Author for correspondence:
T. Alex Perkins
e-mail: taperkins@nd.edu

†Present address: Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, USA.

Royal Society Publishing

Theory and data for simulating fine-scale human movement in an urban environment

T. Alex Perkins[1,2,†], Andres J. Garcia[3,4], Valerie A. Paz-Soldán[5], Steven T. Stoddard[2], Robert C. Reiner Jr[1,2], Gonzalo Vazquez-Prokopec[6], Donal Bisanzio[6], Amy C. Morrison[2], Eric S. Halsey[7], Tadeusz J. Kochel[7], David L. Smith[1,8,9], Uriel Kitron[1,6], Thomas W. Scott[1,2] and Andrew J. Tatem[1,10,11]

[1]Fogarty International Center, National Institutes of Health, Bethesda, MD, USA
[2]Department of Entomology and Nematology, University of California, Davis, CA, USA
[3]Emerging Pathogens Institute, and [4]Department of Geography, University of Florida, Gainesville, FL, USA
[5]Department of Global Health Systems and Development, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
[6]Department of Environmental Sciences, Emory University, Atlanta, GA, USA
[7]United States Naval Medical Research Unit No. 6, Lima, Peru
[8]Department of Epidemiology, and [9]Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[10]Department of Geography and Environment, University of Southampton, Southampton, UK
[11]Flowminder Foundation, Stockholm, Sweden

Individual-based models of infectious disease transmission depend on accurate quantification of fine-scale patterns of human movement. Existing models of movement either pertain to overly coarse scales, simulate some aspects of movement but not others, or were designed specifically for populations in developed countries. Here, we propose a generalizable framework for simulating the locations that an individual visits, time allocation across those locations, and population-level variation therein. As a case study, we fit alternative models for each of five aspects of movement (number, distance from home and types of locations visited; frequency and duration of visits) to interview data from 157 residents of the city of Iquitos, Peru. Comparison of alternative models showed that location type and distance from home were significant determinants of the locations that individuals visited and how much time they spent there. We also found that for most locations, residents of two neighbourhoods displayed indistinguishable preferences for visiting locations at various distances, despite differing distributions of locations around those neighbourhoods. Finally, simulated patterns of time allocation matched the interview data in a number of ways, suggesting that our framework constitutes a sound basis for simulating fine-scale movement and for investigating factors that influence it.

1. Introduction

The importance of mathematical modelling of human movement is far ranging, as movement patterns underlie key issues in public health, economics and urban planning [1,2]. Different questions in each of these arenas, however, necessitate the selection—and sometimes the development—of models with different levels of spatial and temporal resolution.

In infectious disease epidemiology, describing average flows of movement between cities is informative of the spread and persistence of pathogens over broad geographical areas [3–5]. At more local scales, such as within a city,

movements by smaller groups must be acknowledged to capture structure relevant to the dynamics of transmission [6–9] and the uniqueness of exposures of those different groups [10]. Different types of movement over different time scales must also be considered. Migratory and seasonal movements are likely important for circulation over broad regions and time scales, whereas commuting patterns and other routine movements give rise to the structure of contacts on which epidemics spread [1,11].

For modelling movement at fine spatial scales, such as within cities, many models (e.g. [12–14]) focus on scales no finer than several hundred square metres, or the area covered by a mobile phone tower, because that represents the finest scale at which an individual's presence can be deduced with mobile phone call records. In certain locations, however, data informative of an individual's whereabouts at finer scales could be used to motivate and parametrize models at the scale of buildings or lots [15,16]. An existing conceptual framework for modelling movement at this scale is that of an individual's activity space, which is defined as 'the subset of all urban locations with which the individual has direct contact as the result of day-to-day activities' [17].

Modelling the composition of activity spaces and movement within them is essential for any simulation of synthetic human populations in an urban environment. Such simulations are of high utility for a number of applications in infectious disease epidemiology, including planning for the containment of influenza or smallpox outbreaks [18–21] and evaluating the efficacy of a putative dengue vaccine [22,23]. Although these and other applications are of relevance to cities in the developing world, the development of algorithms for simulating the composition and dynamics of activity spaces has focused primarily on cities in North America and Europe. Furthermore, simulation models of human activity spaces typically focus on details that, while important for applications in transportation [24], may be unnecessarily complex for application to the epidemiology of many infectious diseases [15]. Still other models are capable of simulating movement and time allocation within an activity space (e.g. [25,26]), but provide no basis for simulating which locations comprise the activity space. A generalizable framework that can be parametrized with readily attainable data and that can be used to simulate activity space composition and time allocation in a variety of geographical contexts is therefore needed.

To address this need, we developed a modelling framework that integrates five distinct aspects of movement—i.e. number of locations in the activity space, location type and distance from home of locations in the activity space, and frequency and duration of visits to those locations—to generate a cohesive description of time allocation within an individual's activity space. This framework allows for simulation of locations comprising the activity space and parameters that govern a stochastic process of movement between pairs of locations within the activity space. Together, this results in a description of how the individual allocates time across the activity space. To demonstrate the utility of this framework, we fit a model of activity space composition and time allocation to data from retrospective interviews of 157 residents of the city of Iquitos in northeastern Peru. Using these data, we selected among candidate models with varying levels of detail about location type and distance from home, and we used simulations of the best-supported model (e.g. figure 1) to assess the model's ability to reproduce empirical patterns of time allocation in the study population.

2. Material and methods

2.1. Modelling framework

2.1.1. Activity space composition

The first mathematical characteristic of an individual's activity space that we define is that it is comprised of locations belonging to m different classes, each of which is distinguished by how long locations of that class tend to remain in the individual's activity space. For example, many people likely have some locations that they visit routinely (e.g. relatives' houses) as well as some that they do not visit as a matter of routine (e.g. repair shops, airports). We posit that routine locations remain in one's activity space for long periods of time, whereas irregular locations come and go from the individual's activity space over time. Each such class i is defined by a constant rate λ_i at which new locations are added to it and a constant rate μ_i at which a location of that class is removed from the activity space (figure 2a). If we assume that locations in each class are removed from the activity space in the same order in which they were added to it, then each class within an individual's activity space can be modelled as a queue (of M/M/1 type [27]). This convenience means that we can directly calculate some key characteristics of the dynamics of each class, including the stationary distribution of the number of locations of each class in the activity space (geometric with parameter $\rho_i = 1 - \lambda_i/\mu_i$) [27, pp. 548–552]. The stationary distribution of the number of locations across all classes in the activity space is thus a sum of m geometric random variables. In the event that m is finite, the stationary distribution of activity space size follows a negative binomial distribution with parameters m and ρ if all $\rho_i = \rho$. If there is a very large number of classes (i.e. as $m \to \infty$), the stationary distribution of activity space size is a Poisson distribution with parameter $m\rho$. Finally, in the trivial case with only a single class, the stationary distribution of activity space size is geometrically distributed with parameter ρ.

2.1.2. Time allocation within the activity space

Because there is potential for locations to be added to or removed from the activity space at any time, we complement the continuous-time process of activity space turnover with a continuous-time process of movement by an individual within her or his activity space. A simple and general way to model this movement process is with a continuous-time, finite-state Markov process (figure 2b), similar to how geographers have modelled trip behaviour [28,29]. Under such a formulation, the set of n locations in an individual's activity space comprise the states among which an individual moves about according to the $n \times n$ rate matrix Q [27, p. 396]. Each off-diagonal element $q_{i,j}$ of Q represents the instantaneous probability that an individual moves from location i to location j, and the elements of this matrix are stipulated to satisfy the condition that $q_{i,i} = -\sum_{j \neq i} q_{i,j}$. Characteristics of a movement process obeying these dynamics include that the durations of visits to each location i are exponentially distributed with mean $-1/q_{i,i}$ and that the stationary distribution π, which is a vector containing the long-term average proportion of one's time spent at each location, satisfies $\pi Q = 0$ [27, pp. 395 and 398].

