a range of α and β power-law parameters.

At the end of Sec. 2.2 we pointed out that there is significant variability between the observed power laws in Fig. 2 for the real world map data, which is dependent upon the local configuration of sites around a given point of reference. In particular, recall that Fig. 2 indicates some of the locations have a local power law of $\alpha \ll 1$ for some radial range up to the critical scale of r_μ. This variation between locations does not affect the jump distribution significantly from our model when only one jump is considered — see, e.g., the agreement in Fig. 4 — however, multi-jump processes allow for a greater degree of variability as the variance in local power laws is additive.

We have confirmed that the local variability in power laws generates similar modifications to those observed in Fig. 6 and have plotted an example to illustrate this effect in Fig. 7. In the black line of Fig. 7, we have plotted the binned frequency of total distances achieved using that same model as the dashed lines of Fig. 2 but with $\mathcal{J} = 5$ and 10^4 realisations for clarity of presentation. In the red line of the same plot, however, we produce the result from the same process but where 50% of the realisations have been randomly chosen with $\alpha \ll 1$ ($\alpha = 0.05$ in this example), where the shape of this distribution deviates from the original black line in a similar fashion to the observed differences in Fig. 6.

A more sophisticated method to modify the model (than the one shown in Fig. 7) would be to sample from the numerically-obtained power laws from Fig. 2 to emulate the local spatial heterogeneity of the real world map directly in the processes of Sec. 3. This method would be significantly advantageous, in terms of computational complexity, to direct simulation methods over the real world map data such as those used in Fig. 6. We plan to investigate this, and related methods as a means to efficiently simulate large-scale spatial movement patterns of individuals in future work.

4. The effect on reservoirs of infection

Locally-defined epidemiogical variables f, e.g., the prevalence of infection or mean parasite burden of hosts within a given region, will typically have a dependence on the local population number $N(t)$ (which itself may be time t dependent), the amount of material in the reservoir of infection per individual $L(t)$ (such that the total amount is simply $N(t)L(t)$), as well as other local epidemiogical variables, i.e.,

$$f = f[N(t), L(t), R_0, \dots] \,. \tag{16}$$

After some finite period in time δt (which may be, e.g., a day for our purposes), if an individual enters or leaves the spatial region over which L is defined, due to the heterogeneity between regions, the value of $L(t) \neq L(t + \delta t)$, and hence f, may be modified. Therefore, in order to understand how local variables such as f are affected by spatial movements of human hosts, it is essential to correctly assess how $L(t)$ is modified by such migrations.

Due to the fact that $N(t)L(t)$ is a summation over the total quantity of infectious material in the reservoir, if the individual who leaves or enters the region over δt contributes to the reservoir of infection during this time, then they will perturb $N(t)L(t)$ in the following way

$$N(t)L(t + \delta t) = N(t)L(t) \pm \ell(t) \,, \tag{17}$$

where $\ell(t)$ is the amount of infectious material generated by one individual. Therefore, $\ell(t)$ must be a time-dependent random variable that also depends on the local epidemiological variables to the individual who generates this material (which may not be the same as those of the region they have just entered).

Note that in order to obtain Eq. (17) above, we have assumed that $N(t) = N(t + \delta t)$ over the time period of interest, which may not be the case if δt is taken to be as long as, e.g., a year. This is because secular effects which can cause the local population number to vary include periodic seasonal migrant labour or family visits at particular times of year. If one considers variations in the reservoir of infection on shorter timescales, however, Eq. (17) is certainly valid. Hence, choosing a day as our dynamical timescale, 'pulses' in the reservoir of the form given by Eq. (17) occur at the daily rate at which individuals travel to a specified reservoir of infection and then contribute to it.

