Impact of human mobility on the periodicities and mechanisms underlying measles dynamics

Ramona Marguta and Andrea Parisi

Centro de Física da Matéria Condensada, Bioisystems and Integrative Sciences Institute and Departamento de Física, Faculdade de Ciências da Universidade de Lisboa, Campo Grande Edificio C8, 1749-016 Lisboa, Portugal

Three main mechanisms determining the dynamics of measles have been described in the literature: invasion in disease-free lands leading to import-dependent outbreaks, switching between annual and biennial attractors driven by seasonality, and amplification of stochastic fluctuations close to the endemic equilibrium. Here, we study the importance of the three mechanisms using a detailed geographical description of human mobility. We perform individual-based simulations of an SIR model using a gridded description of human settlements on top of which we implement human mobility according to the radiation model. Parallel computation permits detailed simulations of large areas. Focusing our research on the British Isles, we show that human mobility has an impact on the periodicity of measles outbreaks. Depending on the level of mobility, we observe at the global level multi-annual, annual or biennial cycles. The periodicity observed globally, however, differs from the local epidemic cycles: different locations show different mechanisms at work depending on both population size and mobility. As a result, the periodicities observed locally depend on the interplay between the local population size and human mobility.

1. Introduction

A large body of research in recent years has been devoted to the study of the geographical spread of childhood diseases and, in particular, measles. Early work was carried out by Cliff and co-workers on the geographical spread of measles in Iceland, showing that the disease is not self-sustained, but depends on the intensity of imports from outer countries [1,2]. Examining measles incidence between 1901 and 1970, they showed [1] that pre-1938 measles in Iceland was characterized by isolated outbreaks separated by long (up to 8 years) disease-free periods. Increased communication after 1945 led to a reduction in the interepidemic periods: between 1955 and 1968, the epidemic peaks were separated by 4 years, during which the disease was again extinct. These observations were consistent with the rate of imports being the major determinant of the occurrence of epidemics in disease-free lands.

A series of studies on the incidence of measles in England and Wales [3–6] showed that the situation was different when large populated areas were investigated. In particular, measles shows complex nonlinear dynamics that can be explained as a result of a switching between close-by attractors (annual and biennial) caused by the strength of the seasonal forcing [5]. This mechanism is shared by other childhood diseases; for instance, for rubella, the coexistence of 1 and 5 year attractors was similarly found [5]. Further analysis of a large set of administrative areas in England and Wales using wavelets demonstrated that measles propagates from large cities to smaller populated areas, with large populated areas synchronized by the seasonal forcing [6].

Recently, another approach has been suggested for the dynamics of childhood diseases conferring lifelong immunity. The idea comes from the observation that in finite populations the stochastic fluctuations around the
endemic equilibrium can become the source of recurrent epidemics. McKane & Newmann [7] studied a simple predator–prey model with fixed carrying capacity, and showed that the sustained oscillations around the equilibrium can be explained as a result of an amplification of stochastic fluctuations around a preferred frequency. The whole system behaves as a forced oscillator, except that the forcing is not provided by any external source: it is the result of the intrinsic stochasticity of the system and its nonlinear response. The same approach was demonstrated in a susceptible–infected–recovered (SIR) model with demography [8]. In particular, it was shown that, when parametrized for measles, the oscillations around the equilibrium have a preferred frequency of 2.1 years, in line with the frequency of outbreaks reported in historical time series. Further research showed that, as soon as spatial [9] or temporal correlations [9,10] are taken into account, the amplitude and coherence of these fluctuations increase considerably, making them more regular.

The three different mechanisms described (import-dependent spread, multi-stable dynamics driven by seasonality, stochastic amplification) have been suggested as major determinants in the observed incidence data, but they have been studied separately. Our aim is to explore their relative importance in a more realistic geographical setting. The increase in computational power witnessed in recent years already permits detailed simulations of geographical disease spread in both humans and animals. An important example in which the geographical spread was studied in detail is that of foot-and-mouth disease [11]: with detailed information of farm locations and cattle movements, individual-based simulations provided accurate predictions on the geographical dynamics of the disease. For human-related diseases, information about mobility is much harder to obtain. Nevertheless, simulations of the spread of influenza [12–15] at the global level and severe acute respiratory syndrome (SARS) at local level [16] have been performed in recent years. Our aim here is to perform simulations on regional scales with a description of human mobility as detailed as possible.