2.2. Model refinements

Applying this framework to specific populations requires specifying which locations are included in an individual's activity space and how an individual spends time at and moves among those locations. Here, we investigated two primary characteristics of locations—location type and distance from home—that affect whether individuals visit those locations and if so how frequently and for how long they visit (table 1).
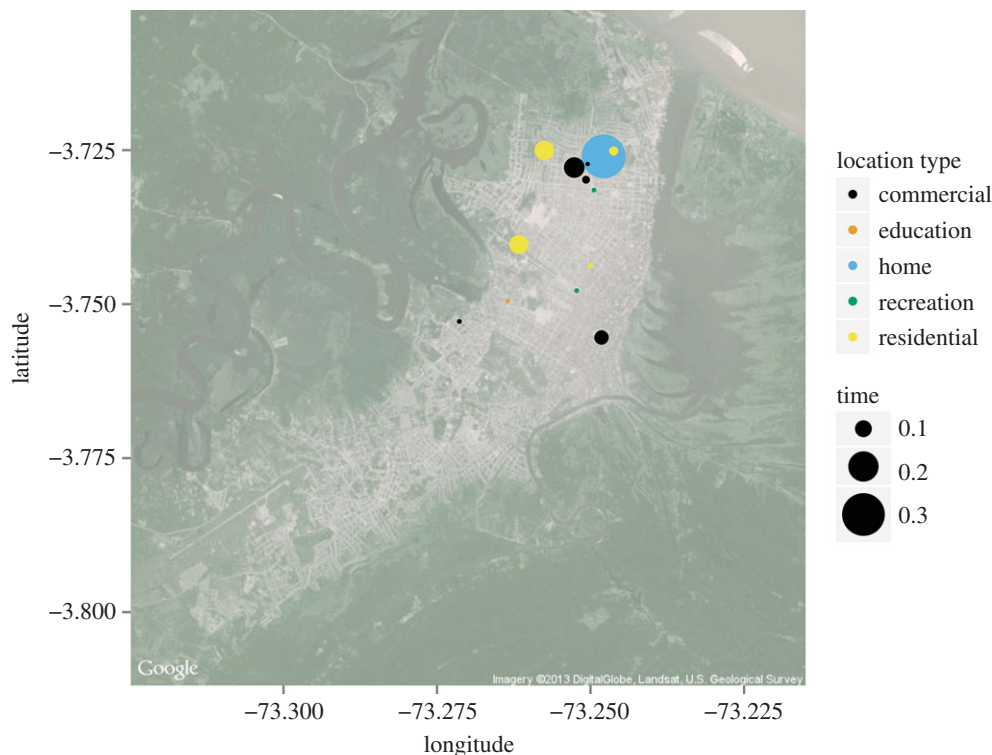
Figure 1. Example of an individual's activity space and the proportion of the individual's time spent at each location, simulated with the fitted model. (Online version in colour.)
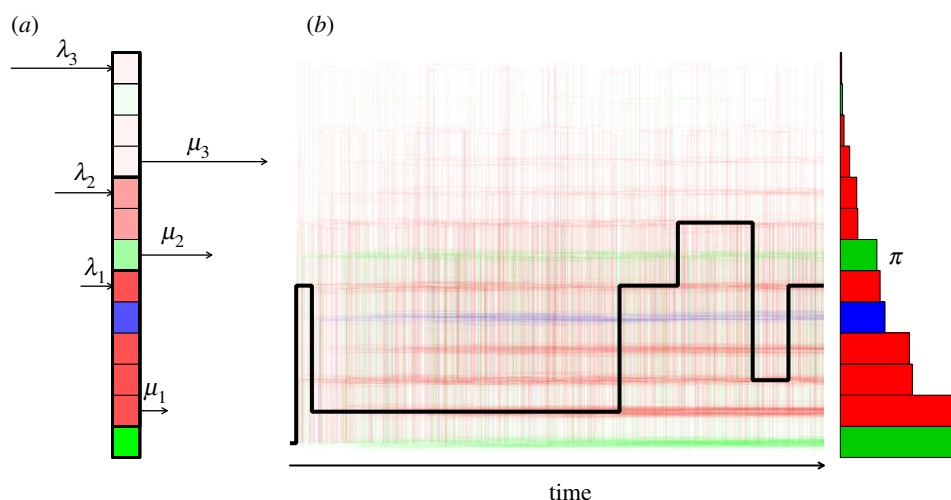

Figure 2. Model schematic. (a) Snapshot of an individual's activity space at a given point in time, with the individual's home on the bottom, a relatively permanent class above it, an intermediate class above that, and a relatively transient class at the top, with classes separated by thick lines. The composition of each class is governed by an M/M/1 queue with 'birth' rate λ_i and 'death' rate μ_i. Different colours indicate different location *types* (e.g. houses, shops, parks), which are distinct from the permanent, intermediate and transient location *classes*. (b) An individual's whereabouts over time, with a single example in black and 500 replicates in faint colours. On average, the overall proportion of time spent at each location tends to the distribution π. (Online version in colour.)

2.2.1. Location types

Given some way of classifying locations according to T types, such as residential or commercial, we propose that whenever a new location is added to one's activity space it has a probability p_τ of being a location of type τ, where $\sum p_\tau = 1$. Consequently, the number of locations of each type in an activity space comprised of n locations is a multinomial random variable with parameters n and $\boldsymbol{p} = \{p_{\tau_1}, \ldots, p_{\tau_T}\}$. These location types are not necessarily the same as the location classes. Rather, we envision location types as readily distinguishable based on observable characteristics (e.g. houses, schools) and location classes as defined solely by the extent of their transience in

one's activity space (e.g. a workplace and a market could belong to the same class).

2.2.2. Distance from home

We also consider the possibility that individuals tend to display a preference for locations that are closer to their homes. To investigate this possibility, we must first consider how far away locations of a certain type are from one's home. If an individual chooses locations randomly with respect to distance from home, then distances at which more locations of a given type are present would be more likely to be chosen. Thus, we must first

Table 1. Summary of model components, model subcomponents and how variation in each is represented probabilistically in the best-supported model.

model component	model subcomponent	probability distribution
locations visited	number of locations	negative binomial (r, ρ)
	locations of a given type	multinomial (p_τ)
	locations of a given distance	$\propto \text{pdf}(\delta; \tau) \cdot \exp(-\mu_\tau \delta^{\eta_\tau})$
time allocation	frequency of visits	lognormal $(\mu(\delta; a_{d,\tau}, b_{d,\tau}, c_{d,\tau}), \sigma_{d,\tau})$
	duration of visits	lognormal $(\mu(\delta; a_{f,\tau}, b_{f,\tau}, c_{f,\tau}), \sigma_{f,\tau})$

account for the distribution $p(\delta; \tau)$ of locations of type τ at various distances δ and then determine whether the realized distribution of locations visited is weighted in favour of those closer to home. We do so with a weighting function of the form

$$\exp(-\mu_\tau \delta^{\eta_\tau}), \tag{2.1}$$

which is similar to an exponential distribution with rate μ but with an additional parameter η to allow for more flexibility to weight locations very close to home more strongly. Once the distance from home of a new location is determined, we select a specific location randomly from all locations of a given type τ at that distance. We furthermore consider the possibility that different location types have different weighting functions, as determined by the parameters μ_τ and η_τ.

2.2.3. Time allocation

To determine how individuals allocate time across locations in their activity spaces, we assign each individual a frequency of visits, f_i, and a mean duration of each visit, d_i, to each location i in the activity space. The collection of all f_i and d_i for $i \in 1, \ldots, n$ is then used to populate the matrix Q. Because the diagonal entries of Q are directly related to the mean duration of visits to each location, these entries are defined simply as $q_{i,i} = -1/d_i$. To populate the off-diagonal entries of Q, which describe movements between locations, we herein make the simplifying assumption that consecutive movements are uncorrelated and proportional to the frequency at which an individual visits each location, such that

$$q_{i,j} = \frac{1}{d_i} \frac{f_j}{\sum_{j \neq i} f_j}. \tag{2.2}$$

To select f_i and d_i for different individual–location pairs, we model these parameters as bivariate normal random variables on a natural log scale with means $\mu_{f,\tau}$ and $\mu_{d,\tau}$, standard deviations $\sigma_{f,\tau}$ and $\sigma_{d,\tau}$ and correlation ρ_τ. Moreover, we consider the possibility that the means and standard deviations differ for different location types and that the means depend on distance from home according to the function

$$\mu(\delta) = \frac{1}{2}a + \frac{ab \exp(c\delta)}{1 + \exp(c\delta)}. \tag{2.3}$$

This function is defined such that the mean frequency or duration of visits on a natural log scale to a location of a given type τ approaches $(1/2)a(1 + b)$ near home and $(1/2)a$ as $c\delta \to -\infty$. Finally, we assume that the frequency and mean duration of visits to one's own home and to locations outside the city are also jointly distributed lognormal random variables but that they are characterized by their own parameters, which do not depend on δ.