Let us define T_Λ as the average daily rate of individuals travelling greater than a distance of r_Λ to a pre-defined reservoir of infection and subsequently contributing to it. Equivalently, if T_μ is the average daily rate of individuals travelling greater than a distance of r_μ and then subsequently contributing to the reservoir, then the following scaling between these quantities should exist which accounts for the difference in probability mass arising from the rescaling of distance

$$\frac{T_\Lambda}{T_\mu} = \frac{\int_{r_\Lambda}^{\infty} \mathrm{d}r\, p(r;\alpha,\beta,r_\mu)}{\int_{r_\mu}^{\infty} \mathrm{d}r\, p(r;\alpha,\beta,r_\mu)} = \frac{1 - \int_0^{r_\Lambda} \mathrm{d}r\, \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{r_\mu^\beta r^{\alpha-1}}{(r_\mu+r)^{\alpha+\beta}}}{1 - \int_0^{r_\mu} \mathrm{d}r\, \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{r_\mu^\beta r^{\alpha-1}}{(r_\mu+r)^{\alpha+\beta}}} \,. \tag{18}$$

Note that to derive Eq. (18) we are assuming that all individuals jump once per day so that the marginalised jump PDF

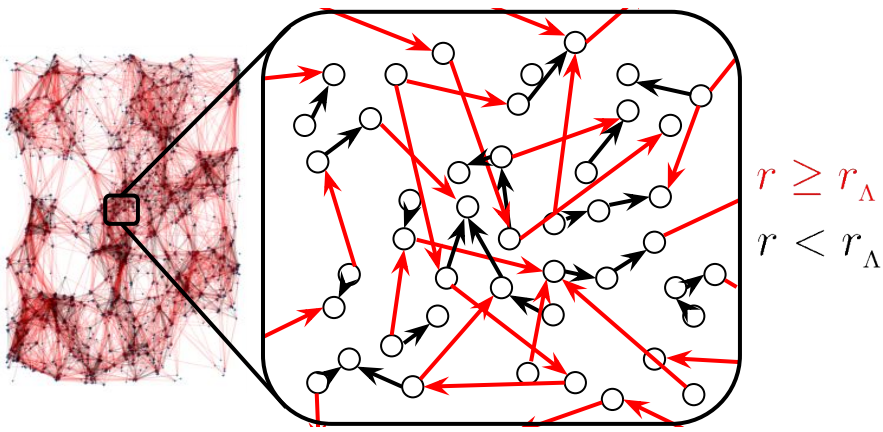

Figure 9: A diagram of real map data (based on buildings in central Malawi) and a zoomed illustration which indicate the coarse-graining procedure of Eq. (18) — in which the black arrows are 'removed'. Arrows depict individual movements between sites (given by white dots on the zoomed version) where black arrows on both the map and its zoomed counterpart correspond to jumps over a radial distance $r < r_\Lambda$ (where r_Λ is the spatial coarse-graining scale) and red arrows correspond to jumps over a radial distance $r \geq r_\Lambda$.

$p(r; \alpha, \beta, r_\mu)$ of Sec. 2 may be used directly. By integration of Eq. (18), one generally obtains

$$\frac{T_\Lambda}{T_\mu} = \frac{1 - \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)} \left(\frac{r_\Lambda}{r_\mu}\right)^\alpha {}_2F_1\left(\alpha, \alpha+\beta; \alpha+1; -\frac{r_\Lambda}{r_\mu}\right)}{1 - \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)} {}_2F_1\left(\alpha, \alpha+\beta; \alpha+1; -1\right)}, \quad (19)$$

where ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ is the (Gauss) hypergeometric function. In Fig. 8 we use the solution given by Eq. (19) to compute the quantity T_Λ/T_μ as a function of r_Λ/r_μ for a range of parameter choices.

Note that, by construction, r_Λ acts as a spatial coarse-graining or 'renormalisation' scale, since the removal of jumps from distances below this value in T_Λ is tantamount to renormalising the average daily rate of contribution of individuals to reservoirs — the rate at which pulses of the form given by Eq. (17) occur — up to this specified spatial resolution. The effect of this spatial coarse-graining is illustrated Fig. 9 on a digram with real map data based on buildings in central Malawi. The black (red) arrows on the map and zoomed illustration show jumps below (above or equal to) the radial distance of r_Λ which has been arbitrarily chosen.