The main problem facing numerical work on regional geographical scales is the lack of detailed information on human mobility at such scales. For instance, Riley et al. [16] used a metapopulation model to describe the spread of SARS in Hong Kong, where each metapopulation represents a different district, but interaction among districts was based on their adjacency, with a fixed rate among neighbouring districts and a different fixed rate among more distant districts independent of their distance and population. Colizza et al. [13,14] instead describe global spread using the large-scale aeroplane connection network, but mobility at lower scales needs additional information on the mobility of individuals. Recent studies however have shown that human mobility can be statistically described on multiple scales using a truncated power law [17,18]. Moreover, a model for human mobility was recently developed [19] that depends only on a single parameter that can be fitted to available datasets on mobility for a specific country. The model overcomes several limitations of the traditionally used gravity model, and incorporates the scale invariance of commuting patterns. This model permits a coherent description of mobility on multiple scales. In this study, we present computer simulations of disease spread using a detailed geographical description of human mobility on a specific area, the British Isles, and explore how the dynamics for an SIR model parametrized for measles changes when mobility among individuals is enhanced or reduced.

2. Computer model

The computer model used for our simulations combines three level of descriptions: (i) a population level, providing a detailed geographical distribution of human population; (ii) a mobility level, detailing how individuals move around the geographical area under study; and (iii) an epidemiological level, detailing how the disease evolves.

The first layer is thus the population description: it uses data from the Gridded Population of the World (GPW) database [20], which provides estimates of the resident population of a given geographical area on a regular grid of cells of angular size 2.5 arc-minutes. This corresponds to a longitudinal size of approximately 5 km at the Equator, and around 3.5 km in northern Europe. The database also provides maps with a 30 arc-second resolution, but we did not use these highly detailed maps here. Because a single cell is our basic geographical unit, we consider each of these cells as a well-mixed population in which the disease can evolve following a specified epidemiological model. Hence, transmission within the cell occurs as in a well-mixed model; however, transmission between different cells is also taken into consideration by simulating human mobility. We allow each individual to move to a different cell with a given probability that depends on the propensity of individuals to move to that cell. We assume that individuals commute daily between different locations and, in order to evaluate the probability of reaching a specific location, we are interested in knowing what the daily flux is between the home location and the destination location. The description of human mobility we use represents the second layer of our model and is based on the recently introduced radiation model [19], which was shown to accurately predict human mobility patterns and overcomes a series of limitations of the widely used gravity model [19]. The model gives the flux of individuals \( T_{ij} \) travelling from source \( i \) to destination \( j \) characterized by a population \( n_i \) to the destination \( j \) characterized by a population \( n_j \), the two being separated by a distance \( r_{ij} \). The flux is given by the following formula:

\[
T_{ij} = T_i (m_j + s_j)(m_i + n_i + s_i),
\]

where \( s_j \) is the total population within the circle of radius \( r_{ij} \) around the source \( i \), excluding both populations \( m_i \) and \( n_i \). Here, \( T_i = \sum_{j \neq i} T_{ij} \) is the total flux of commuters leaving location \( i \) and it is proportional to the population of the source location. Hence,

\[
T_i = m_i N_c/N,
\]

where \( N_c \) is the total number of commuters and \( N \) is the total population in the country. The fraction \( N_c/N \), which represents the fraction of individuals in the population that participate in the long-distance displacements, controls the level of transmission to different cells. This fraction can be estimated by existing data on human mobility. The model has a few notable properties that are particularly advantageous to us. The flux between locations is controlled by a single parameter \( N_c/N \), which represents the fraction of individuals in the population that participate in the
long-distance displacements and thus controls the level of transmission to different cells. By modifying this parameter, the fluxes between locations can be varied with an explicit meaning in terms of the number of moving individuals. Additionally, the fluxes are source limited: increasing $N_i/N$ the number of individuals leaving each cell never exceeds the population of the cell, a property that does not hold for the widely used gravity model.