2.3. Data

The data used to fit the model were collected in the Amazonian city of Iquitos, Peru. Iquitos is an isolated city of approximately 377 000 inhabitants surrounded on three sides by rivers. Only one regional road leaves the city, meaning that most local

travel outside the city takes place by boat. Besides walking short distances, movement within the city typically occurs by motorcycle, motorcar or bus. Neighbourhoods within the city vary in the availability of services and quality of construction. Most contain key services (schools, health centres, markets), but specialized commerce and other services are concentrated in specific areas of the city (e.g. two main hospitals, a shopping district in the downtown area). Participants in this study lived in either of two neighbourhoods (electronic supplementary material, figure S1), Maynas and Tupac Amaru, which exhibit modest differences in housing construction and population density [30]. These neighbourhoods were initially chosen for other studies because they are mostly self-contained (having schools and health centers) and are distant enough to limit inter-neighbourhood movement [9,30]. Historically, dengue virus transmission and mosquito densities have been higher in Maynas [31].

2.3.1. Geographical information system

Construction of a geographical information system (GIS) for the city of Iquitos began in 1998 [32] and was ongoing at the time of this study [30,31,33]. We used coverages that encompassed nearly 50 000 lots using the coordinate system Universal Transverse Mercator WGS-84 Zone 18S. In areas with ongoing epidemiological and entomological research projects, field personnel assigned location types to each lot. It should be noted though that many sites were mixed use; e.g. residential and commercial in many cases. During the retrospective interview study, unmapped locations were physically located to obtain Geographical Positioning System coordinates and to update the GIS.

2.3.2. Retrospective interviews

Retrospective interviews were conducted using a semi-structured interview (SSI) [34,35]. The SSI was developed to address the issues of recall, reliability, reproducibility, compliance, behavioural change and privacy that are typically associated with classic movement survey methods, such as interviews, diaries or direct observation. This tool was designed for use with people who became infected with dengue or who were in shared spaces with people who had become infected with dengue (potential future dengue cases), to find out where they had been in the past two weeks (to identify potential exposure sites). Hence, we were limited to methods that could be applied retrospectively. Based on focus groups conducted to develop the SSI, we found that: (i) people could identify routine locations they visit, but triggers were needed for certain types of locations, (ii) the best aid for recalling locations visited was to begin the SSI by thinking about a 'typical day' and to gradually add memory triggers over the course of the interview, and (iii) there were clear 'common activity spaces' identified for all. During the development and validation of the SSI, we found that high-resolution satellite imagery of participants' neighbourhoods, combined with street labels, was not very useful in triggering recall of locations, but aided participants in describing and locating where they had been. If participants were unable to find a location on the map, they would either call someone at the location to ask for the address, or give our research

Table 2. Summary of candidate models for each model subcomponent.

model subcomponent	candidate models
number of locations	geometric, Poisson, negative binomial
locations of a given type	all possible location type groupings in which types within a group have identical probabilities of being found in an activity space
locations of a given distance	all possible location type groupings in which types within a group have identical effects of distance from home, which depend on either one or two parameters per location type grouping
frequency of visits	all possible location type groupings in which types within a group have identical effects of distance from home or no effect of distance from home. Zero or non-zero correlation with duration of visits
duration of visits	all possible location type groupings in which types within a group have identical effects of distance from home or no effect of distance from home. Zero or non-zero correlation with frequency of visits
time at home	zero or non-zero correlation between frequency and duration of times at home

team the best description they could. Our research team would then find the location and record its position. The SSI followed with a section on common activity spaces already identified by others (e.g. schools, markets, health facilities). The SSI concluded with a section that listed categories of locations that tended to require triggers to be recalled or else would otherwise likely go unreported. The result of each interview was a list of the locations that an individual visited in the two-week period preceding the interview, as well as estimates of the frequency and duration of visits to each location.

A total of 120 participants provided information about time spent at home, 138 about time spent elsewhere and 101 of 157 provided information about both. These interviews were conducted with residents of two neighbourhoods (electronic supplementary material, figure S3), from which study participants were obtained by convenience sampling and are thus not completely representative of the local population in terms of age, sex, occupation and possibly other factors; they are nonetheless diverse (electronic supplementary material, figures S2–S4). Interviews were mostly conducted during the dengue transmission season (electronic supplementary material, figure S5), which spans several months, as well as variation in seasons that could impact movement behaviour (e.g. times when children are in or out of school).

2.4. Analysis
2.4.1. Model fitting and selection
The first aim of our analysis was to fit and select among candidate models for each of the five aspects of movement. For example, candidate models that we considered for the distribution of the number of locations visited included geometric, Poisson and negative binomial distributions. Candidate models for all aspects of movement are listed in table 2. For each aspect of movement, each candidate model was fitted separately to the interview data by numerically estimating maximum-likelihood parameter values. To compare a pair of models with nested parameters, we performed a likelihood ratio test and selected the more complex model if $p < 0.05$. For situations with more than two candidate models, we assessed the relative support for each by computing each model's Akaike information criterion corrected for finite sample size (AIC_c), which balances goodness of fit and model complexity, and then Akaike weights, which are a measure of relative support [36].

To explore the set of possible models with different levels of detail about location type, we employed a form of backward elimination. To do so, we first fitted the candidate model with the finest breakdown of location types that we considered, such that each location type had its own set of parameters. We then chose the pair of location types with the most similar parameter values and fitted a new model in which the parameters of those two location types were constrained to be equal. Repeating this procedure of agglomerating location types based on parameter similarity, we obtained a set of candidate models with a range of location-type categorizations, from the case in which each location type had its own parameter values to the case in which all location types had the same parameter values. Additional details of the model selection procedure specific to different aspects of movement are elaborated on in the electronic supplementary material, table S2.

For all aspects of movement, we also fitted models separately to interviews of residents of two distinct neighbourhoods in Iquitos to assess the robustness of our fitted model to possible neighbourhood-specific differences (which have been found elsewhere [37]). In doing so, we applied the same model selection procedures as described above, and we assessed support for either the aggregated model (denoted M_{M+T}) or the disaggregated model ($M_M + M_T$) by $\Delta AIC_c > 10$ (as recommended for a pair of non-nested models [36, p. 123]).

2.4.2. Comparison of model outputs against data
To assess the realism of patterns of time allocation simulated with our model, we compared simulation results against patterns of time allocation derived directly from the retrospective interviews. Although data from these interviews were also used to parametrize the model, the model's ability to reproduce patterns of individual time allocation does not necessarily follow, because the features of the interview data with which the model was parametrized are distinct from those against which model outputs were compared. This comparison therefore allowed us to assess the model's ability to translate basic aspects of movement into descriptions of time allocation across locations, which is the objective of the model and a goal that is common to many applications in epidemiology and other fields.

To that end, we first calculated the empirical pattern of time allocation within each individual's activity space. For the same number of individuals as participated in retrospective interviews of time spent at home and elsewhere ($n = 101$), we simulated patterns of time allocation 10^3 times (e.g. figure 1). Given these empirical and simulated patterns of time allocation, we then examined (i) the proportion of individuals that allocated a certain proportion of their time at a single location of each location type as well as at their home and (ii) how time allocation was distributed over distance from home.