In Fig. 9, there is an emergence of black clusters of connected locations on a spatial scale specified by r_Λ which are connected by red arrows at longer distances. By increasing r_Λ, we may therefore identify structures in the map of locations of real data on increasing spatial scales, e.g., buildings, villages, towns, cities, etc. The variation of T_Λ as a function of r_Λ hence provides an average reservoir pulse rate for clustered communities at different spatial coarse-graining scales — the latter of which may be fixed to identify movement of individuals between structures at different scales, e.g., villages or cities.

It was shown in Refs. [13, 14] that if the average daily rate at which reservoir pulses occur exceeds the average death rate of the infectious material in the reservoir per day, d_{res}, then the effect of infected human migration on the transmission dynamics becomes particularly strong. By combining these two concepts above, one may derive a new and important spatial scale for reservoirs of infection which depends on the type of helminth, as well as the local geometry of locations and human movement behaviour.

The spatial epidemological transmission process of the helminth disease defines a relevant scale for reservoirs of infection, e.g., household, village, cluster, etc., which is obtained through statistical inference with relevant data — see, e.g., [28, 4, 29]. If the spatial scale of this 'epidemiological unit' is smaller than the critical scale r^*_Λ, the average daily rate of pulses will always be larger in magnitude than its value using the critical scale T^*_Λ because the curves shown in Fig. 8 always decrease monotonically. This implies that when the epidemiological unit is of the size of the critical scale or smaller, pulses of the form given by Eq. (17) cannot be safely neglected and should become important to take into account when modelling the transmission dynamics (see Refs. [13, 14]). If, however, the spatial scale of the epidemiological unit is larger than r^*_Λ, then effects from migration between reservoirs of infection may potentially be safely neglected.

Using Eq. (19) the critical spatial scale r^*_Λ at which T^*_Λ becomes equal to the death rate of the infectious material in the reservoir per day, i.e., $T^*_\Lambda = d_{\text{res}}$, has been plotted for hookworm (which has $d_{\text{res}} \simeq 0.071$ [30]) in Fig. 10 as a function of the average daily pulse rate T_μ from individuals travelling distances at or above r_μ. In this plot we note that nearly all parameter combinations indicate a sharp decline in the critical scale ratio if T_μ is found to be below d_{res} — where, in particular, for values of $\alpha = 2$ the scale ratio appears to fall extremely sharply and hence one can no longer find a finite critical scale below which migration becomes important.

In contrast, for values $T_\mu > d_{\text{res}}$ in increasing orders of magnitude, one infers from Fig. 10 that the value of α is nearly irrelevant (which makes sense by construction of T_μ) and, particularly for $\beta = 2, 3$, the increase in the critical scale ratio is very gradual. Such a relationship is also consistent with our expectation as the decline in the average pulse rate from people travelling from distances $r_\Lambda \gg r_\mu$ is particularly sharp for $\beta = 2, 3$ — see, e.g., Fig. 8 — and so one requires a significant