Detailed studies on human mobility, based on the tracking of a large number of mobile phone users [18], have shown that individuals have a set of preferred locations they visit with specific frequencies. In particular, an individual’s probability of going to visit a specific location was generally found to follow a Pareto law [18]; a detailed study on human mobility within the city of Portland, OR, USA [21] showed it to follow a power law. We mimic a similar behaviour by assigning to each individual a set of preferred locations: each individual has its own specific set, and he will visit them with a frequency that reflects his propensity to visit as dictated by the radiation model. These preferred locations are chosen randomly for each individual, but in such a way that the overall fluxes between locations correspond to those predicted by the radiation model. In practice, we build the set of preferred locations and their frequencies as follows: for individuals in each cell $i$, we calculate the probability of visiting cell $j$. This is simply given by $p_{ij} = T_{ij}/m_i$. The probability $p_i = \sum_j p_{ij}$ corresponds to the probability that an individual leaves cell $i$ and through equation (2.2) it corresponds to $p_i = N_i/N$; it follows that $1 - p_i$ represents the probability of not leaving cell $i$. Hence, the full set of choices for an individual are all the cells of the map $j$ (with $j \neq i$), each chosen with probability $p_{ij}$, and location $i$ with probability $1 - p_i$. The associated probabilities form a discrete distribution that can be used to extract preferred locations. We sample this probability distribution $M$ times, obtaining a list of locations in which those with high probability will have been chosen more often. We can thus count the number of times each distinct location has been chosen and transform this count into a frequency by dividing it by the size $M$ of the sample. The result is a list of distinct locations with an associated frequency of visit that can be used to decide where the individual should be moved to. The size of the sample $M$ was chosen to be equal to 50 in all simulations: values larger than one order of magnitude did not lead to significantly different results even for low values of $N_i/N$.

Individuals are moved daily: at the start of each day, a destination is chosen for each individual among his set of preferred locations according to their frequency of visit. The individual is then moved to that preferred location and participates in the dynamics of that cell, thus permitting transmission of the disease to different cells. At the end of the day, the individual is moved back to his original location. The use of a set of preferred locations is particularly advantageous for us, because it reduces the problem of choosing a destination from a number of possible destinations as large as the total number of grid cells with non-uniform probability, to that of choosing among a small set of preferred locations. This is justified by the fact that most locations will have a very low probability of visit. For the simulations here presented, the average number of preferred locations per individual was at most close to 15, compared with a total number of cells close to 90,000.

As mentioned earlier, our aim is to study large geographical areas and thus large populations. In order to easily handle the amount of data we are dealing with, we parallelize the algorithm by assigning chunks of the geographical map under study to different computing nodes. To achieve this, an algorithm is used that partitions the map into regions that are compact in shape and support similar population sizes. The algorithm performs a multi-level coarsening of the grid map in order to obtain a coarsed map that is composed by a small number of coarsed grid elements. An optimal subdivision of the map is then rapidly found using a simulated annealing algorithm. The map is then decoarsened one level at a time, and, each time, the final part of a simulated annealing algorithm is performed to adapt the optimal solution found at the coarser level to the new decoarsed level. The procedure is repeated until decoarsening leads to the original grided map. This complex algorithm allows us to partition large maps in very short times. The full procedure is detailed in the electronic supplementary material, SI.

The third layer in our simulations is the epidemiological model, which describes the disease dynamics in each cell. The simulation program supports any compartmental model provided we define the compartments and the set of transitions with their corresponding transition rates. In this work, we focused on the SIR model. In each cell, individuals can be either susceptible (S), infected (I) or recovered (R). When infected, they can transmit the disease to other susceptibles, whereas recovered individuals have a lifelong immunity to the disease. Demography acts by removing infected and recovered individuals and replacing them with new susceptible individuals, thus replenishing the susceptible class. We also include gamma-distributed recovery profiles [22] by using multiple infective classes. The number of infective classes $L$ interpolates among an exponential recovery for $L = 1$ and a deterministic recovery for $L = \infty$. Simulations with $L = 1$ and $L = 2$ show a remarkable qualitative and quantitative difference: this is understandable, because the distribution of recovery times is exponential for $L = 1$ (thus corresponding to a constant rate of recovery), whereas it is non-exponential with a maximum away from 0 for $L > 1$ (corresponding to a recovery rate which depends on the time from infection), hence changing $L$ from 1 to 2 introduces an important difference. No qualitative difference appears instead among simulations with $L = 2$ and simulations with higher values of $L$. In this work, our simulations were performed with $L = 2$: we did not explore any further dependence on $L$, reserving an eventual exploration for later studies.