Our quantitative approach to the comparison of simulated and empirical patterns of time allocation was based on that of statistical hypothesis testing. For a given statistic (e.g. proportion of time
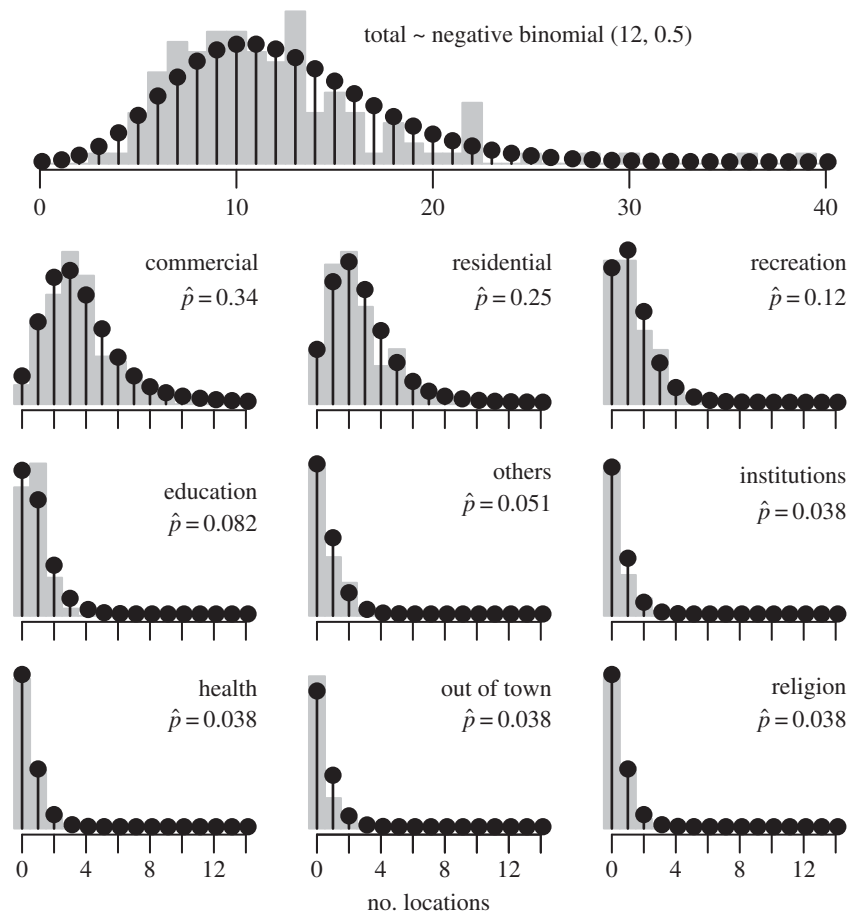
Figure 3. Numbers of locations in individuals' activity spaces, in total (top) and by location type (other panels). Grey bars show empirical data from retrospective interviews ($n = 138$). Black stems show a fitted negative binomial distribution (top) and a fitted multinomial distribution conditioned on the empirical distribution of total numbers of locations and marginalized for each location type (other panels).

allocated at a distance of 1 km from home), the set of simulated values of that statistic comprised the null distribution against which the empirical value was compared. Formally, and assuming a two-tailed test, if the empirical value of a given statistic fell between the 2.5th and 97.5th percentiles of this null distribution, then our interpretation was that we could not reject the possibility that the empirical value was generated by the process codified by our model. Informally, given that any single statistic that we considered fell along a continuum of related statistics (e.g. distances from home of 100 m, 200 m, etc.), we found it instructive to visually compare how well the entire set of empirical statistics tracked the set of simulated statistics.

3. Results

3.1. Model fitting and selection

3.1.1. Activity space composition

Among the three distributions that we considered, we found clear support for the negative binomial (Akaike weight ≈ 1, table S3, figure 3). This result is consistent with the hypothesis that activity spaces comprised $m = 12$ classes each with identical parameter $\rho = 0.4997$, or, alternatively, that activity space size was Poisson distributed with variation among individuals' parameters that was gamma distributed with hyperparameters equal to r and $\rho/(1 - \rho)$.

We found support for uneven representation of location types within individuals' activity spaces, with the best-supported model being those in which there were six distinct

location-type groups, and with models with between five and nine such groups having nearly all Akaike weight (table S4). The best-supported model assigned 34% of locations as commercial, 25% as residential, 12% as recreational and others at less than 10% each (figure 3). All models with some fitted vector p fit the data better than a model in which p equalled the proportions of locations of each type within the city (table S4), which represented the hypothesis that locations were chosen irrespective of type.

For each location type that we considered, there was a distinct effect of distance from home on its inclusion in one's activity space, as the model with separate parameters for each location type had an AIC_c value 49 lower than the next best model (table S5). For locations in the recreation or others categories, individuals were more likely to visit locations of an intermediate distance from home (approx. $0.5 - 2$ km) than they were locations nearby or far away (figure 4). For all other location types, individuals were less likely to visit locations farther from home, with this effect being particularly strong for locations in the residential, commercial, education and institutions categories (figure 4). These different relationships for different location types are likely attributable to aspatial considerations, given that we accounted for the spatial distribution of each location type relative to study participants' homes.

3.1.2. Time allocation

The frequency and mean duration of visits to home were best described by a bivariate lognormal distribution with
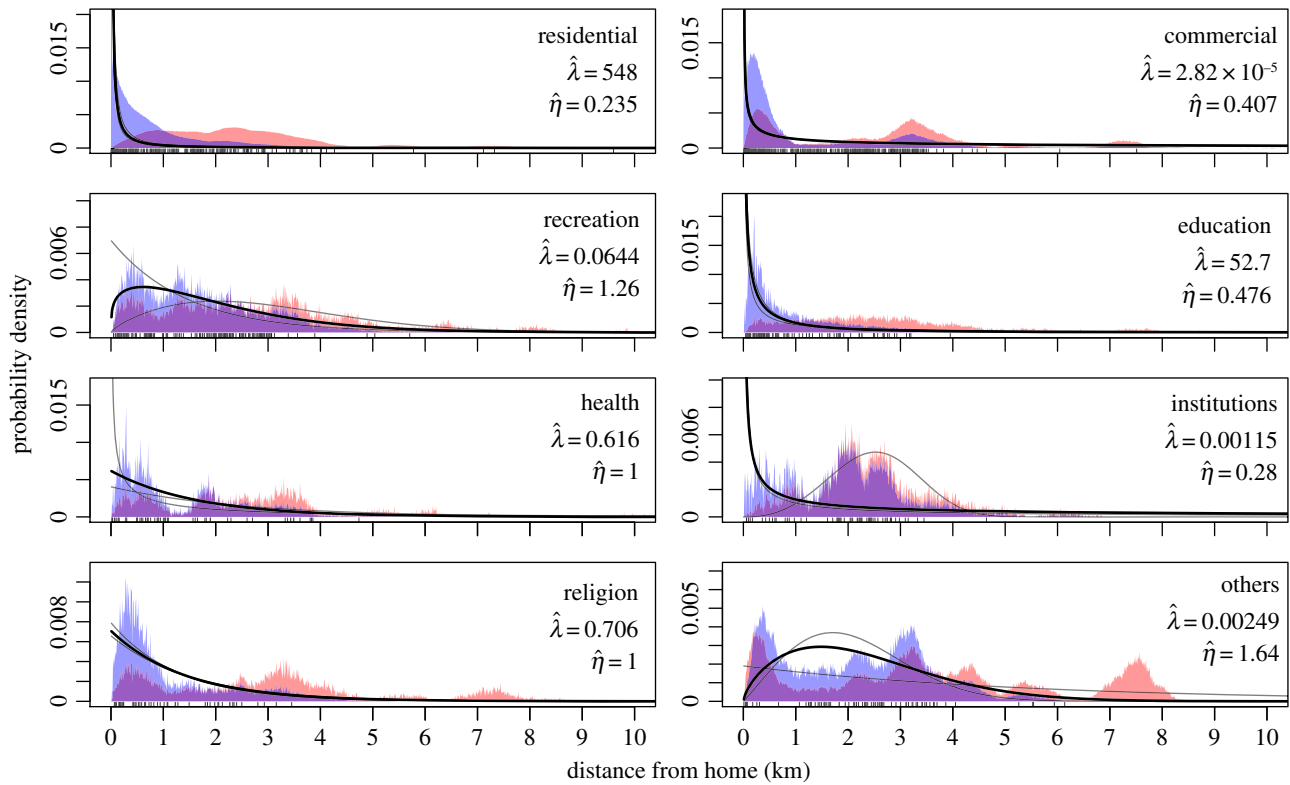
Figure 4. Relationship between distance from home and the probability that a location of a given type is chosen for one's activity space. The probability density of distances from home at which activity space locations of a given type are located (blue area) is obtained by weighting the probability density of distances from home at which all locations of a given type are located (red area) by a function for how distance from home affects the probability of being chosen for one's activity space (curves). The thick black curves (whose parameter values are listed in each panel) were fitted to all interview data without consideration of the neighbourhood in which interviewees live, whereas the thin grey curves were fitted separately to interviews from residents of the Maynas and Tupac neighbourhoods. Rug plots along the bottom of each panel indicate the distances from home at which locations in the activity spaces of study participants were located. (Online version in colour.)

an average frequency of 2.79 times per day, an average mean duration of 4.2 h per visit (between the hours of 5.00 and 22.00), and a significant negative correlation between those quantities ($\rho = -0.59$; likelihood ratio test: $\chi_1^2 = 51.8$, $p = 6.2e - 13$) (figure 5).