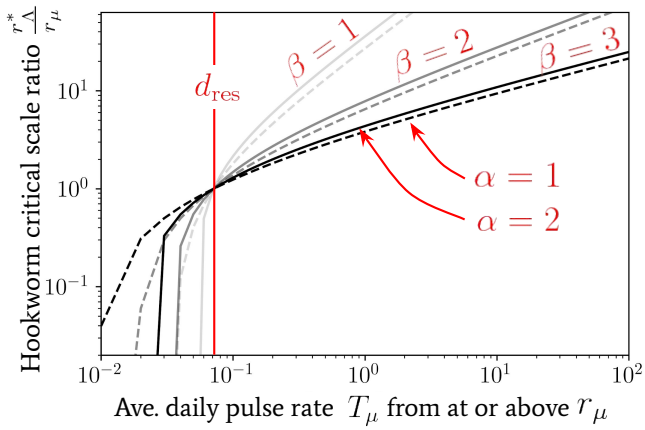

Figure 10: The critical spatial scale as a ratio r^*_Λ at which $T^*_\Lambda = d_{\rm res}$, for hookworm — where $d_{\rm res} \simeq 0.071$ [30] — plotted as a ratio of r_μ. The value of this scale is shown against the value of the average daily pulse rate T_μ from individuals travelling distances at or above r_μ. We have used Eq. (19) to generate this relationship with a (bisection) root-finding algorithm.

increase in the amplitude of T_μ to achieve a significant change in the critical scale.

Despite the interesting results above, we must acknowledge that there are important caveats to this relationship (and the one shown in Fig. 8) which arise from the assumptions made in obtaining Eq. (18), i.e., the statistical homogeneity and isotropy of people living in the surrounding households who all jump only once per day. By relaxing the assumptions made in obtaining Eq. (18), the rate of pulses will vary between different reservoirs of infection due to the effects of local spatial heterogeneity, as we have observed in Fig. 5 and replicated in Fig. 7. In addition to this effect, the inclusion of multiple unidirectional jumps with different initial preferred directions may no longer be isotropically configured when viewed collectively over the course of the day — which would be the case when multiple individuals all converge to a globally preferred location such as a place of work or a school. We propose to consider such modifications in future work on specific case studies.

5. Discussion and conclusions

In this work we have developed and studied spatial models of human movement between reservoirs of infection — the important drivers of new infections for many parasitic organisms in humans. Our models have been generated by combining information from the observed universality of human work-home commute patterns in, e.g., Ref. [24], with data about the spatial patterns of, e.g., building locations, obtained from real-world datasets, such as those of Ref. [27], to provide a generalised description which can be adapted to specific situations. Future work on specific locations will likely require a numerical approach where the model developed in Secs. 2 and 3 will be very useful for computational efficiency.

At its most basic, our model of human mobility assigns a single daily journey (or 'jump') for each individual, however, we have discussed many extensions to this description, which include: multiple Poisson-distributed jumps, each of which may either be in random directions or unidirectional; and substantial local heterogeneity of available travel locations. By combining these extensions into a generalised framework, one has the capability to generate a wide variety of possible daily human movement behaviours between reservoirs of infection under a unifying description which offers a model-focussed explanation for the observed patterns that is also computationally efficient. In future work, it would be interesting to extend our approach to include time dependence in the location of reservoirs, which may exist in diseases such as LF, where the reservoir, in the form of mosquito vectors infected by filarial larvae, can vary spatially over time.

We note here that the human movement models we have studied are comparable to those developed in Refs. [31, 32] for vector-bourne diseases, which label the single daily journey models as so-called 'Lagrangian' mobility models and the multiple successive movement models as 'Eulerian'. We consider our stochastic approach to generating frequency-distance distributions useful and informative in understanding the underlying probability distributions for these works.

Our human movement models also highlight the most important information necessary to efficiently (by which we mean low numbers of parameters) describe the effect of migration patterns on reservoirs of infection and hence the disease transmission dynamics. The necessary data can be collected in well-designed field studies. Based on our work here, information to obtain for a particular study may include:

1. Obtaining the power-law for the expected cumulative number of locations available to travel to by an individual — see, e.g., Fig. 2. These data can be obtained from geographic maps.
2. Evaluating the amplitude of the effect of local heterogeneity (and perhaps anisotropy) on the power-law cumulative distribution of locations available to travel to by an individual.
3. Obtaining the average number of journeys performed within a single day per individual. In this case some metric for having 'completed' a journey associated to the time spent at each location will likely need to be determined.
4. Evaluating the average number of people entering or leaving the relevant reservoir of infection at some particular spatial scale, e.g., this can be T_μ or evaluated at some other spatial scale since the reasoning of Eq. (18) can be easily adapted. Note that the true spatial scale of the reservoir does not need to be known for collecting these data.