The processes considered are thus described by the following set of events:

1. transmission: $SI \xrightarrow{\beta} I_1$
2. recovery $I_j \xrightarrow{1} I_{j+1}$ ($1 \leq j < L$)
3. recovery $I_j \xrightarrow{\gamma} R$
4. death/birth $[I_j, R] \xrightarrow{\mu} S$.

We use the same rate $\mu$ for death and birth ensuring, as is commonly done, that the population size remains constant. The parameter $\mu$ does not change during a simulation, thus we do not consider changes in birth and/or death rate. We include seasonality by letting the parameter $\beta$ be time dependent—the form of the seasonality is assumed to be term-time forcing [5]:

$$\beta(t) = \beta_0[1 + \beta_1 \text{term}(t)],$$

(2.3)
where term \((t)\) is a function taking values +1 during school time and -1 otherwise \([5,23,24]\), and \(\beta_1\) is a parameter that controls the amplitude of the forcing.

The dynamics is simulated in each cell soon after moving individuals to new locations chosen from their sets of preferred locations according to their frequency of visit as detailed earlier in this section. Each cell thus will be reduced in population owing to the individuals who leave the cell for other destinations, and will increase owing to individuals coming from other cells. Individuals who, after the mobility phase, are in the cell (hence residents who did not leave the cell and visitors coming from other cells) can interact for that day. Dynamics is simulated using the Gillespie event-selection algorithm for 1 day \([25]\), and disease can hence be transmitted between current residents and visitors of the cell. At the end of the day, individuals are moved back to their home cell, and then the whole procedure is repeated.

The system is initially placed close to endemic equilibrium. The basic reproductive number and the equilibrium values for the number of susceptible, infective and recovered individuals in a well-mixed population can be calculated from the deterministic approximation. For an SIR model with \(L\) infective classes, the basic reproductive number depends on \(L\) and is given by \(R_0 = \beta / \mu [1 - (L \gamma / \mu + \mu)^j]\). The equilibrium can be expressed in terms of \(R_0\) as

\[
S^* = \frac{1}{R_0},
\]

\[
I_j^* = \left( \frac{L \gamma}{L \gamma + \mu} \right)^j \frac{\mu}{L \gamma + \mu} \left(1 - \frac{1}{R_0}\right), \quad 0 \leq j < L
\]

\[
R^* = \left(1 - \frac{\mu R_0}{\beta}\right) \left(1 - \frac{1}{R_0}\right).
\]

Note that the total number of infectives at equilibrium is \(I^* = \sum_{j=1}^{L} I_j^* = \frac{\mu}{\beta} (R_0 - 1)\) depending on the number of infective classes \(L\) only through \(R_0\). Individuals are initially assigned one epidemiological class with a probability proportional to the corresponding endemic equilibrium prevalence, thus building in each cell a random discrete realization close to the endemic equilibrium. Owing to the mobility of individuals, the resulting set-up does not correspond to the local endemic equilibrium, but it is sufficiently close to it: the system settles after about 7–8 years of simulated time.

3. Methods and data

For the investigation described in this work, our area of study corresponds to that of the British Isles. The choice comes from the wealth of data available for England and Wales for childhood diseases, and the fact that the British Isles can be considered to some extent an isolated unit, neglecting imports from outer regions. The population count maps are obtained by combining the GPW maps for the UK, Ireland and the Isle of Man, thus excluding any other surrounding country. It is of course possible to extend the study to include the influence of neighbouring countries, but we did not pursue this goal in this study.

Disease parameters for measles were chosen following \([5]\). We set the transmission rate to \(\beta_0 = 1.175 \text{ d}^{-1}\); the recovery rate was set to \(\gamma = 1/13 \text{ d}^{-1}\) and the death/birth rate to \(\mu = 5.5 \times 10^{-5} \text{ d}^{-1}\). The seasonal forcing was set to \(\beta_1 = 0.25\) following \([5]\), except when we were interested in comparing with the non-forced model; in that case, we used \(\beta_0 = 0\).

