PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A Scaling Analysis of Measles Epidemics in a Small Population

C. J. Rhodes and R. M. Anderson

Phil. Trans. R. Soc. Lond. B 1996 **351**, 1679-1688
doi: 10.1098/rstb.1996.0150

A scaling analysis of measles epidemics in a small population

C. J. RHODES* AND R. M. ANDERSON

Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.

SUMMARY

We present a detailed analysis of the pattern of measles outbreaks in the small isolated community of the Faroe Islands. Measles outbreaks in this population are characterized by frequent fade-out of infection resulting in long intervals when the disease is absent from the islands. Using an analysis of the distribution of epidemic sizes and epidemic durations we propose that the dynamical structure observed in the measles case returns reflects the existence of an underlying scaling mechanism. Consequently the dynamics are not as purely stochastic as is usually thought for epidemiological systems of this sort. We use a lattice-based epidemic model to provide a theoretical estimate of the scaling exponents and show that a conventional compartmental SEIR model is unable to reproduce this result. The methods discussed in this paper are general and represent a novel way to consider the dynamics of any other communicable disease where there is frequent fade-out in the case returns.

1. INTRODUCTION

In recent years the study of the population biology of ecological, epidemiological and immunological systems has stimulated much discussion into the possible role of nonlinear effects in population dynamics (May 1976; May 1987; Olsen *et al.* 1988; Grenfell 1992; Tidd *et al.* 1993; Grenfell *et al.* 1994; Anderson 1994). Many systems which were previously thought of as noisy limit cycles or purely stochastic have been re-interpreted as manifestations of real nonlinear processes at work in the underlying interacting populations. This has led to the development of an extensive range of new mathematical instruments for the analysis of time series generated when monitoring the prevalence of a pathogenic agent, or the abundance of a species in an ecosystem (Sugihara & May 1990; Sugihara *et al.* 1990). The influence of these ideas has been felt far beyond biology and now extends into many other areas of science and engineering (Mullin 1993).

Central to this appraisal of biological dynamics has been the study of the incidence of measles virus infection within large urban communities in the developed world. The extensive and often quite detailed measles data sets have been compared with a variety of mathematical models, often based on the compartmental SEIR structure. Quantities such as characteristic frequencies sand Liapunov exponents in observed and stimulated time series have been compared. Such analysis are typically based on trends in populations where the infection is endemic prior to the introduction of mass immunization. For measles, the critical population size above which chains of transmission can persist is estimated to be around 250–300,000 individuals (Bartlett 1957, 1960; Black 1966). Populations smaller than this are subject to frequent extinctions of the infection only for it to be re-introduced at a later date when an infective individual enters the community. In these circumstances the resulting temporal dynamics are highly irregular and traditional methods of analysis, such as power spectra or autocorrelation (Anderson & May 1991), do not yield useful insight into the epidemiology of the disease. Bartlett (1957) classified measles epidemics of this intermittent form as Type III epidemics and, until recently, they have been the subject of somewhat less theoretical attention than the data sets from populations larger than the critical community size.

We have recently shown (Rhodes & Anderson 1996*b*) that by measuring the distribution of epidemic sizes and epidemic durations, it is possible to uncover regularities in the Type III dynamics in small populations. In this paper we present a full analysis of the available measles data for the Faroe Islands showing how the scaling relations can be extracted from the time-series. We also show how it is possible to estimate likely epidemic sizes and durations that might be expected in a given time-interval. The methods we present are general and can be applied to any communicable disease where Type III dynamics are observed. Given the success of much recent epidemiological modelling of measles we attempt to recover the observed dynamics with the results of a frequently used compartmental measles model. However, it turns out that for these small populations and fast dynamics it is not possible to estimate the scaling exponents using this

* Author to whom correspondence should be addressed.

type of model, so instead we use a recently introduced lattice-based model to provide a theoretical estimate of the exponents.

2. MEASLES EPIDEMICS IN THE FAROE ISLANDS

Located midway between Scotland and Iceland, the Faroe Islands are a sparsely population, geographically isolated cluster of islands. Historically the economy of the islands has been closely tied to fishing and whaling, resulting in contacts with European and Icelandic fishing fleets. There has also been regular trading contact with Scandinavia and the United Kingdom. It is believed that these are predominant routes by which the measles virus has entered the population. The Faroe Island monthly measles case returns for the years 1912–1969, before mass-vaccination was introduced, are shown in figure 1 and the intermittent nature of the epidemics with frequent epidemic fade-out is clear to see. Due to the small size of the population the accuracy of these measles records is believed to be high (Cliff *et al.* 1993). The extensive morbidity and occasional mortality that arose during measles epidemic outbreaks in the largely susceptible population meant that few of those manifestating the visible symptoms of the disease escaped notice. A detailed study of the epidemiological patterns of measles infection in many different island and mainland populations is given in the book by Cliff *et al.* (1993).

There are 649 months recorded, with measles cases present in 188 of those months. Looked at in detail show many consecutive months with no cases present followed by a rapid outbreak of infection, when an infected index case arrives in the community. After a variable period of some months there is a return to an absence of cases. The smallest of these epidemic outbreaks has only one reported case, whereas the largest outbreak recorded 4456 measles cases. In figure 2*a* is plotted the number of epidemic outbreaks that commenced in a given month. There appears to be a slight excess of epidemics that begin between May and July, but this is to be expected since there would be more frequent contact with the outside world during the summer months. Also, as shown in figure 2*b*, the largest epidemic recorded began its course in June, coinciding with a whaling gathering, though the spread of epidemic initiations throughout the year is fairly uniform. The same analysis performed for the epidemic durations is shown in figure 2*c*. Again, the longest epidemic (though not the largest in terms of total case numbers) began in June and lasted for the next 20 months, though there is a slight tendency for longer epidemics to begin between the months of April and July. Finally, the average number of cases per year to arise each month is shown in figure 2*d*. Despite the fact that the number of epidemic initiations is lower than during summer, as shown in figure 2*a*, the winter months record the highest number of cases per month. This is probably due to the fact that the population is confined indoors and individuals are in contact longer. An increase in the case reports in the winter compared with summer months is a feature that is also observed in the measles case returns of large urban populations (Fine & Clarkson 1982).