Due to the apparent universality of the power-laws obtained from the work-home commute in Ref. [24], we have left out obtaining values for β (see Eq. (4)) from the list compiled above. However, for a rigorous analysis of a particular problem this value should also be investigated.

Lastly, we comment here on the intriguing possibility that the methodology we have developed could be useful in defining the

spatial scale for evaluation units of, e.g., mass drug administration programmes [21, 20]. These should ideally be congruent with epidemiological units defined by the spatial extent of infectious reservoirs. We plan to investigate this idea further in future work.

Acknowledgements

RJH, JET and RMA gratefully thank the Bill and Melinda Gates Foundation for research grant support via the DeWorm3 (OPP1129535) award to the Natural History Museum in London (http://www.gatesfoundation.org/). CV gratefully acknowledges funding from the NTD Modelling Consortium (OPP1184344) by the Bill and Melinda Gates Foundation in partnership with the Task Force for Global Health (http://www.taskforce.org/). The views, opinions, assumptions or any other information set out in this article are solely those of the authors. All authors acknowledge joint Centre funding from the UK Medical Research Council and Department for International Development.

References

[1] R. L. Pullan, P. W. Gething, J. L. Smith, C. S. Mwandawiro, H. J. Sturrock, C. W. Gitonga, S. I. Hay, S. Brooker, Spatial modelling of soil-transmitted helminth infections in kenya: a disease control planning tool, PLoS neglected tropical diseases 5 (2) (2011).

[2] N. Schur, E. Hürlimann, A.-S. Stensgaard, K. Chimfwembe, G. Mushinge, C. Simoonga, N. B. Kabatereine, T. K. Kristensen, J. Utzinger, P. Vounatsou, Spatially explicit schistosoma infection risk in eastern africa using bayesian geostatistical modelling, Acta tropica 128 (2) (2013) 365–377 (2013).

[3] R. L. Pullan, J. L. Smith, R. Jasrasaria, S. J. Brooker, Global numbers of infection and disease burden of soil transmitted helminth infections in 2010, Parasites & vectors 7 (1) (2014) 37 (2014).

[4] C. L. Moyes, A. J. Henry, N. Golding, Z. Huang, B. Singh, J. K. Baird, P. N. Newton, M. Huffman, K. A. Duda, C. J. Drakeley, et al., Defining the geographical range of the plasmodium knowlesi reservoir, PLoS neglected tropical diseases 8 (3) (2014) e2780 (2014).

[5] D.-A. Karagiannis-Voules, P. Biedermann, U. F. Ekpo, A. Garba, E. Langer, E. Mathieu, N. Midzi, P. Mwinzi, A. M. Polderman, G. Raso, et al., Spatial and temporal distribution of soil-transmitted helminth infection in sub-saharan africa: a systematic review and geostatistical meta-analysis, The Lancet infectious diseases 15 (1) (2015) 74–84 (2015).

[6] P. Moraga, J. Cano, R. F. Baggaley, J. O. Gyapong, S. M. Njenga, B. Nikolay, E. Davies, M. P. Rebollo, R. L. Pullan, M. J. Bockarie, et al., Modelling the distribution and transmission intensity of lymphatic filariasis in sub-saharan africa prior to scaling up interventions: integrated use of geostatistical and mathematical modelling, Parasites & vectors 8 (1) (2015) 560 (2015).

[7] N. A. Hamm, R. J. Soares Magalhães, A. C. Clements, Earth observation, spatial data quality, and neglected tropical diseases, PLoS neglected tropical diseases 9 (12) (2015) e0004164 (2015).