The only remaining parameter needed to perform simulations of mobility is the ratio \(N_c/N\) that controls human mobility. We did not consider vaccination in this work. Hence, the value of this ratio should be referred to the interval of years corresponding to that of available pre-vaccination data that, for measles, range between 1944 and 1966 \([26]\); however, there is no available information about mobility for those years. As a consequence, we used mobility as a proxy to investigate its influence on the dynamics. First, we estimated the mobility ratio using recent census data available from the Office of National Statistics of the United Kingdom \([27]\). We were able to obtain a value of \(N_c/N = 0.176 \pm 0.001\) for our simulations of the UK by comparing the distribution of the average flux of commuters, with respect to distance covered as predicted by our radiation model based on gridded maps, with that derived from census data of the UK. The procedure used for this estimate is detailed in the electronic supplementary material, \(S2\), and takes into account the specific set-up of our model. We then assumed this value to be valid for the whole area under investigation, and used it as a reference for the mobility ratio. We explored different values of \(N_c/N\) to investigate the influence of this ratio on the dynamics, varying it between 0 (for which no mobility occurs) and 0.2.

Incidence data for measles in England and Wales were made available by the authors of \([26]\). Data provide biweekly incidence for 60 cities in England and Wales between 1944 and 1966, as well as average population size of each of the cities. The original unaggregated weekly data come from the Office of Population Censuses and Surveys \([26,28]\). Data provide biweekly incidence for 60 cities in England and Wales between 1944 and 1966, as well as average population size of each of the cities. The original unaggregated weekly data come from the Office of Population Censuses and Surveys \([26,28]\).

Under-reporting was estimated to be around 50% \([26,29]\): the correct infective incidence is estimated in \([4]\) through a time-series analysis that takes into account the variation in birth rates during those years. Because we assume constant birth rates, here we use a simplified approach, and we assume that the correct number of infective cases is twice the reported number.

4. Results

Before illustrating results obtained through simulations, it is worth discussing briefly two extreme cases. The first case occurs when \(N_c/N\) is set to zero: this corresponds to lack of mobility of individuals who are constrained to their residence cell. Thus, the model describes a collection of isolated well-mixed populations. In the map under consideration, all cells have a population below the critical community size (the population size of the inhabited cells varies between 1 and 137 646 individuals, the latter corresponding to a cell located in the area of London), and because no interaction between cells is possible, they will experience permanent disease extinction. On the other hand, for values of \(N_c/N\) close to 1, the model describes a set of cells where all individuals travel to other cells: hence, the dynamics is more well-mixed, but still it does not correspond to that of a well-mixed population. In fact, each individual will most probably interact with individuals having the same preferred location, hence the probability of interaction with any other individual is not uniform. This should not be surprising, because it reflects a realistic description of human mobility where
individuals interact most intensively with those visiting the same preferred locations, be it for work, shopping, etc. As a result, this model does not interpolate between isolated and well-mixed populations, but rather it approximates realistic human mobility at given values of the mobility ratio, with the two extremes of $N_c/N = 0$ and $N_c/N = 1$ both describing unrealistic situations.

First, we are interested in the global behaviour in the area under study for different values $N_c/N$. Because the model does not describe a well-mixed population for any value of the mobility ratio, it becomes important to understand whether the endemic equilibrium is maintained in these conditions, and how the dynamics around the endemic equilibrium change. In our simulations, we observe extinction for levels of $N_c/N$ lower than 0.0005: this value suggests that persistence of the disease can occur even with extremely low mobility levels; however, the specific threshold is influenced by the size of the grid cells and could be different if a population not artificially partitioned into gridded cells was considered. For higher values of $N_c/N$, the average infective incidence achieves a level close to that of the well-mixed population, but rather it approximates realistic situations.

We obtain time series for the cities by considering the aggregated incidence data of the cells within the area occupied by the city. We simplify the shape of the city to a circle centred on the geographical centre of the city, covering a surface equivalent to the surface area of the city itself. The radius of the circle is fine-tuned to get a population close to that reported in census data. We specify in the caption of the figures the resulting population used for simulation data as well as the historical city population to which the measles datasets refer.