For locations other than home, one attribute that we found to affect both the frequency and mean duration of visits was location type. Comparing models of the frequency and mean duration of visits with different numbers of location-type groups, we found strong support for the model with the most detailed representation of location types (Akaike weight = 0.99, table S6). For each of those location types, we also found differences in how frequencies and mean durations of visits to locations of those types were affected by distance from home and by a correlation between those quantities. Distance from home had a significant effect only for residential and commercial location types (Akaike weights ≈ 1, table S7), with the frequency and mean duration of visits decreasing and increasing, respectively, with increasing distance from home. Statistical support for a correlation between the frequency and mean duration of visits to locations of each type was only evident for the education and institutions location types (Akaike weights > 0.95, electronic supplementary material, table S6), which had correlations of 0.27 and 0.49, respectively (electronic supplementary material, table S8).

The frequency and mean duration of visits outside the city were best described by a bivariate lognormal distribution with an average frequency of once per 6.4 days, an average mean duration of 4.4 h, and no correlation between those quantities (likelihood ratio test: $\chi_1^2 = 2.56$, $p = 0.11$) (figure 5).

3.1.3. Neighbourhood comparison

Applying our model fitting and selection procedures separately to interviews from residents of two distinct neighbourhoods, we found strong statistical support for the disaggregated neighbourhood model ($M_M + M_T$) over the aggregated model ($M_{M + T}$) only for the model subcomponent concerning the effect of distance from home on inclusion in the activity space ($\Delta \text{AIC}_c = 59.24$, table S9). Performing a similar comparison of $M_M + M_T$ and $M_{M + T}$ by location type (table S10), we found statistical support for differences in the effect of distance from home only for the recreation ($\Delta \text{AIC}_c = 34.86$), institutions ($\Delta \text{AIC}_c = 11.66$) and others ($\Delta \text{AIC}_c = 13.64$) location types (figure 4).

3.2. Comparison of model outputs against data

Both empirical and simulated data displayed wide variation in the proportion of time that individuals spent at home (figure 6). However, the fitted model tended to predict that very few people spent a majority of their time at home. For locations other than home, simulated patterns of time allocation were consistent with the empirical pattern that most people spent relatively little time at any single location (i.e. 10% or less). The most significant departures from this pattern in the empirical data were that approximately 10% of study participants who visited an educational location or a location outside of Iquitos spent significantly more time at those locations (approx. 20–30%) than was exhibited in the simulated data. For location types that were relatively infrequently included in one's activity space (i.e. health, institutions, religion, others), the empirical pattern of very few
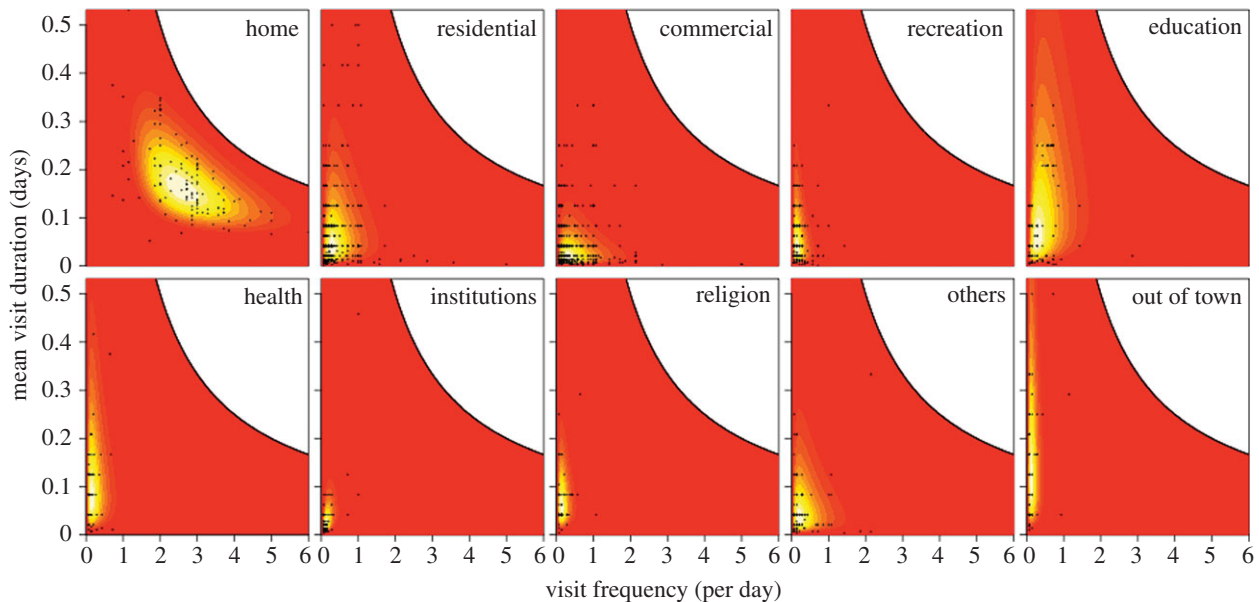
Figure 5. Joint distributions of the frequency and mean duration of visits to one's own home, to locations of various types within Iquitos, and to locations outside the city. Data from retrospective interviews (black dots) are superimposed on bivariate lognormal probability density surfaces (relative density: red, low; orange, intermediate; yellow, high) with maximum-likelihood parameter estimates. Note that (i) surfaces for residential and commercial location types reflect maximum-likelihood parameter estimates at a distance of 100 m from home and that other distances have somewhat different surfaces and (ii) it is not possible for data to occupy the white region above the black curve, which represents the combinations of frequencies and mean durations that result in a person spending all of their time at a single location. (Online version in colour.)
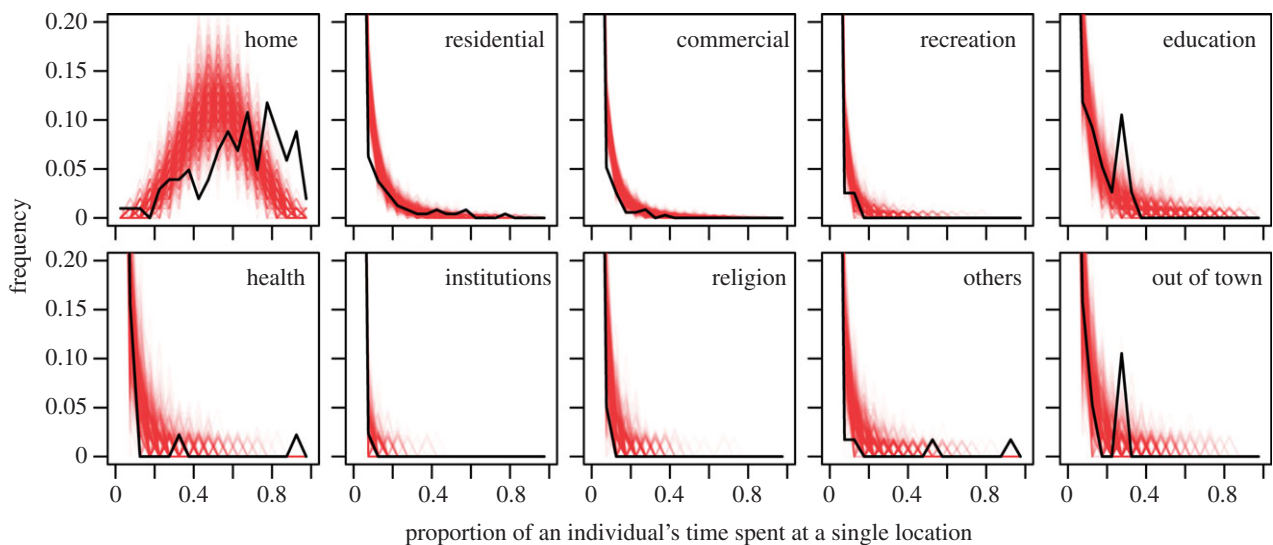
Figure 6. Distributions of the overall proportion of an individual's time spent at a single location of each of several location types. Black lines show the proportion of participant – location pairs in the retrospective interview study for which a given proportion of the participant's time was spent at that location. Red lines show 10^3 analogous distributions resulting from simulating activity spaces, simulating movement parameters within those activity spaces, and calculating π for the same 101 individuals that participated in both parts of the retrospective interview study. (Online version in colour.)

individuals spending a significant amount of their time at a single location of one of those types was consistent with individual replicates of the simulation (figure 6).