[8] S. J. OHanlon, H. C. Slater, R. A. Cheke, B. A. Boatin, L. E. Coffeng, S. D. S. Pion, M. Boussinesq, H. G. M. Zour, W. A. Stolk, M.-G. Basez, Model-based geostatistical mapping of the prevalence of onchocerca volvulus in west africa, PLOS Neglected Tropical Diseases 10 (1) (2016) 1–36 (01 2016). doi:10.1371/journal.pntd.0004328.
URL https://doi.org/10.1371/journal.pntd.0004328

[9] M. Ciddio, L. Mari, S. H. Sokolow, G. A. De Leo, R. Casagrandi, M. Gatto, The spatial spread of schistosomiasis: A multidimensional network model applied to saint-louis region, senegal, Advances in water resources 108 (2017) 406–415 (2017).

[10] L. Mari, M. Gatto, M. Ciddio, E. D. Dia, S. H. Sokolow, G. A. De Leo, R. Casagrandi, Big-data-driven modeling unveils country-wide drivers of endemic schistosomiasis, Scientific reports 7 (1) (2017) 489 (2017).

[11] R. M. Flueckiger, E. Giorgi, J. Cano, M. Abdala, O. N. Amiel, G. Baayenda, A. Bakhtiari, W. Batcho, K. H. Bennawi, M. Dejene, et al., Understanding the spatial distribution of trichiasis and its association with trachomatous inflammationfollicular, BMC infectious diseases 19 (1) (2019) 364 (2019).

[12] C. Vegvari, J. E. Truscott, K. Kura, R. M. Anderson, Human population movement can impede the elimination of soil-transmitted helminth transmission in regions with heterogeneity in mass drug administration coverage and transmission potential between villages: a metapopulation analysis, Parasites & vectors 12 (1) (2019) 438 (2019).

[13] R. J. Hardwick, C. Vegvari, J. E. Truscott, R. M. Anderson, The breakpointof soil-transmitted helminths with infected human migration, Journal of theoretical biology (2019) 110076 (2019).

[14] R. J. Hardwick, M. Werkman, J. E. Truscott, R. M. Anderson, Stochastic challenges to interrupting helminth transmission, medRxiv (2019).

[15] S. Brooker, J. Bethony, P. J. Hotez, Human hookworm infection in the 21st century, Vol. 58 of Advances in Parasitology, Academic Press, 2004, pp. 197 – 288 (2004). doi:https://doi.org/10.1016/S0065-308X(04)58004-1.
URL http://www.sciencedirect.com/science/article/pii/S0065308X04580041

[16] D. W. T. Crompton, Preventive chemotherapy in human helminthiasis : coordinated use of anthelminthic drugs in control interventions : a manual for health professionals and programme managers, World Health Organization, 2006 (2006).
URL http://www.who.int/iris/handle/10665/43545

[17] Ending the neglect and reaching 2020 goals, https://unitingtocombatntds.org/london-declaration-neglected-tropical-diseases/.

[18] Progress reports of the london declaration, https://unitingtocombatntds.org/reports/.

[19] Accelerating work to overcome the global impact of neglected tropical diseases : a roadmap for implementation : executive summary, World Health Organization, 2012 (2012).
URL http://www.who.int/iris/handle/10665/70809

[20] K. E. Halliday, W. E. Oswald, C. Mcharo, E. Beaumont, P. M. Gichuki, S. Kepha, S. S. Witek-McManus, S. H. Matendechero, H. El-Busaidy, R. Muendo, et al., Community-level epidemiology of soil-transmitted helminths in the context of school-based deworming: Baseline results of a cluster randomised trial on the coast of kenya, PLoS neglected tropical diseases 13 (8) (2019) e0007427 (2019).