Figure 2 shows the disease incidence in the area of London for two values of the mobility ratio: $N_c/N = 0.05$ (figure 2a) and $N_c/N = 0.2$ (figure 2b). In both cases, the plot shows a regular series of biennial epidemics, compatible with the data available for the area of London (figure 2d): except for the incidence level, the frequency of outbreaks seems unaffected by this difference in mobility ratios. This is remarkably in contrast with the global time series, especially when considering that the area of London accounts for about 12% of the whole population in the area under consideration. This suggests that for London the dynamics are fundamentally internal: imports from other locations do not play a relevant role. The different behaviour of the global time series for $N_c/N = 0.05$ must hence be due to dynamics within lower populated areas. Of the three mechanisms we are interested in, seasonality for London plays the major role. For such large locations, the sequence of biennial outbreaks must be a consequence of the seasonal forcing. This is illustrated in figure 2c: where the incidence for London is shown in the absence of seasonality ($B_t = 0$). In this case, the incidence level fluctuates around the endemic equilibrium, and the resulting time series differ remarkably from the behaviour observed in datasets for London. Note that

![Figure 1](http://rsif.royalsocietypublishing.org/)

**Figure 1.** (a) Daily average number of infectives for the whole British Isles as a function of the mobility ratio $N_c/N$: circles correspond to simulations with seasonality and squares to simulations without seasonality. The dashed line corresponds to the endemic equilibrium for a well-mixed population. All values are close to the endemic equilibrium; simulations for $N_c/N = 0.0005$ and lower values drop to zero owing to extinction events. The other plots show the time series of seasonally forced simulations for the different values of the mobility ratio: (b) $N_c/N = 0.001$, (c) $N_c/N = 0.008$, (d) $N_c/N = 0.05$, (e) $N_c/N = 0.2$. Simulations with seasonality were performed with parameters as described in the main text, resulting in $\langle \beta(t) \rangle = 1.32$. Simulations without seasonality were performed with $\beta_0 = (\beta(t)) = 1.32$. (Online version in colour.)
since the definition of the extent of the area of London changed in 1965, in our simulation, we use the current definition while datasets (figure 2d) are based on the previous definition and thus refer to a smaller population size.

In figure 3, we show the time series for Sheffield, a city with a population of about 500,000 individuals, considerably smaller than London, but still beyond the critical community size. The mobility among the high-populated cells of the area leads to persistence and to a biennial cycle for $N_c/N = 0.1$ (figure 3a). For $N_c/N = 0.05$ (figure 3b), the interepidemic period is slightly larger although no extinction occurs in the area: the level of imports affects the local dynamics causing this light shift. For lower values of the mobility ratio (figure 3c), the decrease in import leads to extinctions and to a larger change in frequency of the outbreaks, being more distant in time.

The situation is different when we look at low-populated areas. Figure 4a–d shows the infective incidence for the city of Chester. This is a city with a population of 57,000 residents, about 30 km from the larger city of Liverpool, and 60 km from Manchester. For $N_c/N = 0.1$ (figure 4b), the time series of measles incidence are quite irregular, but a
detailed analysis shows several extinction events. For lower values of the mobility ratio, extinction events combined with lower mobility become increasingly important: figure 4a shows the time series for $N_c/N = 0.05$. The lower number of imports results in more spaced outbreaks, with a frequency for the main peaks roughly triennial. Here, seasonality has a limited effect; the amplitude and frequency in the time series in (figure 4a) and (figure 4c) barely differ. This suggests that, in contrast with the case of London, the frequency at which outbreaks are observed in datasets (figure 4a) for small populated areas is a direct consequence of the level of mobility.

The irregularity of the time series might result from the proximity of Chester to larger cities. In the electronic supplementary material, S3, we include a movie for $N_c/N = 0.1$ showing how the disease spreads in the area we are studying. Travelling waves originating from high-populated areas can be observed: these propagating waves can avoid extinction during the epidemic troughs and can trigger recurrent epidemics in low-populated areas. Some of these low-populated areas are subject to the occurrence of multiple waves leading to irregular time series. The movie in the electronic supplementary material shows that the region of Chester is exposed to waves propagating from the neighbouring large cities and to waves coming from the south and west. Figure 5 shows some snapshots taken from this movie.