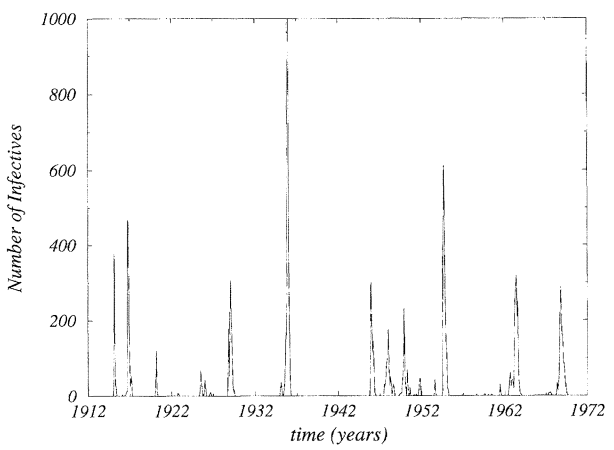


Figure 1. The Faroe Island monthly case returns 1912–1969 inclusive.

The time series for the Faroe Islands measles case returns is composed of many distinct and temporally separated epidemic events, each with a given size and duration. We have sought an analysis of the epidemiological record that directly addresses the discrete nature of these epidemic events, and it appears that simple power laws govern their size and duration (Rhodes & Anderson 1996*b*).

3. SCALING ANALYSIS

Recently it has been demonstrated that the distribution of epidemic sizes and durations are strongly suggestive of power-low behaviour. There has been much discussion concerning the possible emergence of power-law phenomena in biological systems, particularly with regard to issues relating to the dynamics of evolution (Bak & Sneppen 1993; de Boer *et al.* 1994). Power laws are characteristic of self-similar (fractal) phenomena, that is, there is an absence of any particular scale (Bak *et al.* 1988; Solé & Manrubia 1995; Solé & Bascompte 1996). If, on the other hand, an exponential type of distribution is present, then some characteristic scale is relevant and will dominate the dynamics.

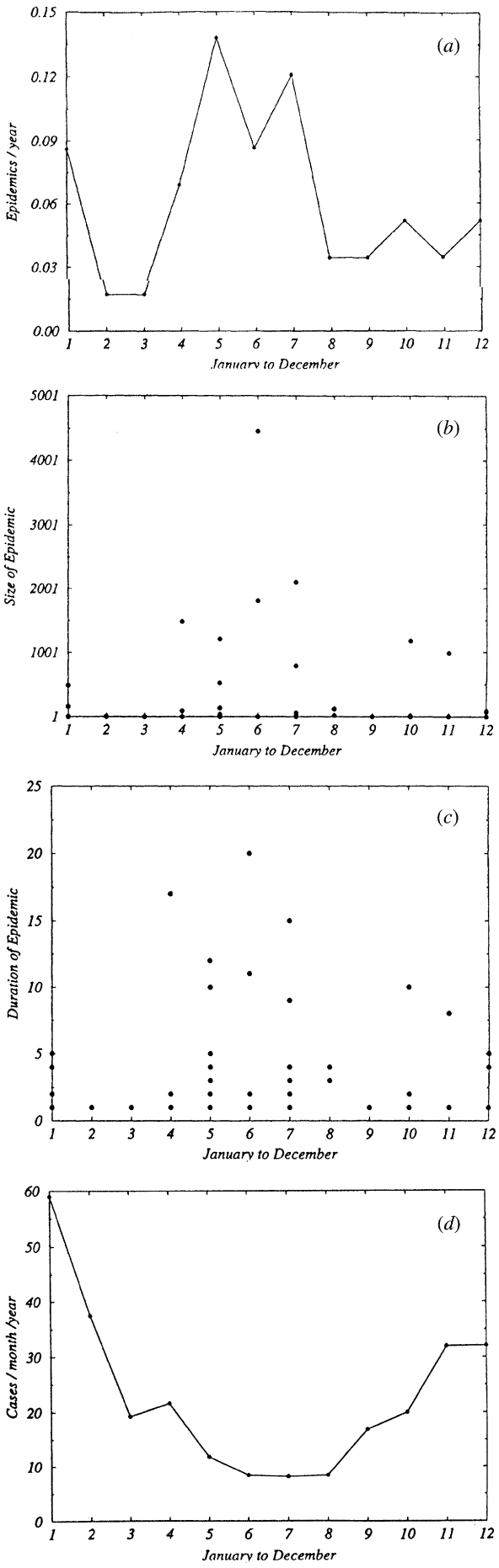
(*a*) *Analytic background*

The quantity we wish to calculate is the number of epidemics of a given size, but because we only have limited data, we are only able to indirectly calculate this quantity. Our approach is analogues to that used in geophysics when studying the dynamics of earthquakes (Sornette & Sornette 1989; Chen *et al.* 1991). We assume that, in a given time interval, the number of epidemics of size s, $N(s)$, scales as

$$N(s) \propto s^{-1-b} \tag{1}$$

or, more concisely

$$N(s) = as^{-\beta} \tag{2}$$



where $\beta = 1 + b$. The short duration of the epidemiological time-series makes this quantity difficult to assess directly, so instead we calculate the number of epidemics greater in size than a given size, s_c, namely

$$N(s > s_c) = \int_{s_c}^{\infty} a s^{-\beta}\, ds \tag{3}$$

The integration is straightforward, so we can say that if we observe

$$N(s > s_c) = \frac{a}{b} s_c^{-b} \tag{4}$$

then the number of epidemics of size s scales as

$$N(s) = a s^{-1-b} \tag{5}$$

From equation 4 if we plot the number of epidemics greater than a given epidemic size s then we can estimate the scaling exponent b. Taking the logarithm of equation 4 we have

$$\log N(>s) = \log\left(\frac{a}{b}\right) - b \log s \tag{6}$$

Such a log-log plot allows an estimate of b from the gradient and an estimate of a from the intercept, allowing a full parameterisation of equation 5.

An identical analysis can be carried out for the scaling of epidemic durations. Assuming that the number of epidemics of duration t scales as

$$N(t) \propto t^{-1-c} \tag{7}$$

a log-log plot of the number of epidemics of duration greater than t, $N(>t)$, against t allows an estimate of the scaling exponent c.

We are assuming the existence of a power-law fit to the data here, but we later show how attempted fits using an exponential distribution of epidemic event sizes and durations does not lead such a good description.