Simulated data also captured some patterns of time allocation over a range of distances from home but were deficient in other ways. One of the primary deficiencies of the simulated data was that they under-predicted the proportion of time allocated within 100 m of an individual's home (figure 7). That time was instead allocated elsewhere, thereby shifting the simulated patterns of time allocation at distances beyond 100 m above the empirical pattern at those distances. The simulated data captured the mean pattern of time allocation between distances of approximately 100 m to 5 km

relatively well otherwise, but they appeared to under-represent the variability in time allocation at different distances (i.e. the simulated patterns were relatively smooth compared with the empirical pattern). The simulations also somewhat over-predicted time allocated beyond 5 km, but in individual replicates this discrepancy was not as severe as it appeared in the ensemble. Examining the mean distance from home of where time was allocated (figure 7, green) and the mean distance of locations visited (figure 7, blue), it appeared that over-predicting the distance from home at which time was allocated resulted from both over-predicting the distance of locations that were visited and over-predicting the frequency and duration of visits to faraway locations.
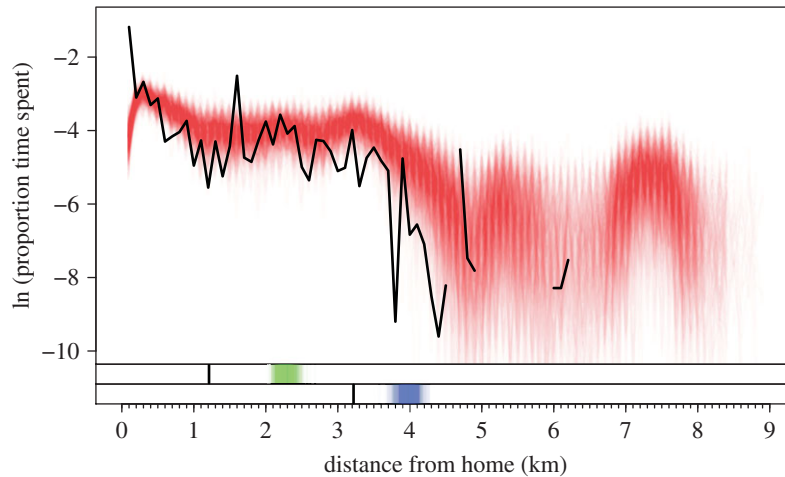
Figure 7. Distribution of time allocation over distance from home, as measured by the retrospective interviews (black) and in 10^3 simulations of the fitted model (red). The bar directly below the main panel shows the mean distance at which time is allocated for the empirical (black) and each simulated (green) distribution. The bar along the bottom shows the mean distance of locations visited in the empirical (black) and each simulated (blue) activity space. (Online version in colour.)

4. Discussion

We developed a new mathematical framework for simulating movement of individual people at fine spatial scales in an urban environment. This framework has advantages over existing alternatives in that it is specified at the scale of individual buildings or lots, it allows for *a priori* simulation of both where individuals go and how they allocate their time, and it can be fitted to a variety of data sources about human movement at fine scales. It is also more general than other simulation models of fine-scale urban movement that are structurally sensitive to details of specific locations. Instead, our framework represents a minimal model for realistically simulating patterns of individual time allocation, yet it is based on conventional mathematical tools that make it readily extensible for various applications where additional detail is warranted.

As a case study for this new framework, we fitted candidate versions of such a model, each with a different level of detail, to a set of 157 interviews of residents of Iquitos, Peru. In doing so, we found that location type and distance from study participants' homes were important factors in determining which locations comprised their activity spaces and how they allocated time across those locations. Simulations of the best-supported model provided a realistic description of aggregate patterns of time allocation, thereby validating the utility of this framework for its intended purpose of realistically simulating activity spaces and time allocation therein for a synthetic population. The attainment of this best-supported model thus represents a significant advance in realistically simulating fine-scale human movement in Iquitos, where empirical and modelling studies of dengue virus epidemiology are ongoing, and a starting point for simulating fine-scale human movement in other resource-poor urban environments. The availability of our code at https://github.com/TAlexPerkins/ActivitySpace_JRSI2014 should facilitate adaptation of the model to other cities and extension of the model around additional complexities that other datasets may be well suited to address.

activity space turn over at different rates. Currently, we are unable to examine this process of turnover empirically, because we do not possess data appropriate for doing so. Ideally, data to address this question would consist of records of when individuals started and stopped visiting locations over a span of months or even years. Nonetheless, positing such a process provides us with a set of first principles from which to derive candidate models for the distribution of the total number of locations that one visits on a time scale that is short relative to that of activity space turnover. By comparing three theoretically motivated candidate models, we found support for a negative binomial distribution of this quantity. Even for applications in which our model is not used in its entirety, knowledge of the distribution of the number of locations visited should be useful for informing realistic degree distributions for network models [38,39].

Another aspect of activity space composition that our framework accounts for is that some location types appear in activity spaces in proportions much greater (or smaller) than their representation in the city. Indeed, we found clear support for this hypothesis in Iquitos, with 71% of the locations that study participants frequented belonging to the commercial, residential, or recreational location types. An implication of this result is that models that do not incorporate information about location types may miss important aspects of the structure of human movement at intra-urban scales. Although the importance of location types has been recognized in the time-space geography literature, models of human movement developed for coarser scales (e.g. [40,41]) generally ignore information about location types and may therefore perform poorly when applied at increasingly finer scales where location-type composition is highly heterogeneous. Incorporating information about location types depends not only on the details of a model, however, but also on the availability of comprehensive and reliable spatial data, which in many cases will either not be available or will have some uncertainty about location type designation. Although such information may not be widely available at present for many cities and countries, it is becoming increasingly available where research programmes are ongoing and is likely to become increasingly so in coming years, given advances in satellite image processing [42]. For certain applications in epidemiology, this information will be indispensible, as risk is greater at some location types than others [43–45].

4.1. Interpretations of model and results
4.1.1. Activity space composition
Our framework posits that activity space composition is dynamic and that different classes of locations within the

We also found a clear effect of distance from home on the probability that a location was included in one's activity space. As a result, networks of co-location or contact generated by a process such as the one described herein have a distinct spatial dimension, which has a number of implications for network topology and for the spread of infectious diseases on such networks [46,47]. Importantly, for seven of the 10 location types we considered, study participants from two geographically distinct neighbourhoods possessed statistically indistinguishable values of the parameters that govern the effect of distance from home on inclusion of a location in one's activity space. This suggests that differences in the spatial distribution of locations within and around a neighbourhood, rather than possible social or economic differences, drive differences in the realized distributions of distances over which activity space locations of these seven types are found. One implication of this result is that realized distributions of locations visited may not be comparable across different spatial contexts, as those differences could reflect differences in the distribution of locations rather than differences in the underlying movement process. On the other hand, this result suggests that the capacity of a model like ours to extrapolate to other neighbourhoods in Iquitos, and perhaps elsewhere [48], appears promising provided that comprehensive spatial data on location types are available.