Besides the two cases of large and small populated cities, a third case occurs when a city sustains a set of epidemic outbreaks with rare extinction events. This typically occurs for cities with intermediate population sizes: in the literature, the city of York is often used to illustrate this case. In our simulations, we fine-tune the city size to obtain a population of about 100 000 individuals, in line with the population reported in historical datasets for measles outbreaks [26]. We observe that a moderate mobility ratio permits sustained epidemics. Figure 6a,b shows the time series for $N_c/N = 0.1$ and $N_c/N = 0.05$. The results in figure 6a show a biennial cycle for major outbreaks with no extinctions. For $N_c/N = 0.05$ instead, the time series (figure 6b) shows a slightly larger cycle close to 2.5 years: the increase of the cycle period might be related to the occurrence of extinction events occurring owing to the lower mobility. Figure 6c shows the time series for the city of York for a non-seasonal simulation ($\beta_1 = 0$) and a mobility ratio $N_c/N = 0.1$. The frequency of outbreaks remains close to 2 years: this is in line with the fact that the stochastic fluctuations have a periodicity of 2 years for measles.
Figure 7 shows a comparison of the power spectra of the time series for $N_c/N = 0.1$, $0.08$ and $0.05$ for the seasonal and non-seasonal case as a function of the frequency. The seasonal simulations show a peak at the frequency corresponding to a cycle of 1 year. Simulations for $N_c/N = 0.1$ and $0.08$ show a stochastic peak close to 0.5, corresponding to a biennial cycle; for $N_c/N = 0.05$, the stochastic peak corresponds to a larger cycle of 2.5 years. Comparing seasonal and non-seasonal simulations with the same $N_c/N$, we can observe that the stochastic peaks are close in amplitude. For $N_c/N = 0.05$, the seasonal forcing does not affect the stochastic peak, whereas a mild dependence is observed for larger values.

In figure 8, we report a detailed plot of the observed dependence of the period of outbreaks versus the mobility ratio $N_c/N$ for the cities discussed in the text as well as the global time series. We can distinguish four behaviours depending on $N_c/N$. For the highest values used ($N_c/N \approx 0.15$), the period of outbreaks converges to a biennial cycle for all cases. Below this level of mobility instead the period observed for cities and that observed for the global time series differ. For intermediate values of the mobility ratio ($0.025 \leq N_c/N \leq 0.10$), the global time series show annual outbreaks. For these same values, London shows its characteristic biennial cycle, in line with our observations discussed earlier. Smaller cities show instead a period of fluctuations that depends on $N_c/N$ and increases when mobility is decreased. Chester has a more pronounced dependence than Sheffield or York owing to its small population size, which leads to local extinctions and a strong dependence on imported cases from nearby locations for
most values of $N_c/N$. For low values of the mobility ratio $(0.001 < N_c/N \leq 0.01)$, the city of London joins the global time series in showing annual cycles, whereas smaller cities maintain their strong dependence on $N_c/N$. Finally, for very low values of mobility ($N_c/N \approx 0.001$), the global time series become multi-annual, and most cities follow the same periodicity. Interestingly, London shows a smaller period of fluctuations, once again deviating from the behaviour of the other cities.

5. Discussion and conclusions

A large body of research has been devoted to measles in order to understand its dynamics. Three mechanisms have been identified as most relevant: import-dependent spread, which permits propagation in areas where the disease went extinct; multi-stable dynamics driven by seasonal forcing; and stochastic amplification. These mechanisms have been studied in detail at the level of cities on models where the population is well mixed. Recently, some attempts have been proposed to investigate how some of these mechanisms interact on a larger collection of cities. Here, we have explored a more detailed description of mobility using a model that simulates propagation on a detailed scale, reaching large cities as well as low-populated areas of the British Isles.

The new ingredient that we have introduced is a realistic description of mobility at a level of detail and on an area which extends to the whole British Isles. This level of description has two relevant characteristics: first, it provides the possibility to investigate \textit{in silico} the role of high-populated as well as low-populated areas; second, it uses a description of mobility that goes beyond the well-mixed modelling of mobility used in recent publications.