4. COUNTING EPIDEMIC SIZES AND DURATIONS

In the 58 years of recorded Faroe measles data there are 43 distinct epidemic outbreaks. An epidemic outbreak has a duration, t, where

$$t = \tau_{end} - \tau_{start} \tag{8}$$

and τ_{start} is the first month when cases in an event first appear and τ_{end} is the next mont when there are no more cases present. An epidemic outbreak can have a duration of 1 month up to any integer number of

Figure 2. (*a*) The average number of epidemic outbreaks occurring each calendar month. (*b*) Size of the epidemic outbreaks as a function of the calendar month in which they began. (*c*) Duration of the epidemic outbreaks (in months) as a function of the calendar month in which they began. (*d*) Average number of new measles cases to arise each calendar month.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

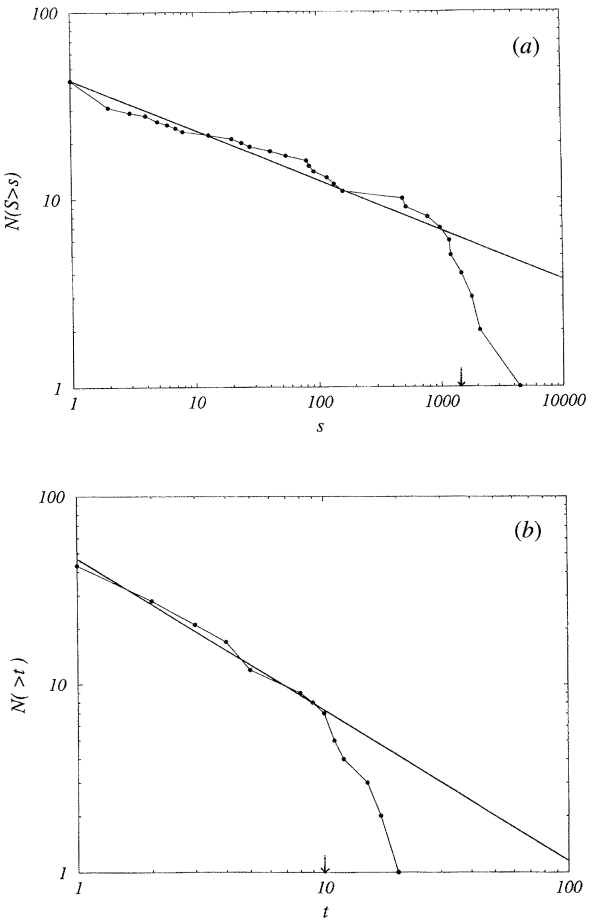


Figure 3. (*a*) Epidemic size distribution for the Fåroe Island data (log-log plot). The best-fit line shown is calculated so as only to apply to epidemics up to size 1500. The correlation for the regression $r = -0.908$. (*b*) Epidemic duration distribution for the Faroe Island data (log-log plot). The best-fit line shown is calculated so as only to apply to epidemics of duration up to 10 months, with $r = -0.99$.

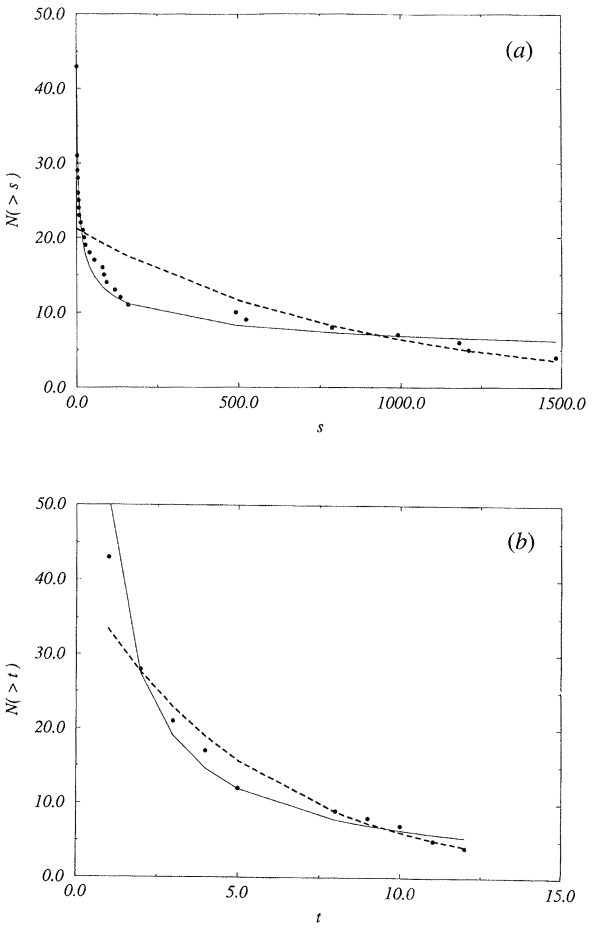


Figure 4. (*a*) Epidemic size distribution (linear plot) fitted with power-law (solid line) and exponential (dashed line) functions. (*b*) Epidemic duration distribution (linear plot) fitted with power-law (solid line) and exponential (dashed line) functions.

months. Similarly, an epidemic event has a size, s, where

$$s = \sum_{\tau_{start}}^{\tau_{end}} C(\tau) \qquad (9)$$

and $C(\tau)$ is the number of recorded cases of measles in the month τ.

In figure 3*a* is plotted the epidemic size distribution and, in figure 3*b*, the epidemic duration distribution on log-log plots. Each of the 43 epidemics is plotted on these graphs. The scaling law described by equation 4 appears to fit the epidemic size distribution very well. However, there is a tail-off for epidemics that infect more than 1500 individuals. This might happen because, for the very largest epidemics, the dynamics are fundamentally different from those of the smaller epidemics, causing either a different exponent to arise, or a complete breakdown in scaling. Alternatively, it is possible that due to the infrequent occurrence of the very largest epidemics, the Faroe Island time series is not long enough to sample a representative number of big epidemics. There is, at present, no way to distinguish between these two hypotheses. What is certainly true is that when the largest measles epidemics occurred in the islands, there would have been enough time for a modification of behaviour and mixing patterns to curtail as much as possible spread of the infection into those regions as yet unaffected by the disease.

Given the behaviour of the largest epidemics, we can plot the distribution so that all epidemics greater than 1500 cases are summed together. The gradient of the best-fit line of this plot gives a value of $b \simeq 0.27$, and this is shown in figure 3*a*.

The distribution of the epidemic durations likewise shows the agreement with scaling, but for the reasons discussed above, appears to break down for epidemics longer than 12 months duration. For epidemics longer than this there is also the possibility of other measles cases being generated by the next index case to arrive in the community, raising the possibility of running two epidemic outbreaks into each other. The best fit line shown in figure 3*b* is calculated for the epidemics up to 10 months in duration. This gives $c \simeq 0.8$.

By way of comparison, we illustrate fits of the data using power law and exponential functions. Figures 4*a* and 4*b* show the data plotted on a linear scale with the associated fits. Over the range of data shown, the power law is much the better fit for the epidemic sizes and we do not think that an exponential distribution

can be used to explain the scaling. For the epidemic durations the power law also appears to be the better fit over that of the exponential function, though it is not as clear cut.

We have also demonstrated that the scaling exponents, b and c, in the measles case returns from Bornholm and Reykjavik are of the same magnitude as the Faroe Islands (Rhodes & Anderson 1996*b*).