4.1.2. Time allocation within the activity space

Location type was also a major factor in determining how much time individuals spent at locations in their activity spaces. Distance from home, on the other hand, only had an impact on the frequency and mean duration of visits to locations in the residential and commercial categories. These results again highlight the value that spatial data on location types can add to simulations of individual activity spaces, although a number of refinements could still be made to improve the model's capacity to simulate activity spaces that are more quantitatively consistent with empirical data. Indeed, accurately quantifying heterogeneity in the time that individuals spend at locations has the potential to affect forecasts of epidemic behaviour and predictions about the conditions under which a disease will persist [49].

4.2. Model limitations and extensions

One important extension of the framework will be identifying how differences in individual attributes such as age, sex and occupation can be used to deterministically account for variation in the movement behaviours of different individuals—for which there is evidence in our study population [9]. Our goal here, however, was not to describe every possible detail that could affect human movement, but to develop a mathematically cohesive description of how five distinct aspects of movement combine to generate patterns of time allocation. Rather than predicting the whereabouts of a single individual, the most reasonable use of our fitted model in its present form is the simulation of patterns of time allocation aggregated over many individuals in a synthetic population. Indeed, our model is suited for this purpose because it explicitly accounts for variation in movement behaviours among the diverse individuals in the study population. In applications in which certain attributes are of particular interest (e.g. age in transmission models of immunizing pathogens),

model parameters could be specified as functions of those variables and fitted to data.

Although our results clearly show that location type and distance from home are important determinants of activity space composition, the reality is that other factors likely play a role, too. One such factor that we have omitted is the existence of correlations between locations visited by cohabitants or members of certain social groups [50,51], which are likely to be of importance to the transmission of many infectious diseases [52–54]. We also assume no spatial correlation among locations in one's activity space. One factor that probably generates such patterns is that locations other than home, such as the workplace, could serve as anchor points for making choices about other locations to visit [55]. Extending our methodology to account for this phenomenon is possible, but doing so will require empirically discerning (i.e. adapting the SSI) and statistically describing (i.e. identifying associated factors and quantifying variation) which locations serve as anchor points and which are chosen secondarily. One approach to describing which locations act as anchor points could involve gathering data on which location types tend to be visited following visits to other location types, with likely anchor points being location types frequently visited following time at home [56, figs 5.7 and 5.15]. The choice of locations visited secondarily to those anchor points might then be better informed by distance from the anchor than by distance from home. It also may be worthwhile to consider differences in the identities of anchor locations for individuals of different ages, occupations, etc., as well as whether certain types of locations tend to be associated with anchor locations of certain other types.

For applications in which sequences of movements, or 'trip chains' [57], are of interest, the movement process we assume could benefit from a number of elaborations. As much of time-space geography is concerned with such sequences rather than with aggregate distributions of time, a wealth of factors have been proposed that could be incorporated into our framework. Examples include elements of the Q matrix that are specific in various ways to different origin–destination pairs [58], an increased probability of returning home with increasing trip length [59], and elements of Q that depend on day of week [60] or time of day [61]. Given a proposal for a more detailed formulation of Q, existing theory of Markov processes could be used to calculate likelihoods of parameter values given data on sequences of movement. Such realistic accounting of trip chains could also allow for consideration of time in motion between locations, which is of specific interest in some applications and of implicit interest in others vis-à-vis the fact that time in motion reduces time spent at locations in the activity space.

References

1. Prothero RM. 1977 Disease and mobility: a neglected factor in epidemiology. *Int. J. Epidemiol.* **6**, 259–267. (doi:10.1093/ije/6.3.259)

2. Coulibaly S *et al.* 2008 *World Development Report 2009: reshaping economic geography.* Washington, DC: World Bank.

3. Grenfell BT, Bjornstad ON, Kappey J. 2001 Travelling waves and spatial hierarchies in measles epidemics. *Nature* **414**, 716–723. (doi:10.1038/414716a)

4. Riley S, Ferguson NM. 2006 Smallpox transmission and control: spatial dynamics in Great Britain. *Proc. Natl Acad. Sci. USA* **103**, 12 637–12 642. (doi:10.1073/pnas.0510873103)

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

6. Eubank S, Guclu H, Kumar VSA, Marathe MV. 2004 Modelling disease outbreaks in realistic urban social networks. *Nature* **429**, 180–184. (doi:10.1038/nature02541)

7. Keeling MJ, Danon L, Vernon MC, House TA. 2010 Individual identity and movement networks for disease metapopulations. *Proc. Natl Acad. Sci. USA* **107**, 8866–8870. (doi:10.1073/pnas.1000416107)

8. Dalziel BD, Pourbohloul B, Ellner SP. 2013 Human mobility patterns predict divergent epidemic dynamics among cities. *Proc. R. Soc. B* **280**, 20130763. (doi:10.1098/rspb.2013.0763)

9. Vazquez-Prokopec GM *et al.* 2013 Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment. *PLoS ONE* **8**, e58802. (doi:10.1371/journal.pone.0058802)

10. Perchoux C, Chaix B, Cummins S, Kestens Y. 2013 Conceptualization and measurement of environmental exposure in epidemiology: accounting for activity space related to daily mobility. *Health Place* **21**, 86–93. (doi:10.1016/j.healthplace.2013.01.005)

11. Danon L, House T, Keeling MJ. 2009 The role of routine versus random movements on the spread of disease in Great Britain. *Epidemics* **1**, 250–258. (doi:10.1016/j.epidem.2009.11.002)

12. González MC, Hidalgo CA, Barabási AL. 2008 Understanding individual human mobility patterns. *Nature* **453**, 779–782. (doi:10.1038/nature06958)

13. Song C, Koren T, Wang P, Barabási AL. 2010 Modelling the scaling properties of human mobility. *Nat. Phys.* **6**, 818–823. (doi:10.1038/nphys1760)

14. Lu X, Wetter E, Bharti N, Tatem AJ, Bengtsson L. 2013 Approaching the limit of predictability in human mobility. *Sci. Rep.* **3**, 2923. (doi:10.1038/srep02923)

15. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz-Soldán V, Kochel TJ, Kitron U, Elder JP, Scott TW, Kittayapong P. 2009 The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl. Trop. Dis.* **3**, e481. (doi:10.1371/journal.pntd.0000481)

16. Pindolia DK, Garcia AJ, Wesolowski A, Smith DL, Buckee CO, Noor AM, Snow RW, Tatem AJ. 2012 Human movement data for malaria control and elimination strategic planning. *Malar. J.* **11**, 205. (doi:10.1186/1475-2875-11-205)

17. Horton FE, Reynolds DR. 1971 Effects of urban spatial structure on individual behavior. *Econ. Geogr.* **47**, 36–48. (doi:10.2307/143224)

18. Halloran ME, Longini IM, Nizam A, Yang Y. 2002 Containing bioterrorist smallpox. *Science* **298**, 1428–1432. (doi:10.1126/science.1074674)

19. Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S, Burke DS. 2005 Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209–214. (doi:10.1038/nature04017)

20. Longini IM, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DAT, Halloran ME. 2005 Containing pandemic influenza at the source. *Science* **309**, 1083–1087. (doi:10.1126/science.1115717)

21. Fournié G, Guitian J, Desvaux S, Cuong VC, Dung DH, Pfeiffer DU, Mangtani P, Ghani AC. 2013 Interventions for avian influenza A (H5N1) risk management in live bird market networks. *Proc. Natl Acad. Sci. USA* **110**, 9177–9182. (doi:10.1073/pnas.1220815110)

22. Johansson MA, Hombach J, Cummings DAT. 2011 Models of the impact of dengue vaccines: a review of current research and potential approaches. *Vaccine* **29**, 5860–5868. (doi:10.1016/j.vaccine.2011.06.042)

23. Chao DL, Halstead SB, Halloran ME, Longini IM. 2012 Controlling dengue with vaccines in Thailand. *PLoS Negl. Trop. Dis.* **6**, e1876. (doi:10.1371/journal.pntd.0001876)

24. Buliung RN, Kanaroglou PS. 2007 Activity–travel behaviour research: conceptual issues, state of the art, and emerging perspectives on behavioural analysis and simulation modelling. *Transp. Rev.* **27**, 151–187. (doi:10.1080/01441640600858649)

25. Eagle N, Pentland AS. 2009 Eigenbehaviors: identifying structure in routine. *Behav. Ecol. Sociobiol.* **63**, 1057–1066. (doi:10.1007/s00265-009-0739-0)

26. Schneider CM, Belik V, Couronné T, Smoreda Z, González MC. 2013 Unravelling daily human mobility motifs. *J. R. Soc. Interface* **10**, 20130246. (doi:10.1098/rsif.2013.0246)

27. Taylor HM, Karlin S. 1998 *An introduction to stochastic modeling*, 3rd edn. London, UK: Academic Press.

28. Marble DF. 1964 A simple Markovian model of trip structure in a metropolitan region. In *Papers, Regional Science Association, Western Section*, pp. 150–156.