Our results show that the level of mobility of individuals influences disease incidence both locally and globally. At the global level, the time series change from showing multi-annual cycles for very low mobility levels, to annual cycles, to biennial cycles for increasing mobility ratio. Locally, the change in disease incidence depends on the population size of the areas under study. In particular, we observe different mechanisms at work: in the large populated areas, the seasonal forcing drives the system out of equilibrium creating the typical biennial outbreaks observed in datasets for most values of the mobility ratio. Low-populated areas instead have the frequency of outbreaks dependent on the level of mobility owing to the extinction following each outbreak, until the mobility ratio is high enough to sustain the epidemic. When this occurs, the time series is still heavily influenced by the local mobility, with waves propagating from the surrounding areas that make the time series strongly irregular. An intermediate regime occurs for populations large enough to be less influenced by the surrounding areas, for which extinction is a rare event. In this situation, with an adequate level of mobility, the population sustains the epidemic and, owing to the low population size, amplifies the stochastic fluctuations of the disease.

The results summed up in figure 8 show the complex dependence of the period of outbreaks on the mobility ratio. Two behaviours are particularly interesting. First, the response of the city of London to variations of the level of human mobility is unique when compared with that of smaller locations. While we have discussed the cause of the biennial cycles observed for intermediate values of $N_c/N$, we have no clear explanation for the annual cycles observed for low levels of mobility. Second, the global time series show biennial cycles for high levels of mobility, whereas intermediate and low levels of mobility lead to annual cycles, in contrast with what occurs at the local level. The origin of this discrepancy might be linked to the contribution coming from locations which happen to be out of synchrony with respect to other locations. For low levels of $N_c/N$, the range of periodicities observed at different locations might concur to a similar result. The analysis discussed here is by no mean exhaustive. For instance, the movie in the electronic supplementary material, S3, shows an area in the north of England, which includes Liverpool, Manchester and Leeds, which behaves as the beating heart of the region (see also [6]), not dissimilarly from what happens in the area of London. However, the origin of such behaviour is probably related to synchronization effects between these cities and the high-density populated surrounding area, and requires further studies.

A few observations are required at this point. First, the level of mobility reported in this work should not be compared too closely with the level of mobility taken from census data. The latter refers to the level of commuting for work as measured from census data for 2011. Here, we study a disease which typically affects children, and as such the mobility patterns for these young individuals might differ from those of adult individuals. Our model does not have any age structure, therefore no distinction is attempted among younger and older individuals. We do not have data for mobility between 1944 and 1966, hence we have used the value of $N_c/N$ obtained for census data only as a reference value. The mobility level that we use should be regarded as an effective mobility that weights the mobility for different age classes in a way similar to a single contact parameter SIR when compared with an age-structured SIR. Second, our simulations have been performed using an SIR model with two infective classes. We have not explored higher values of the number of infective classes nor SEIR (susceptible, exposed, infectious, recovered) simulations. We selected seasonality and disease parameter values as used in previous simulation works for measles, but did not perform any sensitivity analysis on them. It has also been shown that the inclusion of non-constant birth/death rates leads to improved agreement with the historical incidence data. Using the variation in birth rates in England and Wales, it was possible to reconstruct the dynamics of the susceptibles and thus explore the effects of variation in birth rates on the time series both at global level [4] and at the level of cities [26]. This could explain for instance the pre-1950 annual cycles observed in time series [4]. There are several aspects of the epidemiological modelling that require further investigation: we will eventually report results on these issues in a separate publication.

On a more computational ground, we presented a computer model that enables the study of the spread of infectious diseases in geographically detailed areas, using a statistical description of mobility based on the radiation model. Our simulation model uses a simulated annealing technique based on coarse-graining of base areal units to partition the geographical map of interest into regions that can be assigned to different computing nodes for parallel computation, thus permitting unlimited complexity. Epidemiological models
other than the SIR can be easily implemented by modifying the number of classes simulated and the corresponding list of transition events, while maintaining the underlying structure describing human mobility unmodified. On the other hand, the radiation model was used by us to infer the probability for an individual to move to a different location. Hence, it is always possible to use a different model for human mobility that can provide an alternative description of the fluxes between locations, while maintaining other features (for instance, the use of a list of preferred locations per individual) unmodified. One possibility could be to estimate such fluxes from data of mobile phone tracking: a recent paper has shown that the radiation model and mobile-tracking data can provide complementary results on the study of the spatial spread of infectious diseases [30].

Acknowledgement. The authors acknowledge funding from the Fundação para a Ciência e a Tecnologia (FCT) under contract no. PTDC/SAU-EPI/112179/2009.

References