5. MODELLING THE EPIDEMIC DISTRIBUTIONS

Given the recent achievements in the applied epidemiological field of modelling the dynamics of measles infection (Schenzle 1984; Olsen *et al.* 1988; Olsen & Schaffer 1990; Anderson & May 1991; Grenfell 1992; Bolker & Grenfell 1993; Tidd *et al.* 1993; Grenfell *et al.* 1994; Boker & Grenfell 1995; Grenfell *et al.* 1995) we expect that it should also be possible to account for the patterns of disease in small populations. However, it turns out that the basic conventional compartmental SEIR models, in either their deterministic or stochastic implementations, are unable to capture the particular form of the epidemic distributions seen in the Faroe Island measles case returns and the possible reasons for this are discussed in more detail below. Instead, we have a lattice-based epidemic model which seems better able to reflect the dynamics.

(*a*) *A simple lattice-based model*

The power law scaling phenomena observed in the raw measles data are strongly suggestive of some underlying scale-free dynamics. Whilst many examples of scale-free patterns have been catalogued in experiments and in the natural world it has only been in recent years that it has become possible to discussion mechanisms by which such patterns may be produced (Bak *et al.* 1988; Chen *et al.* 1991; Solé & Manrubia 1995; Solé & Bascompte 1996). Many nonlinear spatially extended dynamical systems exhibit power law behaviour and we use a latter-based model, recently used in discussions of forest-fire dynamics (Bak *et al.* 1990; Grassberger & Kantz 1991; Drossel & Schwabl 1992; Mosner *et al.* 1992; Christensen *et al.* 1993; Drossel & Schwabl 1993; Grassberger 1993; Clar *et al.* 1994; Drossel & Schwabl 1994), to provide a simple model for the spread of epidemics in small isolated island populations. This is one amongst a number of lattice epidemic models that have been discussed recently, and is closely related to the model of Johansen (1994, 1996).

The earliest attempt at formulating an epidemic model using nearest-neighbour spread between individuals was undertaken by Bailey (1965, 1975), motivated by the early work on percolation theory. Mollison (1977) and Mollison & Kuulasmaa (1985) later developed these early ideas proposing his approach as a general framework with which to describe spatial heterogeneity in epidemic models. This has culminated in the spatial contact model of Cox & Durrett (1988) and Durrett & Neuhauser (1991). An extensive and insightful descriptions of this approach to epidemic modelling and the role of spatial heterogeneity in epidemiological, ecological and evolutionary systems can be found in Durrett & Levin (1994*a*, 1994*b*) and Durrett (1995), where comparisons are made between individual-based lattice models and more conventional coupled differential equation models. There are also related models of epidemic spread through populations distributed on lattices by Grassberger (1983), Cardy (1983) and Cardy & Grassberger (1985). Their approach has been developed by Boccara & Cheong (1992), who demonstrated the importance of population mixing on the rate of epidemic spirad in spatially distributed populations, and later Rhodes & Anderson (1996*a*, 1996*c*). Recently Hassell *et al.* (1991), Rohani & Miramontes (1995) and Rand *et al.* (1995) have used the lattice approach to study spatial host-parasite relationships and the role played by spatial distribution in stabilising otherwise unstable dynamics.

The lattice epidemic model we use is defined as follows, using the same notation as Drossel & Schwabl (1992); each site in the $L \times L$ lattice is in one of three states: Empty, Occupied by a Susceptible or Occupied by an Infective.

The lattice is updated synchronously at each time-step using the following rules:

(i) Susceptibles who are on nearest-neighbour sites to an Infective become Infective themselves.

(ii) Infectives become inactive and the site they occupy becomes Empty.

(iii) Susceptibles are introduced onto Empty lattice sites with a probability p. Periodic boundary conditions are used.

Essentially, these rules define a simple spatial S-I model. It is amongst the simplest possible lattice-based model of epidemic spread and, at best, is a caricature of the real epidemiological processes taking place in the population. A further rule can be added;

(iv) A new Infective can arise when a Susceptible is spontaneously infected with a probability f. This effectively correspond to an immigration term whereby our lattice population is subject to infrequent infection from external sources.

There are 43 epidemic events in the 58 years of the time series, so we can place a lower bound on the infective immigration term f of 0.74 year^{-1}. It is possible that infected index cases reach the islands more often than this, but because they do not generate any secondary infectives we cannot estimate their rate of arrival. We make the assumption, therefore, that infectives arrive on the islands as a Poisson process with a mean rate of arrival no less than f.

For the Faroe Islands population of 25,000 at equilibrium, with an average life-expectancy of 70 years, we expect approximately one new Susceptible to appear each day (*i.e.* ≈ 365 year^{-1}). This sets a lower limit to the ratio of f/p to be $0.74/365 \approx 1/493$. A simulation is run for 130 years to run off transients before data is recorded for the next 180 years. A value of $f/p = 1/300$ was actually used to get a reasonable agreement for the overall number of epidemic events,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