[21] K. H. Ásbjörnsdóttir, S. S. R. Ajjampur, R. M. Anderson, R. Bailey, I. Gardiner, K. E. Halliday, M. Ibikounle, K. Kalua, G. Kang, D. T. J. Littlewood, A. J. F. Luty, A. R. Means, W. Oswald, R. L. Pullan, R. Sarkar, F. Schr, A. Szpiro, J. E. Truscott, M. Werkman, E. Yard, J. L. Walson, T. D. T. Team, Assessing the feasibility of interrupting the transmission of soil-transmitted helminths through mass drug administration: The deworm3 cluster randomized trial protocol, PLOS Neglected Tropical Diseases 12 (1) (2018) 1–16 (01 2018). doi:10.1371/journal.pntd.0006166.
URL https://doi.org/10.1371/journal.pntd.0006166

[22] D. A. Raichlen, B. M. Wood, A. D. Gordon, A. Z. Mabulla, F. W. Marlowe, H. Pontzer, Evidence of lévy walk foraging patterns in human hunter–gatherers, Proceedings of the National Academy of Sciences 111 (2) (2014) 728–733 (2014).

[23] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, S. Chong, On the levy-walk nature of human mobility, IEEE/ACM transactions on networking (TON) 19 (3) (2011) 630–643 (2011).

[24] K. S. Kung, K. Greco, S. Sobolevsky, C. Ratti, Exploring universal patterns in human home-work commuting from mobile phone data, PloS one 9 (6) (2014) e96180 (2014).

[25] M. Lenormand, B. Gonçalves, A. Tugores, J. J. Ramasco, Human diffusion and city influence, Journal of The Royal Society Interface 12 (109) (2015) 20150473 (2015).

[26] W. S.-t. Helminthiases, Eliminating soil-transmitted helminthiases as a public health problem in children: progress report 2001–2010 and strategic plan 2011–2020, France: World Health Organization (2012) 19–29 (2012).

[27] High Resolution Settlement Layer (HRSL), Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University, 2016, Source imagery for HRSL ©2016 Digital-

Globe, Accessed (31/10/2019).

[28] C. D. Criscione, J. D. Anderson, D. Sudimack, J. Subedi, R. P. Upadhayay, B. Jha, K. D. Williams, S. Williams-Blangero, T. J. Anderson, Landscape genetics reveals focal transmission of a human macroparasite, PLoS Neglected Tropical Diseases 4 (4) (2010) e665 (2010).

[29] L. Steinbaum, L. H. Kwong, A. Ercumen, M. S. Negash, A. J. Lovely, S. M. Njenga, A. B. Boehm, A. J. Pickering, K. L. Nelson, Detecting and enumerating soil-transmitted helminth eggs in soil: New method development and results from field testing in kenya and bangladesh, PLoS neglected tropical diseases 11 (4) (2017) e0005522 (2017).

[30] J. Truscott, H. Turner, S. Farrell, R. Anderson, Chapter three - soil-transmitted helminths: Mathematical models of transmission, the impact of mass drug administration and transmission elimination criteria, in: M. G. Basez, R. M. Anderson (Eds.), Mathematical Models for Neglected Tropical Diseases, Vol. 94 of Advances in Parasitology, Academic Press, 2016, pp. 133 – 198 (2016). doi:https://doi.org/10.1016/bs.apar.2016.08.002.
URL http://www.sciencedirect.com/science/article/pii/S0065308X1630077X

[31] C. Cosner, J. C. Beier, R. S. Cantrell, D. Impoinvil, L. Kapitanski, M. D. Potts, A. Troyo, S. Ruan, The effects of human movement on the persistence of vector-borne diseases, Journal of theoretical biology 258 (4) (2009) 550–560 (2009).

[32] D. T. Citron, D. L. Smith, S. L. Wu, H. M. Sanchez Castellanos, A. J. Dolgert, J. M. Henry, C. A. Guerra, Comparing metapopulation dynamics of infectious diseases under different models of human movement, medRxiv (2020). arXiv:https://www.medrxiv.org/content/early/2020/04/07/2020.04.05.20054304.full.pdf, doi:10.1101/2020.04.05.20054304.
URL https://www.medrxiv.org/content/early/2020/04/07/2020.04.05.20054304