29. Nystuen JD. 1967 A theory and simulation of intraurban travel. In *Quantitative geography. I. Economic and cultural topics* (ed. WL Garrison, DF Marble), pp. 54–83. Evanston, IL: Northwestern University Studies in Geography.

30. Stoddard ST *et al.* 2013 House-to-house human movement drives dengue virus transmission. *Proc. Natl Acad. Sci. USA* **110**, 994–999. (doi:10.1073/pnas.1213349110)

31. Morrison AC *et al.* 2010 Epidemiology of dengue virus in Iquitos, Peru 1999 to 2005: interepidemic and epidemic patterns of transmission. *PLoS Negl. Trop. Dis.* **4**, e670. (doi:10.1371/journal.pntd.0000670)

32. Getis A, Morrison AC, Gray K, Scott TW. 2003 Characteristics of the spatial pattern of the denguevector, *Aedes aegypti*, in Iquitos, Peru. *Am. J. Trop. Med. Hyg.* **69**, 494–505.

33. Liebman KA *et al.* 2012 Spatial dimensions of dengue virus transmission across interepidemic and epidemic periods in Iquitos, Peru (1999–2003).

PLoS Negl. Trop. Dis. **6**, e1472. (doi:10.1371/journal.pntd.0001472)

34. Paz-Soldán VA, Stoddard ST, Vazquez-Prokopec G, Morrison AC, Elder JP, Kitron U, Kochel TJ, Scott TW. 2010 Assessing and maximizing the acceptability of global positioning system device use for studying the role of human movement in dengue virus transmission in Iquitos, Peru. Am. J. Trop. Med. Hyg. **82**, 723–730. (doi:10.4269/ajtmh.2010.09-0496)

35. Paz-Soldán VA et al. 2014 Strengths and weaknesses of Global Positioning System (GPS) data-loggers and semi-structured interviews for capturing fine-scale human mobility: findings from Iquitos, Peru. PLoS Negl. Trop. Dis. **8**, e2888. (doi:10.1371/journal.pntd.0002888)

36. Burnham KP, Anderson DR. 1998 Model selection and inference. New York, NY: Springer.

37. Kitamura R, Chen C, Pendyala R. 1997 Generation of synthetic daily activity-travel patterns. Transp. Res. Rec. **1607**, 154–162. (doi:10.3141/1607-21)

38. Bansal S, Grenfell BT, Meyers LA. 2007 When individual behaviour matters: homogeneous and network models in epidemiology. J. R. Soc. Interface **4**, 879–891. (doi:10.1098/rsif.2007.1100)

39. Fumanelli L, Ajelli M, Manfredi P, Vespignani A, Merler S. 2012 Inferring the structure of social contacts from demographic data in the analysis of infectious diseases spread. PLoS Comput. Biol. **8**, e1002673. (doi:10.1371/journal.pcbi.1002673)

40. Simini F, González MC, Maritan A, Barabási AL. 2012 A universal model for mobility and migration patterns. Nature **484**, 96–100. (doi:10.1038/nature10856)

41. Liang X, Zhao J, Dong L, Xu K. 2013 Unraveling the origin of exponential law in intra-urban human mobility. Sci. Rep. **3**, 2983. (doi:10.1038/srep02983)

42. Khan O, Davenhall W, Ali M, Castillo-Salgado C, Vazquez-Prokopec GM, Kitron U, Magalhães RJS, Clements ACA. 2010 Geographical information systems and tropical medicine. Ann. Trop. Med. Parasitol. **104**, 303–318. (doi:10.1179/136485910X12743554759867)

43. Morrison AC, Sihuincha M, Stancil JD, Zamora E, Astete H, Olson JG, Vidal-Ore C, Scott TW. 2006 Aedes aegypti (Diptera: Culicidae) production from non-residential sites in the Amazonian city of Iquitos, Peru. Ann. Trop. Med. Parasitol. **100**, S73–S86. (doi:10.1179/136485906X105534)

44. Cauchemez S, Valleron AJ, Boëlle PY, Flahault A, Ferguson NM. 2008 Estimating the impact of school closure on influenza transmission from Sentinel data. Nature **452**, 750–754. (doi:10.1038/nature06732)

45. Kestens Y, Lebel A, Chaix B, Clary C, Daniel M, Pampalon R, Theriault M, Subramanian SVP, Miranda JJ. 2012 Association between activity space exposure to food establishments and individual risk of overweight. PLoS ONE **7**, e41418. (doi:10.1371/journal.pone.0041418)

46. Barthélemy M. 2011 Spatial networks. Phys. Rep. **499**, 1–101. (doi:10.1016/j.physrep.2010.11.002)

47. Onnela JP, Arbesman S, González MC, Barabási AL, Christakis NA. 2011 Geographic constraints on social network groups. PLoS ONE **6**, e16939. (doi:10.1371/journal.pone.0016939)

48. Noulas A, Scellato S, Lambiotte R, Pontil M, Mascolo C. 2012 A tale of many cities: universal patterns in human urban mobility. PLoS ONE **7**, e37027. (doi:10.1371/journal.pone.0037027)

49. Poletto C, Tizzoni M, Colizza V. 2013 Human mobility and time spent at destination: impact on spatial epidemic spreading. J. Theoret. Biol. **338**, 41–58. (doi:10.1016/j.jtbi.2013.08.032)

50. Adams J, Faust K, Lovasi GS. 2012 Capturing context: integrating spatial and social network analyses. Soc. Netw. **34**, 1–5. (doi:10.1016/j.socnet.2011.10.007)

51. Tarasov A, Kling F, Pozdnoukhov A. 2013 Prediction of user location using the radiation model and social check-ins. In Proc. 2nd ACM SIGKDD Int. Workshop on Urban Computing, article no. 8. New York, NY: ACM. (doi:10.1145/2505821.2505833)

52. Mossong J et al. 2008 Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. **5**, e74. (doi:10.1371/journal.pmed.0050074)

53. Salathe M, Jones JH. 2010 Dynamics and control of diseases in networks with community structure. PLoS Comput. Biol. **6**, e100736. (doi:10.1371/journal.pcbi.1000736)

54. Reiner RC, Stoddard ST, Scott TW. 2014 Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. Epidemics **6**, 30–36. (doi:10.1016/j.epidem.2013.12.003)

55. Matthews SA, Yang TC. 2013 Spatial Polygamy and Contextual Exposures (SPACEs): promoting activity space approaches in research on place and health. Am. Behav. Sci. **57**, 1057–1081. (doi:10.1177/0002764213487345)

56. Jakle JA, Brunn S, Roseman CC. 1976 Human spatial behavior: a social geography. North Scituate, MA: Duxbury Press.

57. Southworth F. 1985 Multi-destination, multipurpose trip chaining and its implications for locational accessibility—a simulation approach. Pap. Reg. Sci. **57**, 107–123. (doi:10.1007/BF01935284)

58. Sasaki T. 1971 Estimation of person trip patterns through Markov chains. In Proc. 5th Int. Symp. on the Theory of Traffic Flow and Transportation, pp. 119–130. New York, NY: Elsevier.

59. O'Kelly ME. 1981 A model of the demand for retail facilities incorporating multistop, multipurpose trips. Geogr. Anal. **13**, 134–148. (doi:10.1111/j.1538-4632.1981.tb00721.x)

60. Yamamoto T, Kitamura R. 1999 An analysis of time allocation to in-home and out-of-home discretionary activities across working days and non-working days. Transportation **26**, 211–230. (doi:10.1023/A:1005167311075)

61. Kitamura R, Chen C, Pendyala RM, Narayanan R. 2000 Micro-simulation of daily activity-travel patterns for travel demand forecasting. Transportation **27**, 25–51. (doi:10.1023/A:1005259324588)