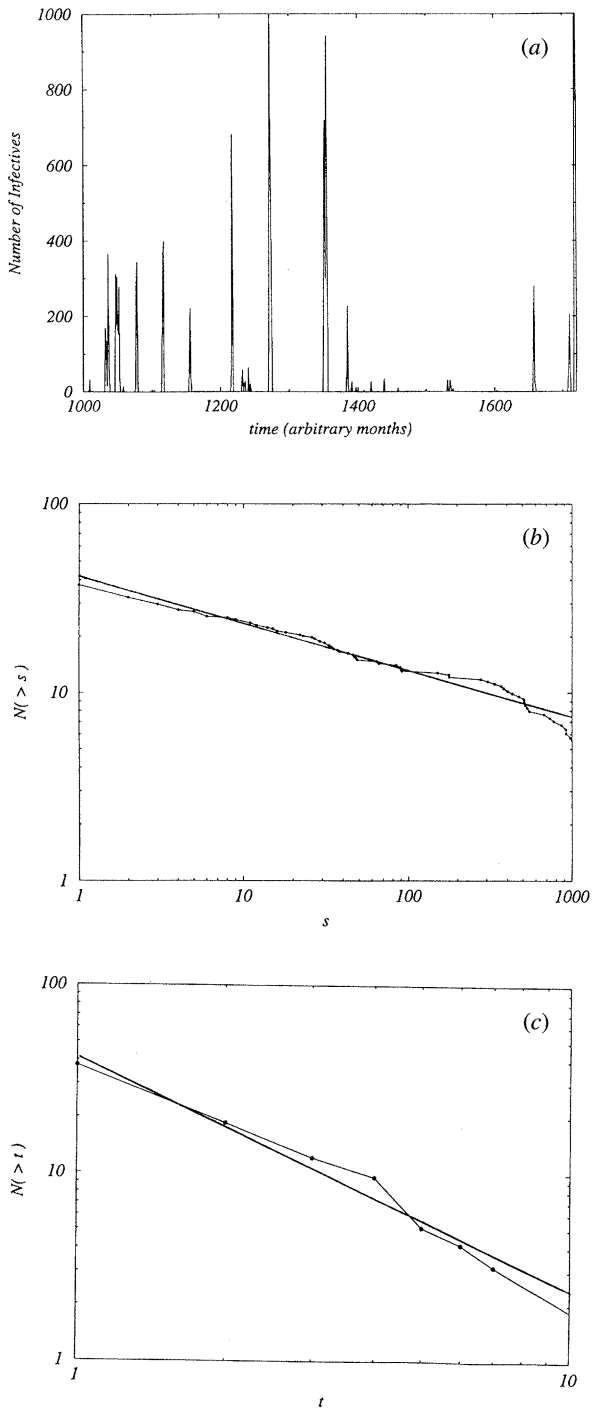


Figure 5. (*a*) Time series of the monthly for the lattice-based model. A sixty year (*i.e.* 720 month) interval of the time series is shown. Setting a lattice size $L = 250$ gives an average population of 25,000 on the lattice. (*b*) Epidemic size distribution for the lattice-based model data (log-log plot). The regression fit has $r = -0.98$. (*c*) Epidemic duration distribution for the lattice-based model data (log-log plot) with $r = -0.99$.

and the dynamics of the model in figure 5*a* closely resembles that seen in the Faroe Island data. The size of the exponent b extracted from the model data, given by the gradient of figure 5*b* is remarkably similar to the exponent derived from the actual epidemiological data; $b_{lattice} \simeq 0.25$. In the distribution of epidemic durations, the lattice model underestimates the number of larger epidemics and the resulting exponent from figure 5*c* is $c_{lattice} \simeq 1.27$. If, however, we were to calculate the gradient by using only those points that appear to be following a scaling relation (up to 4 months) we find $c_{lattice} \simeq 1.0$, which is somewhat nearer the value of c extracted from the Faroe Island data. The connectedness of the spatial distribution of the population on the lattice seems to reflect the social networks that exist in real communities and it appears essential to include this factor in order to accurately describe the epidemiological dynamics. A second important factor in generating this result is the fact (which we clearly observe in the Faroe Island data) that there is a separation of time scales for the replenishment of susceptibles and the immigration of infectives *i.e.* $f/p \ll 1$ as emphasised by Drossel & Schwabl (1992, 1993, 1994).

These simulations indicate that a lattice model seems able to capture the dynamics of epidemic spread in the population. In a sense this is not surprising, since experience with other critical phenomena tells us that the simplest reduced models of a given physical situation that contain the dominant interactions often yield the correct exponents (Binney *et al.* 1993). For this to work the system needs to be in the vicinity of a critical point and recent renormalization-group studies (Loreto *et al.* 1995) suggest that the lattice simulations we use here are close to such a critical point. Our results suggest that this idea of universality might apply in biological systems as well. Universality also tells that exponents are often robust to changes of the details of the underlying model which draws us to speculate that the exponents b and c might also apply to other simple communicable infectious diseases, such as influenza. As discussed by Durret (1995) and Solé & Bascompte (1996), the introduction of this concept into biology might prove fruitful in future work on modelling the complex dynamical systems that often arise. It seems that the lattice approach provides a natural framework in which to describe the spatio-temporal clustering of epidemics often associated with sudden outbreaks of communicable disease in susceptible populations, and we believe this is the first example of the use of a lattice-based method to provide quantitative insight into the dynamics of a real epidemiological data set.

(*b*) *A stochastic* SEIR *simulation*

The measles infection case returns for the health districts in many countries of the developed world are generally quite detailed and long-running and have been extensively analysed and compared with the results of model calculations. This work has considerably furthered the understanding of the influence of various eterogeneities, such as age-structure (Schenzle 1984; Bolker & Grenfell 1993), seasonality in contact rate (Olsen *et al.* 1988; Olsen & Schaffer 1990; Grenfell 1992) and the effects of spatially distributed populations (Bolker & Grenfell 1995), as well as contributing to the debate over the possible effect of nonlinear dynamics in biological systems (Olsen *et al.* 1988; Olsen & Schaffer 1990; Grenfell 1992; Tidd *et al.* 1993). Central to this work has been the standard compartmental SEIR model (Anderson &

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

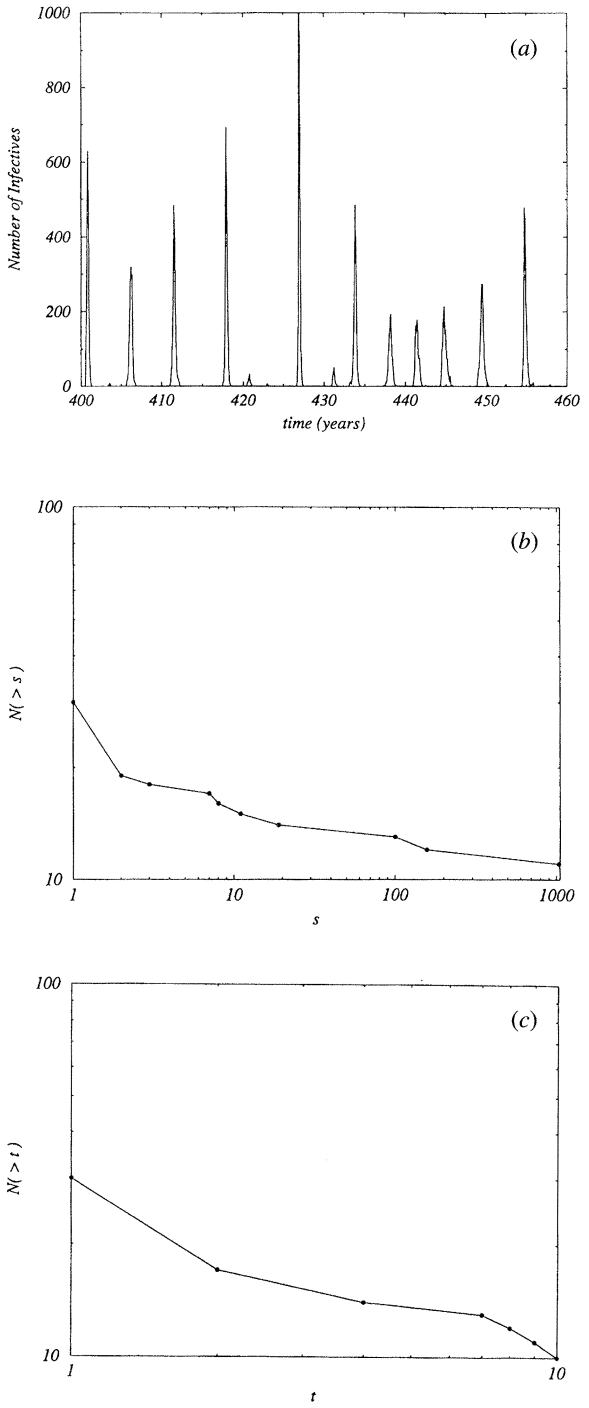


Figure 6. (*a*) Time series of the monthly for a stochastic SEIR model. $N = 25{,}000$, $\mu = 1/70$ years^{-1}, $\gamma = 1/7$ days^{-1}, $\delta = 1.7$ days^{-1} and $R_0 = 14$. (*b*) Epidemic size distribution for the SEIR model data (log-log plot). (*c*) Epidemic duration distribution for the SEIR model data (log-log plot).

May 1991). The heterogeneities thought to be particularly relevant to any epidemiological situation can be straightforwardly grafted onto this mathematical structure.

The total population under consideration is divided up into four compartments; Susceptibles (those who have not yet been infected, but are not yet infectious), Infected (those susceptibles who have been infected, but are not yet infectious), Infected (those who were previously in the exposed class and are now able to infect other susceptibles) and Recovered (those who have been infected, have now recovered and are immune to further reinfection). Susceptibles are born into the population at a certain rate, and we assume there is natural mortality from each of the compartments at a rate which maintains the overall population at a constant level. The following equations describe the time evolution of the population of each of the compartments.

$$dS/dt = \mu N - \mu S - \beta SI \tag{10}$$

$$dE/dt = \beta SI - \mu E - \gamma E \tag{11}$$

$$dI/dt = \gamma E - \mu I - \delta I + \nu \tag{12}$$

$$dR/dt = \delta I - \mu R \tag{13}$$

The total population $S+E+I+R = N$. The average life span of individuals is given by μ^{-1}, β is the contact rate between susceptibles and infectives, γ^{-1} is the average incubation period and δ^{-1} is the average infectious period. The term ν is a small immigration factor representing the occasional introduction of infectives into the system from outside, as happens in the island situation.

For the Faroe Islands, we assume a population of 25,000, an average life span to be 70 years, the incubation is taken to be 7 days and the infectious period is also 7 days. From equations 10 to 13 it is possible to define the contact rate of β in terms of a basic reproductive rate R_0

$$\beta \simeq \frac{\delta R_0}{N} \tag{14}$$

The basic reproductive rate can often be estimated from age-serological profiles (Anderson & May 1991), thus giving an indirect estimation of β. Typically, for a population like the Faroes, the reproductive rate for measles is around 14. Also, because the populations we are considering here are small it is appropriate to utilize the stochastic Monte-Carlo implementation of the above equations (Olsen *et al.* 1988).

The time-series for the monthly measles case returns from the SEIR simulation is shown in figure 6*a* for a representative 60 year interval, showing a qualitatively similarity with the time series for the island populations. Here we have set the ratio of immigration to birth rate to be identical (1/300) to that in the lattice-based simulations shown above. However, the distribution of epidemic sizes and durations shown in figure 6*b* and 6*c* do not show the scaling observed in the Faroe Island data set. Changing the immigration rate, ν, and altering the reproductive rate, R_0, does not serve to improve the agreement with the Faroe Islands distributions.

It is possible to add seasonality in the contact rate. We noted that more cases appeared during the winter months, so, following earlier studies on the effect of seasonality, we assume

$$\beta = \beta_0\{1 + \beta_1 \cos(2\pi t)\} \tag{15}$$

β_0 is given by the right-hand side of equation 14. We varied β_1 between $0 \rightarrow 0.3$, but this had negligible

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

quantitative effect on the distribution of epidemics, making only extremely small differences to the graphs in figures 6*b* and 6*c*.

Additionally, figure 2*a*, suggested a slight excess in the number of epidemics being initiated in the summer months, so we can seasonally force the immigration rate of infectives too. Choosing a cosinusoidsal forcing function as in equation 15, did little to reduce the number of larger epidemics as compared with the smaller epidemics. Thus the distributions retained a concave shape even for extremely strong forcing.

It appears that the stochastic SEIR simulations we have described are not able to accurately capture the specific form of the epidemic size and duration distributions. Consequently we are not able to theoretically confirm the scaling exponents b and c. In other calculations it turns out, the rate of immigration, ν, required to match the observed total number of epidemic events is twice that used in the lattice-based model. The reason for these discrepancies is probably connected to the fact that the SEIR equations, based as they are on homogeneous mass action based as they are on homogeneous mass action based mixing of susceptibles with infectives, overestimate the rate at which infectives can mix amongst the susceptible population. This results in an over-production in the number of larger epidemics which gives the size and duration distributions their characteristic concave shape. We have preliminary evidence which suggests that the introduction of a metapopulation structure onto the total population results in a diminution of the larger epidemics, thus leading to a somewhat straighter scaling distribution (Ferguson, unpublished data). Additionally, an extremely important heterogeneity in measles dynamics in large urban populations is age structure (Schenzle 1984; Bolker & Grenfell 1993). This is important because disease spread amongst school-age children is dominant in the dynamics and at the beginning of each academic year fresh susceptibles are introduced into the school system to be exposed to infection. In contrast, the smaller Faroe Island population measles epidemics are not of a long duration and person-to-person spread is the dominant route for transmission rather than through cohorts of susceptibles entering a school environment which results in the maintenance of endemic infection. In larger cities measles is usually regarded as a disease of childhood which most children will get at some time, whereas in largely susceptible island communities, measles outbreaks afflict all age-groups. Consequently we have not allowed for age-structured transmission in either the lattice or the SEIR models, though a age-structured SEIR model (Ferguson, unpublished data) has been show to behave little differently than a homogeneous model.

6. ESTIMATING EPIDEMIC OUTBREAK FREQUENCIES

The observation of the power laws also allows us to estimate how often epidemics of certain sizes and durations might occur in a given period of time. To calculate the average number of epidemics, E, that (in a 58 year interval) infect, for example, more than 10 individuals but less than 100 individuals we integrate equation 5 between the lower limit, s_l and the upper limit s_u.

$$E = \int_{s_l}^{s_u} a s^{-1-b} ds \tag{16}$$

so

$$E = \frac{a}{b}[s_l^{-b} - s_u^{-b}] \tag{17}$$

For the Faroe data, $b = 0.28$ and $\frac{a}{b} = 43$, so

$$E = 43[s_l^{-0.28} - s_u^{-0.28}] \tag{18}$$

If $s_l = 10$ and $s_u 9 = 100$, then the number of epidemics we expect between these limits in the time series, $E = 10.7$. For the Faroe data, the observed number of epidemics in the size range $E_{obs} = 9$, so the agreement is quite good. In the event that the population parameters remained the same for the Faroes we can say that in the next 60 years we would expect there to be around eleven distinct epidemics that affect greater than 10 people and less than 100 people. The same analysis can be done for the average number of epidemics in a given range of durations.

Whilst we are able to make estimate of the frequencies of epidemics in given size and duration bounds, it is not possible to say when (or where) such epidemics will occur (rather like earthquake prediction), so this is a rather weak form of time series prediction.

7. DISCUSSION

Our results suggest the existence of well defined scaling laws for the size and duration of measles epidemic outbreaks in the Faroe Islands. The distributions are better fit with power law functions than exponential ones. Though we do not discuss it in this paper, measles case returns for Bornholm and Reykjavik show similar scaling. This places the dynamics of Bartlett Type III measles epidemics in small isolated populations, subject to infrequent infection from outside, in the same class as other spatially extended non-linear dynamical systems, where scaling is also observed. These data sets present a picture of what happens when a largely susceptible population is perturbed by the occasional introduction of infection. On the whole, small short duration epidemics predominate with fewer large, long duration epidemics. In practice, this facilitates a form of prediction in which we can calculate the frequency of occurrence of epidemics of given size and duration. Basic forms of the conventional stochastic SEIR model are shown to overestimate the number of large epidemics and cannot provide reasonable estimates of the scaling exponents. However, a simple spatial model can generate similar exponents to those seen in our data analysis.

At present it is not possible to verify if the power laws are a general phenomenon in the dynamics of measles infection. The exponents, b and c, have been shown to

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

be remarkably similar in three different geographical locations (Rhodes & Anderson 1996*b*) but it is not yet known if they can be applied in all epidemiological contexts. Most reasonably accurate measles data are for large populations in urban centres of the developed world. However, with high vaccination coverage in most developed countries, such populations now contain insufficient susceptibles to maintain chains of transmission without introduction of infection by immigrants or visitors. The lower law phenomenon is therefore likely to be of relevance in both the study of infrequent outbreaks of measles in highly vaccinated communities and in isolated low density rural populations in developing countries. We believe that lattice based models of infectious disease open up many new lines of research on how the behaviour of individuals influences the pattern of infection and disease in populations that cannot be easily addressed using conventional mathematical models of disease spread based on population aggregations or compartments. Our analysis suggests that very simple behavioural and biological rules induce measurable epidemiological patterns and hence provides a new tool for epidemiological prediction and interpretation.

We thank L. F. Olsen of the Institute of Biochemistry, Odense University, Denmark for allowing use of the Faroe Island measles data and B. T. Grenfell of the Department of Zoology, Cambridge University for many useful discussions. Research support for C.J.R. and R.M.A. is kindly provided by the Wellcome Trust.

REFERENCES

Anderson, R. M. 1994 The Croonisan Lecture, 1994. Populations, infectious disease and immunity: a very nonlinear world. *Phil. Trans. R. Soc. Lond.* B **346**, 457–505.

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans, dynamics and control*. Oxford University Press.

Bailey, N. J. T. 1965 The simulation of stochastic epidemics in two dimensions. *Proc. 5th Berkeley Symp. on Math. Statist. and Prob.* **4**, 237–257.

Bailey, N. J. T. 1975 *The mathematical theory of infectious diseases and its applications*. Oxford University Press.

Bak, O., Chen, K. & Tang, C. 1990 A forest-fine model and some thoughts on turbulence. *Phys. Lett.* A **147**, 297–300.

Bak, P. & Sneppen, K. 1993 Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* **71**, 4083–4086.

Bak, P., Tang, C. & Weisenfeld, K. 1988 *Self-organised criticality*. *Phys. Rev.* A **38**, 364–374.

Bartlet, M. S. 1957 Measles periodicity and community size. *J. R. statist. Soc.* A **120**, 48–70.

Bartlett, M. S. 1960 The critical community size for measles in the U.S. *J. R. statist. Soc.* A **123**, 37–44.

Biunney, J. J., Dowrick, N. W., Fisher, A. J. & Newman, M. E. J. 1993 *The theory of critical phenomena*. Oxford: Clarendon Press.

Black, F. L. 1966 Measles endemicity in insular populations: critical community size and its evolutionary implications J. THEO. BIOL. **11**, 207–211.

Boccara, N. & Cheong, K. 1992 Automata network models for the spread of infectious diseases in a population of moving individuals *J. Phys. A: Math. Gen.* **25**, 2447–2461.

Bolke, B. M. & Grenfell, B. T. 1993 Chaos and biological complexity in measles dynamics. *Proc. R. Soc. Lond.* B **251**, 75–81.

Bolke, B. M. & Grenfell, B. T. 1995 Space, persistence and dynamics of measles epidemics *Phil. Trans. R. Soc. Lond.* B **348**, 309–320.

Cardy, J. 1983 Field theoretic formulation of an epidemic process with immunisation. *J. Phys. A: Math, Gen.* **16**, L709–712.

Cardy, J. & Grassberger, P. 1985 Epidemic models and percolation. *J. Phys. A: Math. Gen.* **18**, L267–271.

Chen, K., Bak, P. & Obukhov, S. P. 1991 Self-organised criticality in a crack-propagation model of earthquakes. *Phys. Rev.* A **43**, 620–635.

Christensen, K., Flyvberg, H. & Olami, Z. 1983 Self-organised critical forest-fire model: mean-field theory and simulation results in 1 to 6 dimensions. *Phys. Rev. Lett.* **71**, 2737–2740.

Clar, S., Drossel, B. & Schwabl, F. 1994 Scaling laws and simulation results for the self-organised critical forest-fire model. *Phys. Rev.* E **50**, 1009–1018.

Cliff, A., Haggett, P. & Smallman-Raynor, M. 1993 *Measles. An historical geography of a major human viral disease from global expansion to local retreat 1840–1990*. Oxford, Blackwell.

Cox, J. T. & Durrett, R. 1988 Limit theorems for the spread of epidemics and forest fires. *Stoch. Proc. Applics.* **30**, 171–191.

de Boer, J., Derrida, B., Flyvberg, H., Jackson, A. D. & Wettig, T. 1994 Simple model of self-organised biological evolution. *Phys. Rev. Lett.* **78**, 906–909.

Drossel, B. & Schwabl, F. 1992 Self-organised critical forest-fire model. *Phys. Rev. Lett.* **69**, 1629–1632.

Drossel, B. & Schwabl, F. 1993 Forest-fire model with immune trees. *Physica* A **199**, 183–197.

Drossel, B. & Schwabl, F. 1994 Formation of space-time structure in a forest-fire model. *Physica* A **204**, 212–229.

Durrett, R. 1995 Spatial epidemic models. In *Epidemic models: their structure and relation to data*. (ed. D. Mollison). Publications of the Newton Institute, Cambridge University Press.

Durrett, R. & Levin, S. A. 1994*a* The importance of being discrete (and spatial). *Theor. Popul. Biol.* **46**, 363–394.

Durrett, R. & Levin, S. A. 1994*b* Stochastic spatial models: a user's guide to ecological applications. *Phil. Trans. R. Soc. Lond.* B **343**, 329–350.

Durrett, R. & Neuhauser, C. 1991 Epidemics with recovery in d = 2. *Ann. Appl. Prob.* **1**, 189–206.

Ferguson, N., May, R. M. & Anderson, R. M. 1996 Metapopulation models of measles epidemics. In *Spatial ecology: the role of space in population dynamics and interspecific interactions*. (ed. D. Tilman & P. Kareiva) Princeton Monographs in Population Biology. (In the press.)

Fine, P. E. M. & Clarkson, J. A. 1982 Measles in England and Wales – 1: an analysis of factors underlying seasonal patterns. *Int. J. Epidem.* **11**, 5–15.

Grassberger, P. 1983 On the critical behaviour of the general epidemic process and dynamic percolation. *Math. Biosc.* **63**, 157–172.

Grassberger, P. 1993 On a self-organised critical forest-fire model. *J. Phys. A: Math. Gen.* **26**, 2081–2089.

Grassberger, P. & Kantz, H. 1991 On a forest-fire model with supposed self-organised criticality. *J. Stat. Phys.* **63**, 685–700.

Grenfell, B. T. 1992 Chance and chaos in measles dynamics. *J. R. statist. Soc.* B **54**, 383–398.

Grenfell, B. T., Bolker, B. M. & Kleczkowski, A. 1995 Seasonality and extinction in chaotic meta-populations. *Proc. R. Soc. Lond.* B **259**, 97–103.

Grenfell, B. T., Kleczkowski, A., Ellner, S. & Bolker, B. M. 1994 Measles as a case-study in non-linear forecasting and chaos. *Phil. Trans. R. Soc. Lond.* A **348**, 515–530.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Hassell, M. P., Comins, H. N. & May, R. M. 1991 Spatial structure and chaos in insect population dynamics. *Nature, Lond.* **353**, 255–258.

Johansen, A. 1994 Spatio-temporal self-organisation in a model of disease spreading. *Physica* D **78**, 186–193.

Johansen, A. 1996 A simple model of recurrent epidemics *J. theo. Biol.* **178**, 45–51.

Loreto, V., Pietronero, L., Vespignani, A. & Zapperi, S. 1995 Renormalisation-group approach to the critical behaviour of the forest-fire model *Phys. Rev. Lett.* **75**, 465–468.

May, R. M. 1976 Simple mathematical models with very complicated dynamics. *Nature, Lond.* **261**, 459–467.

May, R. M. 1987 Chaos and the dynamics of biological populations. *Proc. R. Soc. Lond.* A **413**, 27–44.

Mollison, D. 1977 Spatial contact models for ecological and epidemic spread. *J. R. statist. Soc.* B **39**, 283–326.

Mollison, D. & Kuulasmaa, K. 1985 Spatial endemic models: theory and simulations. In *The Population dynamics of rabies in wildlife*. (ed. P. J. Bacon) Academic Press, New York, 292–309.

Mossner, W. K., Drossel, B. & Schwabl, F. 1992 Computer simulations of the forest-fire model. *Physica* A **190**, 205–217.

Mullin, T. (ed.) 1993 *The Nature of chaos*. Oxford University Press.

Olsen, L. F. & Schaffer, W. M. 1990 Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science, Wash.* **249**, 499–504.

Olsen, L. F., Truty, G. L. & Schaffer, W. M. 1988 Oscillations and chaos in epidemics: a nonlinear dynamic study of six childhood disease in Copenhagen, Denmark. *Theo. Popul. Biol.* **33**, 344–370.

Rand, D. A., Keeling, M. & Wilson, H. B. 1995 Invasion stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies. *Proc. R. Soc. Lond.* B **259**, 55–63.

Rhodes, C. J. & Anderson, R. M. 1966*a* Dynamics in a lattice epidemic model. *Phys. Lett.* A **210**, 183–188.

Rhodes, C. J. & Anderson, R. M. 1966*b* Power laws governing epidemics in isolated populations. *Nature, Lond.* **381**, 600–602.

Rhodes, C. J. & Anderson, R. M. 1966*c* Persistence and dynamics in lattice models of epidemic spread. *J. theo. Biol.* **180**, 125–133.

Rohani, P. & Miramontes, O. 1995 Host-parasitoid meta-populations: the consequences of parasitoid aggregation on spatial dynamics and searching efficiency. *Proc. R. Soc. Lond.* B **260**, 335–342.

Schenzle, D. 1984 An age-structured model of pre and post vaccination measles transmission. *IMA J. Math. appl. med. Biol.* **1**, 169–191.

Solé, R. V. & Bascompte, J. 1966 Are critical phenomena relevant to large-scale evolution? *Proc. R. Soc. Lond.* B **263**, 161–168.

Solé, R. V. & Manrubia, S. C. 1995 Are rainforests self-organised in a critical state? *J. theo. Biol.* **173**, 31–40.

Sornette, A. & Sornette, D. 1989 Self-organised criticality and earthquakes. *Europhys. Lett.* **9**, 197–202.

Sughihara, G., Greenfell, B. T. & May, R. M. 1990 Distinguishing error from chaos from measurement error in time-series. *Phil. Trans. R. Soc. Lond.* B **330**, 235–251.

Sughihara, G. & May, R. M. 1990 Non-linear forecasting as a way of distinguishing chaos from measurement error in time-series. *Nature, Lond.* **344**, 734–741.

Tidd, C. W., Olsen, L. F. & Schaffer, W. M. 1993 The case for chaos in childhood epidemics, II. Predicting historical epidemics from mathematical models. *Proc. R. Soc. Lond.* B **254**, 257–273.

Received 18 June 1996; accepted 13 August 1996

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES