

PIMS Distinguished Chair Lectures

BRYAN GRENFELL

PIMS Distinguished Chair
University of Alberta
September, 2003

*Modeling the Dynamics
of Infectious Diseases*

Edited by Frithjof Lutscher

Modeling the Dynamics of Infectious Diseases

A series of lectures by Bryan Grenfell

*Department of Zoology
University of Cambridge
Cambridge, UK*

Lecture notes by Frithjof Lutscher

*Department of Mathematical and Statistical Sciences
University of Alberta
Edmonton, Canada*

Abstract

Infectious diseases continue to have a major impact on individuals, populations, and the economy, even though some of them have been eradicated (e.g. small pox). Unlike many other ecological systems, many infectious diseases are well documented by spatio-temporal data sets of occurrence and impact. In addition, in particular for childhood diseases, the dynamics of the disease in a single individual are fairly well understood and fairly simple. As such, infectious diseases are a great field for mathematical modeling, and for connecting these models to data. In this article, we concentrate on three issues, namely (1) comparative childhood disease dynamics and vaccination, (2) spatio-temporal disease dynamics, and (3) evolution in diseases with multiple strains. The mathematical techniques used in the analysis of disease models contain bifurcation theory for ODEs, wavelet analysis, stochastic simulations and various forms of data fitting.

These lecture notes stem from a series of four lectures by Bryan Grenfell as a PIMS distinguished visitor at the University of Alberta and the University of Calgary in the fall of 2003. While none of the individual results presented here is new (and we will give ample reference), the summary and comparative view is meant to give a brief introduction and overview of a subject that has been a hot research topic in recent years.

The following is not meant as an exhaustive review of the relevant literature. More references and technical information are given on:

<http://www.zoo.cam.ac.uk/zoostaff/grenfell/publications/publications.htm>

General introductions to the subject are given by Anderson and May (1981) and Grenfell and Dobson (1995).

1. Childhood Disease Dynamics and Vaccination

In this first section, we model the dynamics of measles in detail. Measles is a highly infectious childhood disease, caused by respiratory infection by a morbillivirus, measles virus (MV). The disease exhibits clearly non-stationary patterns in time series of cases (Bjørnstad et al., 2002). As an example, Figure 1 shows weekly recordings of cases of measles in England and Wales. There is a transition from biennial cycles before vaccination to annual cycles after vaccination. More recently, an overall decline with more irregular patterns can be observed. At first, we are interested how much biological detail has to be incorporated into a mathematical model to accurately capture the observed dynamics. It turns out that a relatively simple SIR model with seasonal forcing, corresponding to the school year, is an adequate description, and that variations in birth and vaccination rate explain the switching between different cycle periods. We then turn to pertussis, a childhood disease caused by a bacterial infection. Although similar to measles in many features, pertussis exhibits fairly different dynamics. The presented modeling approach explains these differences. Finally, we compare vaccination in childhood diseases to vaccination in animal diseases.

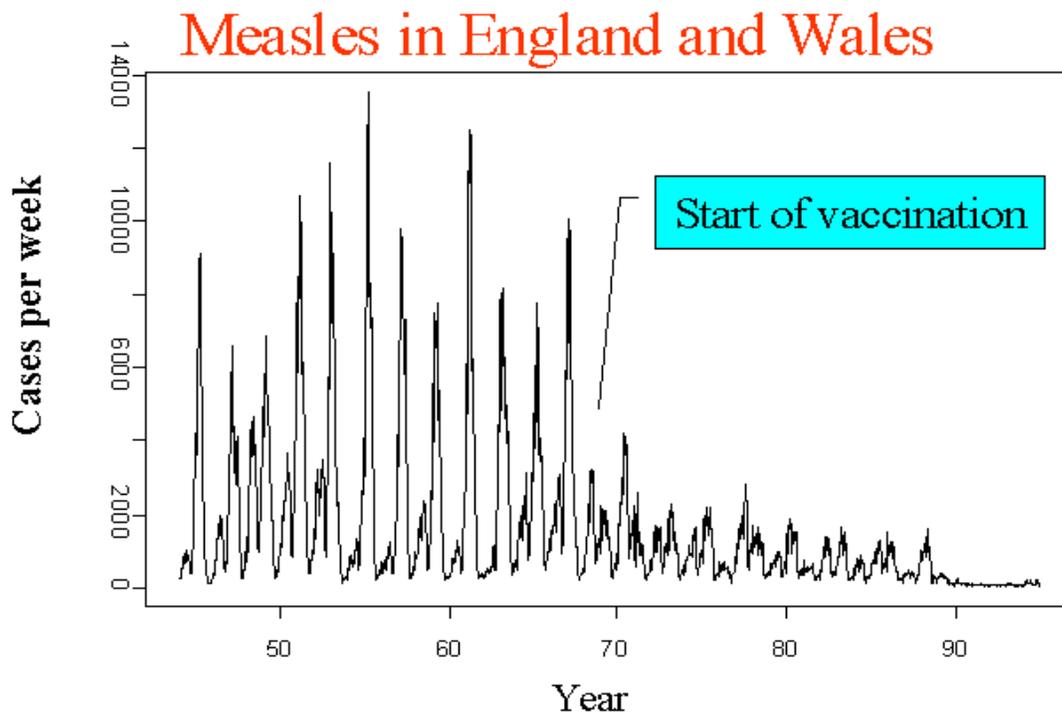


Figure 1: Weekly recording of cases of measles in England and Wales. Until the late 60s, there is a strong biennial cycle with high amplitude. After vaccination, annual cycles with decreased amplitude dominate. Eventually, the number of cases declines and the pattern becomes irregular.

1.1 The epidemic oscillator

When modeling measles dynamics, or those of any other microparasitic disease, the number or density of individuals in each stage of the disease is the important quantity to keep track of whereas the viral load per person is relatively unimportant. Following the classical approach, the population is divided into a susceptible, an infectious (and infected), and a recovered class, denoted by S , I and R , respectively. Individuals contract the disease at a rate $\beta I/N$ and recover from it with rate γ . With its very fast infection process and lifelong immunity after recovery, a single outbreak of measles is probably the only disease to which this most simple SIR model can be applied. To model the course of several years, two changes are necessary. The contact rate β is allowed to vary in time, and births (at rate ν) as well as deaths (at rate μ) are introduced. The total population $N=S+I+R$ is held constant by setting $\nu=\mu$ for now. All newborns are susceptible, vaccination is discussed later. The model equations are given by

$$\begin{aligned}\frac{dS}{dt} &= \nu N - \mu S - \frac{\beta(t)IS}{N} \\ \frac{dI}{dt} &= \frac{\beta(t)IS}{N} - (\mu + \gamma)I \\ \frac{dR}{dt} &= \gamma I - \mu R\end{aligned}$$

During the course of a single season ($\nu=\mu=0$, β constant), the model shows the following simple dynamics as plotted in Figure 2. We assume that, at first, all individuals are susceptible, and a single infectious individual is introduced into the population. Then the number of infectious individuals grows, initially exponentially at rate R_0 , and the number of susceptibles declines. Once the infection reaches its maximum, both the susceptibles and infectious decline. The bilinear contact rate $\beta IS/N$ accounts for R_0 being constant across different population sizes, which is actually observed. For measles, R_0 ranges between 15 and 18.

The most important driving force in repeated measles outbreaks over many years is the seasonal forcing in the contact rate β . During the school year, the contact rate is high, during the holidays it is low. This form of transmission rate is called “term-time forcing” and was modeled most famously in an age-structured model by Schenzle (1984). Whereas these periodic changes have sometimes been modeled by a sinusoidal function, a piecewise constant function seems to approximate the true contact rate better and is consequentially used in all of the following. The epidemic oscillator results from an interplay of seasonal forcing, which triggers a single outbreak as described above, and the birth process, which slowly increases the number of susceptibles between two outbreaks.

Anatomy of a measles epidemic

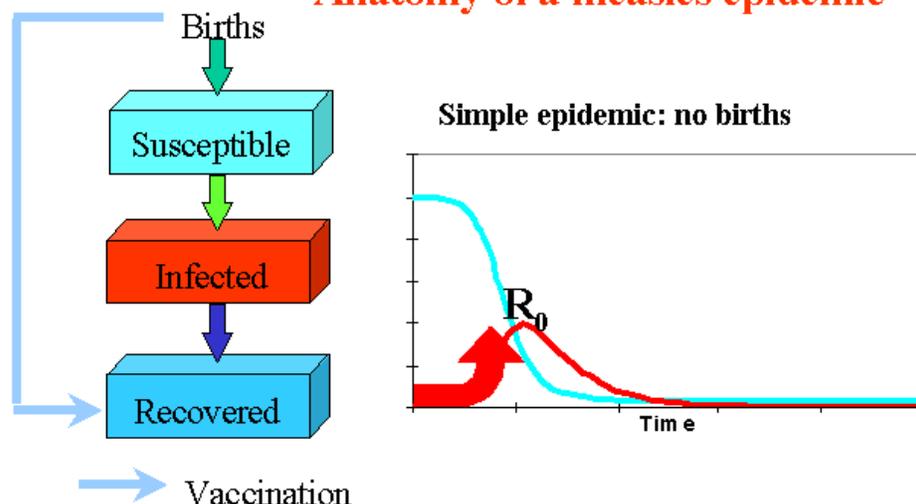


Figure 2: On the left, schematic representation of the SIR model with vaccination. On the right, the course of a single epidemic without population dynamics is plotted, susceptibles are given in light blue, infecteds in red.

The epidemic oscillator produces periodic outbreaks and hence seems to capture the most important ingredients for modeling measles dynamics. However, in the time series in Figure 1, we see changes in outbreak period and irregular non-stationary patterns. In the following, this nonstationarity is analyzed in more detail by use of wavelets. For later reference, we note that vaccination effectively reduces the birth rate since it moves individuals directly into the recovered class (see Figure 2).

1.2 Quantifying Nonstationarity

Time series data, such as cases of measles in Figure 1, contain components of different period and frequency that change in time. Classical Fourier analysis, which works well for stationary patterns, is not appropriate for the analysis required here. Even windowed Fourier analysis (see references 23-26 in Grenfell et al., 2001) does not give satisfying results. Instead, wavelet theory provides the necessary tools in time series analysis. The following account is based on Torrence and Compo (1998). We first define the wavelet used here, the so-called **Morlet** wavelet, explain the technique using a simple example, and then apply the analysis to several time series of measles.

To detect smooth changes in cycle period, the continuous complex Morlet wavelet is used. This wavelet consists of a sinusoidal function (complex exponential) that is damped by a Gaussian function:

$$\Psi_0(\eta) = \pi^{-1/4} \exp(i\omega_0\eta) \exp(-\eta^2 / 2),$$

where $\omega_0 = 6$. The continuous wavelet transform (CWT) of a time series x_n , $n=0 \dots N-1$, with time step δt is now given by the convolution of x_n with a scaled and translated version of Ψ_0 that is localized in time and frequency space. At scale s and time n we have

$$W_n(s) = \sum_m^{N-1} x_m \Psi_0 \left(\frac{(m-n)\delta t}{s} \right).$$

The power of the wavelet is given by $|W_n(s)|^2$. The wavelet transform is normalized to have unit energy across the scale s . For the Morlet wavelet, scale is almost equal to wavelength. In Figure 3, a simple artificial signal is given that changes from a lower frequency to a higher frequency. The wavelet is plotted for two different values of the scale that correspond to the respective frequencies of the signal. On the right, the power of the wavelet is plotted as a function of scale and time. Initially, the power is highest for the scale corresponding to the low frequency or long period, then it changes and the power is highest for the smaller scale, corresponding to the shorter period or higher frequency.

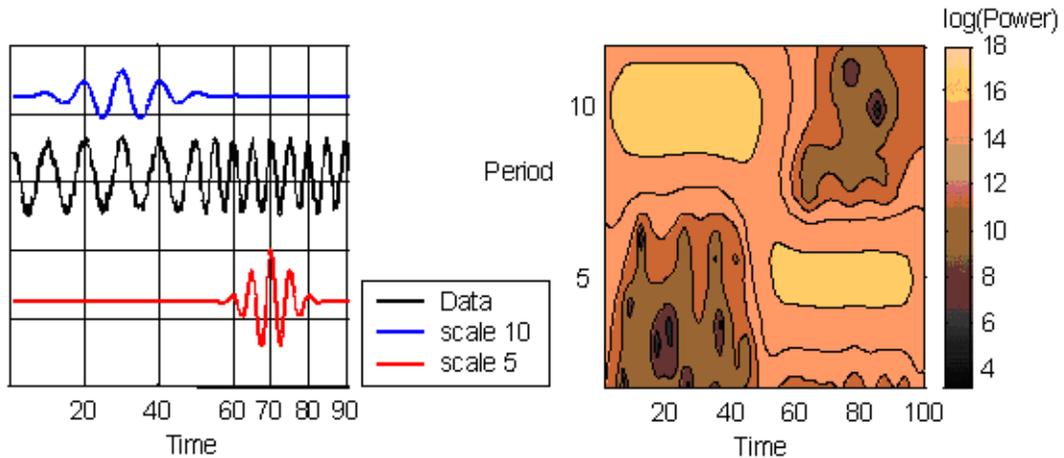


Figure 3: A simple signal that changes frequency and the corresponding wavelets of scale 5 and 10. On the right, the power of the wavelet $W_n(s)$ is plotted. Initially, the period 10 has the highest power, after the signal switches to the higher frequency, period 5 has the highest power.

We apply this wavelet transform now to the measles data from London (Grenfell et al., 2001) and from New York. At first, the data are $\log(x+1)$ -transformed in order to dampen the large amplitudes, then the wavelet transform is applied and the power as a function of scale (period) and time is plotted. For the London data (Figure 4), we see that the biennial period is dominating mostly until the start of vaccination, except for the very beginning

of the time series. As the power of the biennial pattern diminishes after vaccination, there is a short period in which the annual and then the triennial pattern have the most power, and at the end, there is no clear pattern visible. By contrast, the data from New York (Figure 5) show continually high power in the annual pattern. In addition, between 1940 and 1945, a strong triennial pattern emerges that then turns into a biennial pattern. Note that the New York series also shows 2-3 year power before 1940.

Now that we have found a way to quantify and visualize non-stationarity, we are interested in its causes. How does it arise, can it be predicted, and, ultimately, can it be controlled?

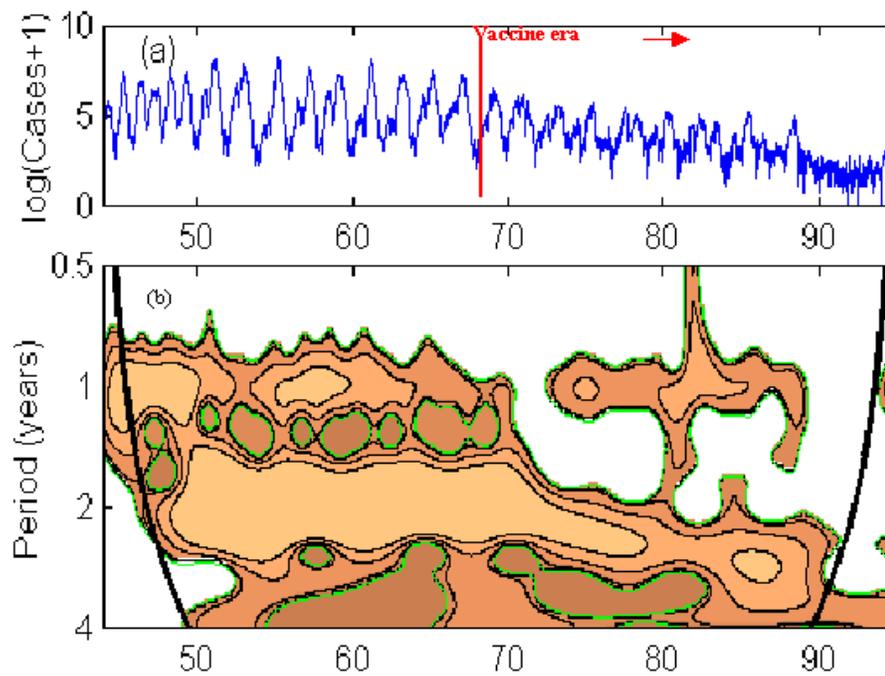


Figure 4: The wavelet power spectrum for the London measles data. For explanation, see text above (from: Grenfell et al. 2001, copyright permission from Nature Publishing Group, <http://www.nature.com>).

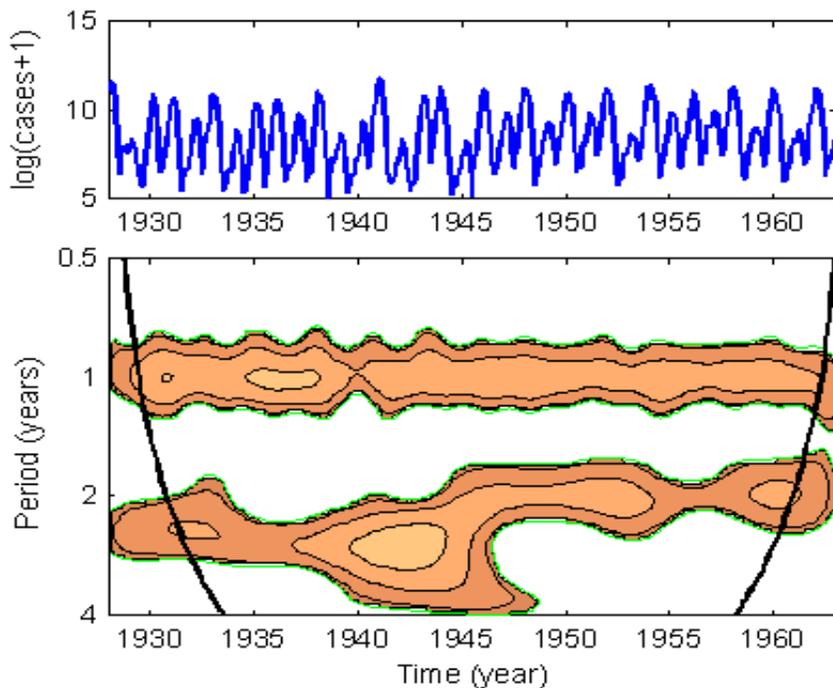


Figure 5: The wavelet power spectrum for the data from New York.

1.3 Explaining Nonstationarity

The highly irregular patterns seen in the measles data from New York have attracted attention from many researchers in dynamical systems theory. The system can be modeled as a periodically forced oscillator. As the seasonality is increased, the models show a cascade of period doubling bifurcations, eventually leading to chaotic behavior (Schaffer, 1985). In this section, we focus on alternative *mechanistic* explanations of the irregularities and non-stationary trends, derived from the underlying important biological features of the process. In particular, we focus on birth rate variations as an external factor rather than internal disease dynamics.

Birth rates, indeed, varied in England during the time that we are interested in, see Figure 6. There was a baby boom in the late 40s and the early 60s in the UK. As mentioned above, births are the way by which the number of susceptibles after a measles infection is increased again, before the next outbreak can occur. Together with the seasonality in the infection rate, the birth rate is the driving force of the epidemic oscillator.

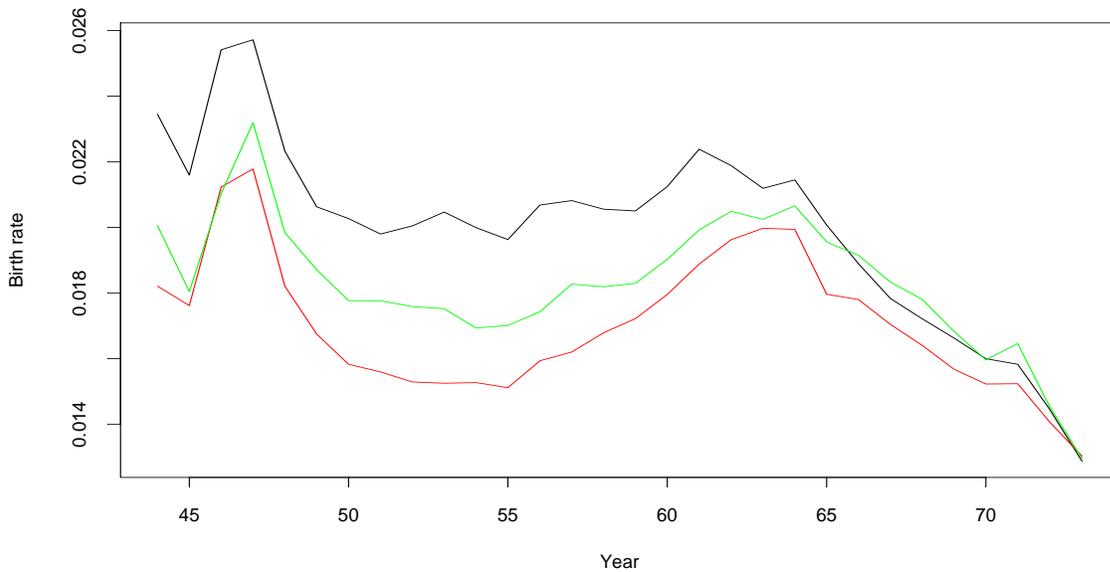


Figure 6: Birth rates in England between 1945 and 1972. The red curve is for London, the green for Manchester, the black one for Liverpool. We can clearly see the baby boom in the 40s and the smaller baby boom in the 60s.

In order to use this information in a model, the crucial observation is that the flow into and out of the susceptible compartment is completely known, that is, birth rates and cases of infections are known. Therefore, the number of susceptibles is known up to a constant. With this knowledge, one can build a discrete time-series model to estimate seasonality in the infection rate (Finkenstädt and Grenfell, 2000; Bjørnstad et al., 2002; Grenfell et al., 2002). It turns out that this simple model captures the change in cycle period nicely. The time step for the discrete time model is chosen to be two weeks in correspondence with the time scale of the disease. Then the seasonality is estimated by fitting the time series SIR-model (TSIR)

$$I(t + 1) = \beta(t)I(t)^\alpha S(t)^\gamma$$

to the data by using one-step ahead predictions as described in Finkenstädt and Grenfell (2000). The exponents α and γ allow for the possibility of spatial or other departures from mass action transmission, as well as discretization of the continuous time infection process. Figure 7 shows that the fit works quite well, and the 20-year simulation for the pre-vaccination era indicates that the model does capture the change from the one-year cycles at the beginning to two-year cycles later on.

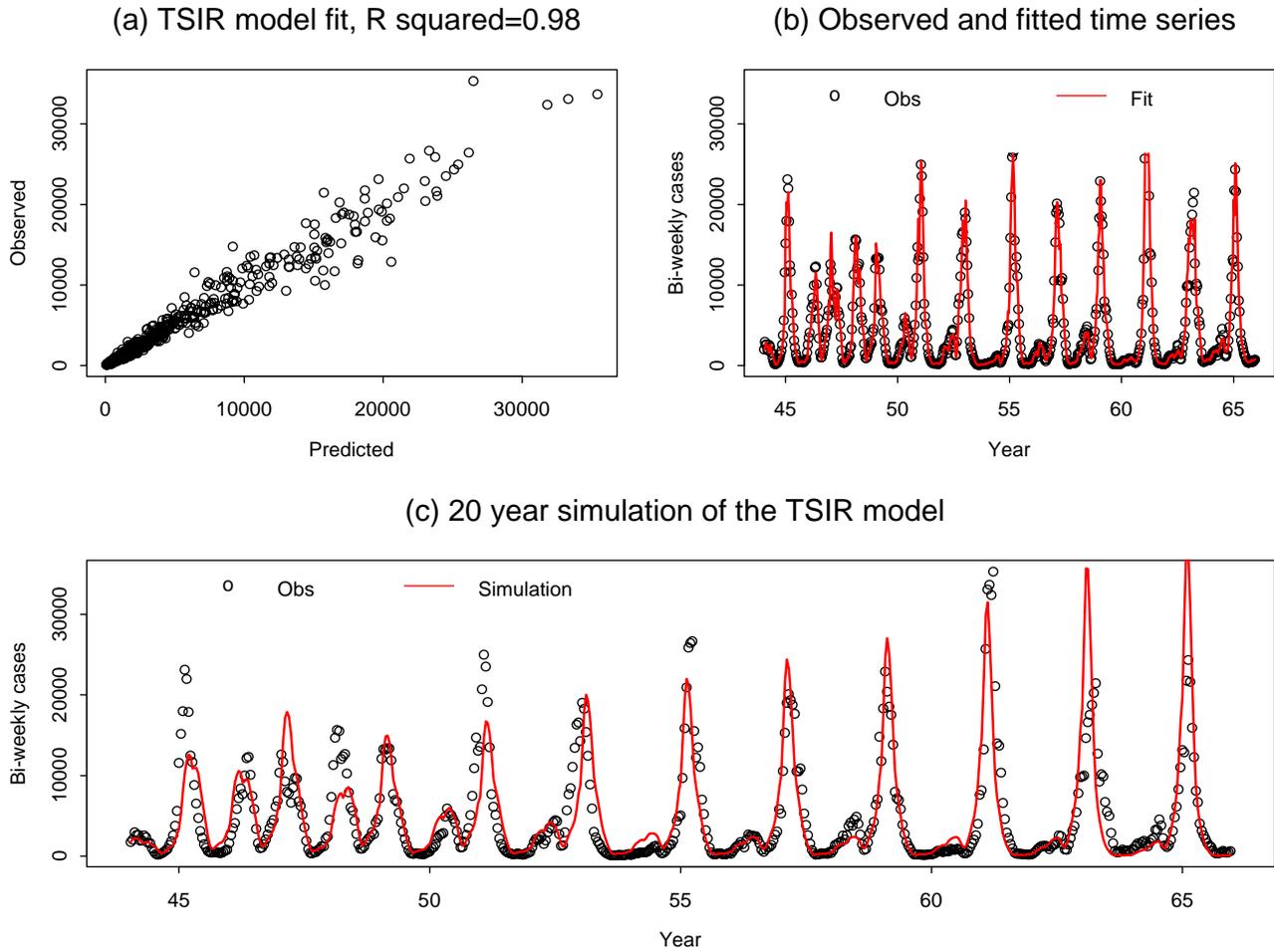


Figure 7: TSIR fit and simulation. The plot in (b) is obtained by one step ahead predictions, whereas (c) is a prediction of the full series based on the fit from (b). More detailed fits as in (a) for different cities can be found in Figure 3 in Bjørnstad et al. (2002).

While not perfect, this simple mechanistic model with varying birth rate and seasonality in transmission demonstrates that the birth rate as an external force explains the change in cyclicity sufficiently well. In Figure 8, the 20-year TSIR simulation is plotted again with superimposed birth rate.

All these simulations were done with the TSIR model, i.e., in discrete time, and it is therefore difficult to get rigorous analytical results. In the next section, we work with differential equations again as we look at the effects of varying birth and vaccination rates on cycle period.

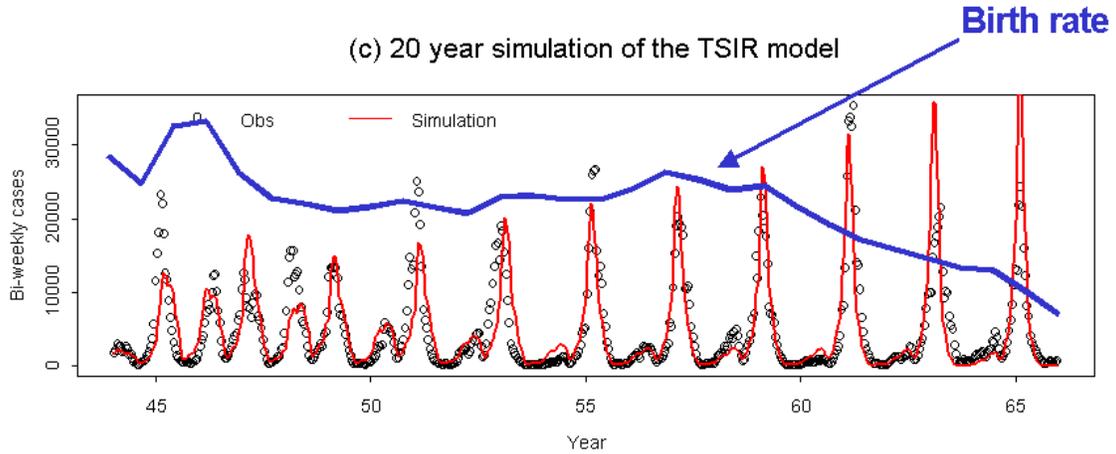


Figure 8: TSIR simulation and birth rate

1.4 External driving forces

From the simulations in the previous section, we see that high birth rates drive measles dynamics to annual cycles, decreasing birth rates lead to biennial patterns. Looking at the original data (Fig. 1), we see that vaccination leads to smaller amplitude and more irregular cycles. As mentioned before, birth rate as well as vaccination influence the recruitment to the susceptible compartment. Therefore it seems conceivable that their effects can be combined in a single variable. Using an appropriate scaling, one can look at variations in birth and vaccination rates as variations of a single parameter. This work is described in detail in Earn et al. (2000).

We start with a slightly more elaborate model for the measles epidemic, namely, we introduce an *exposed* compartment. (This affects the dynamics only qualitatively.) In the simplest form after scaling, the equations read

$$\begin{aligned}\frac{dS}{dt} &= \nu - (\beta I/N + \mu)S, \\ \frac{dE}{dt} &= \beta IS/N - (\sigma + \mu)E, \\ \frac{dI}{dt} &= \sigma E - (\gamma + \mu)I, \\ \frac{dR}{dt} &= \gamma I - \mu R,\end{aligned}$$

compare the simple model in section 1.1. Introducing vaccination of a certain proportion p of all newborns changes the equations for S and R to

$$\frac{dS}{dt} = \nu(1 - p) - (\beta I + \mu)S,$$

$$\frac{dR}{dt} = \nu p + \gamma I - \mu R,$$

whereas the equations for E and I remain the same. Changing variables to

$$S = \hat{S}(1 - p), E = \hat{E}(1 - p), I = \hat{I}(1 - p), R = \hat{R}(1 - p) + (\nu / \mu)p,$$

gives back the original SEIR model with β replaced by $\beta(1-p)$. Therefore, changes in birth rate as well as proportion vaccinated act together and can be expressed as a simple scaling of the transmission rate.

Figure 1 in Earn et al. (2000) shows the proportion of infected individuals at the attractor, plotted in a bifurcation diagram with bifurcation parameter $\langle \beta \rangle$, the average transmission rate. For high average transmission rate, there is a unique annual attractor. As the transmission rate decreases, biennial cycles occur, at further reduction, 3-, 4-, 5-, 6-, 7-, and 8-year cycles appear, sometimes simultaneously. For very low mean transmission rate, only the annual attractor remains.

This means that high birth rates and the absence of vaccination lead to annual cycles in measles. As the birth rate decreases and/or mass vaccination is introduced, more complex patterns are observed as several local attractors coexist. It seems likely that demographic stochasticity leads to switching between attractors, the observations may appear chaotic.

In Figure 9 and 10, the cases of new infections are plotted at the same time as the birth rate and the recruitment rate into the susceptible compartment. After the onset of vaccination, the recruitment rate is obviously lower than the birth rate. The colored bars on the top indicate the predicted attractors according to the bifurcation diagram for the given mean transmission rate. The single green line indicates the high transmission rate regime, for which there is a unique annual attractor. Two blue lines correspond to the existence of a biannual attractor, several lines in different colors correspond to coexisting attractors of different cycle length.

The data for London (Figure 9) show that between the baby boom before 1950 and the onset of vaccination in the late 1960s, the birth rate is more or less constant and low enough for the biennial attractor to appear. After the onset of vaccination, combined with a decline in the birth rate, the system moves into the regime of several coexisting attractors. In contrast, the birth rate in Liverpool is much higher than in London, such that we see an annual attractor there almost until the onset of vaccination. The steep drop in birth rate combined with vaccination brings the effective recruitment rate in Liverpool down to the same level as in London, and several local attractors coexist.

In the USA, the patterns are somewhat different. Birth rates are low during the great depression and then increase between 1930 and 1960, which results in measles shifting from irregular patterns in regimes of multiple coexisting attractors to biannual (New York) or even annual (Baltimore) regimes, see Figure 10.

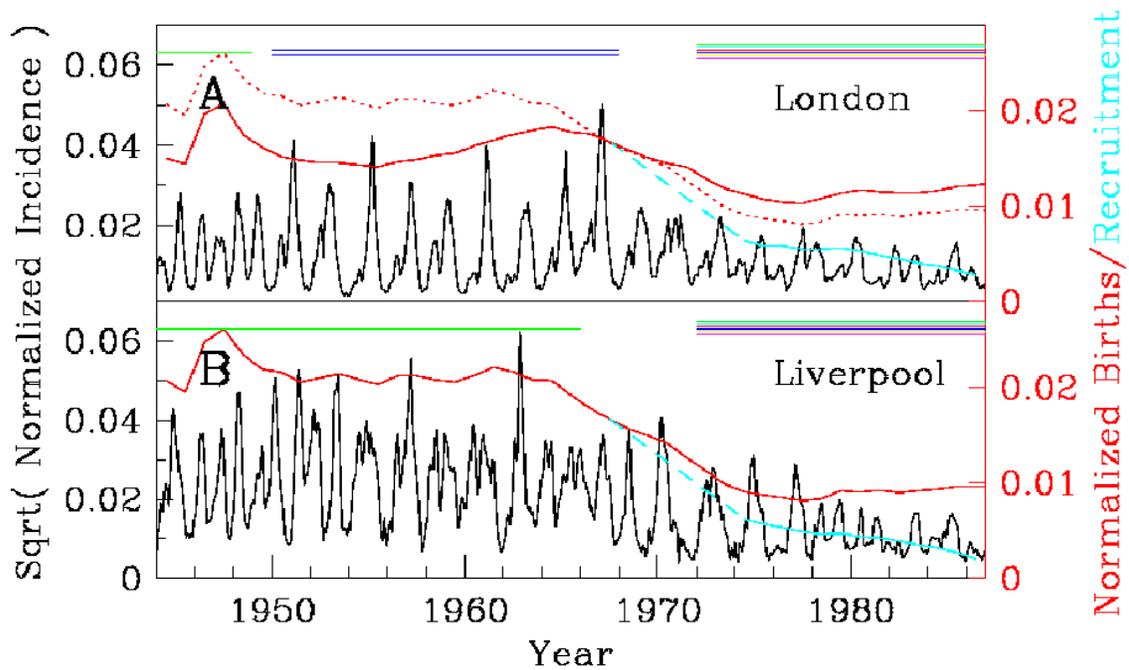


Figure 9: Measles, birth rate and attractors in London and Liverpool (First published in: Earn et al., 2000, Science)

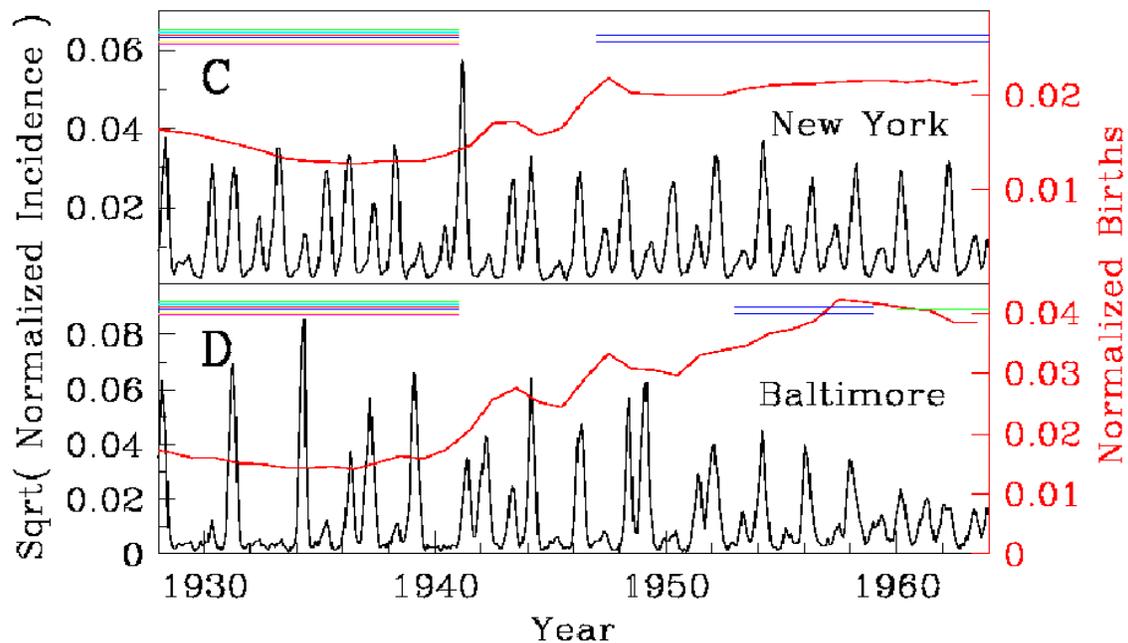


Figure 10: Measles and birth rates in New York and Baltimore (First published in: Earn et al., 2000, Science)

1.5 Comparing different childhood diseases

In the previous sections, we developed a simple SIR model for measles and found that the variation in birth- and vaccination rates explain the data well. In this section, we apply the model to a different childhood disease, whooping cough, and compare the model outcome to data and to the results for measles. The details are given in Rohani et al., (1999).

From a modeling point of view, whooping cough, or pertussis, is quite similar to measles in that it is mostly contracted by children and that it follows a similar course of infection and (usually) recovery. Also, the basic reproductive numbers for both diseases are almost equal. The main difference is that the infectious period in whooping cough is longer than in measles by a factor of 2 to 3. This difference leads to very different outcomes in model prediction and collected data. There is also significant loss of immunity in older people in pertussis, though this does not affect the following results qualitatively.

In Figure 11, the reported cases of measles and pertussis in England are plotted for comparison. The upper plot is for measles, the lower for pertussis. In each plot, the total number of cases is plotted as a black curve on top on a logarithmic scale. The vaccination rate is plotted in red. Reported cases are plotted for different cities, starting with London at the top as the largest city and then progressively moving towards smaller cities below. Dark red means high number of cases, blue corresponds to low numbers, white means absent.

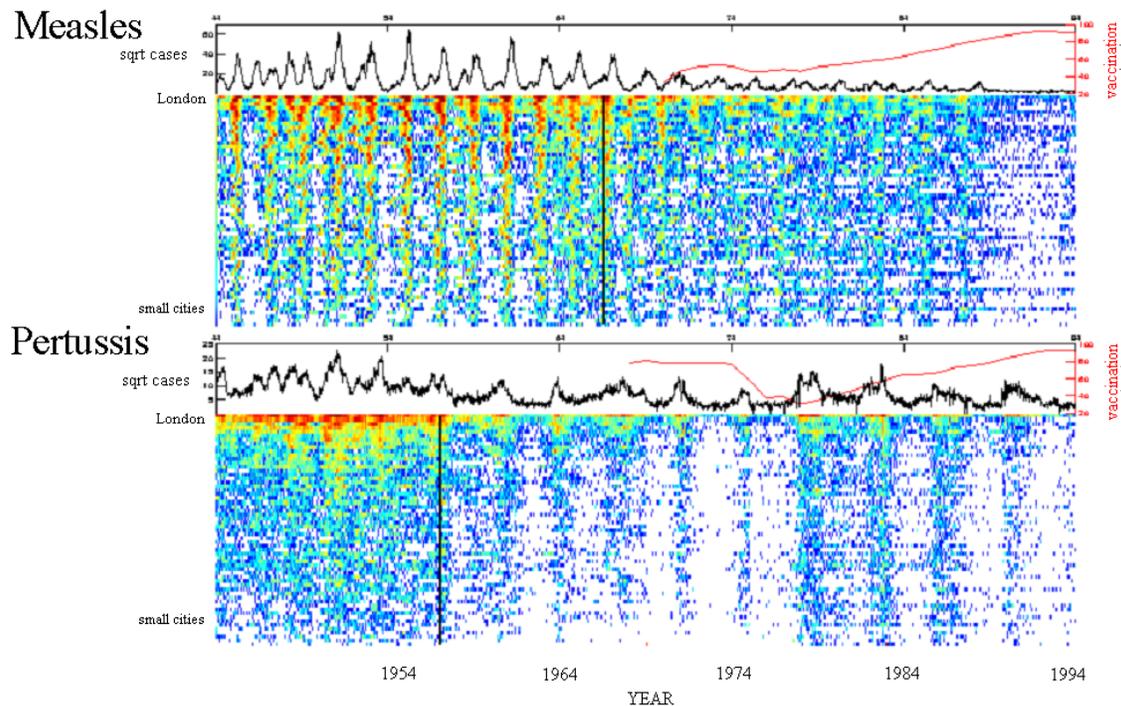


Figure 11: Measles and Pertussis in England, semi-spatial plot (First published in: Rohani et al., 1999, Science)

The upper plot for measles repeats what we know from above and gives new insight about spatial synchronization. Prior to vaccination, we see regular cycles that are also synchronized in space over most cities. With vaccination, the number of cases decreases, the patterns become irregular in time and also desynchronized in time. We will study these changes in detail in Section 2. The lower plot in Figure 11 for pertussis shows a completely different picture. Prior to vaccination, cases of pertussis are irregular and not spatially synchronized. With vaccination, the number of cases drops but not as significantly as with measles. More importantly, patterns become synchronized in space and also turn into fairly regular 3-4 year cycles. It should be noted that the simplistic SIR model from above with parameters for pertussis predicts that the attractor consists of annual cycles, given the current birth and vaccination rates. How can this discrepancy be explained?

Several authors have worked on detailed models for pertussis (Hethcote, 2002; Rohani et al., 2000). Here, we give a descriptive rather than detailed answer to the above question. The general idea is that the seasonal forcing and disease dynamics in measles exhibit much stronger resonance and are hence approaching the attractor very fast whereas the resonance in pertussis is much weaker and solutions take much longer to approach the attractor. This effect can be seen by numerical perturbation experiments, plotted in Figure 12. The experiment consists of introducing a large perturbation to the periodic dynamics on the attractor for a given parameter set, and recording the transient back to the attractor. In this case, the perturbation was the removal of 50% of all infectives. In the simulation for measles (upper plot), the effect on the solution is barely visible. A slight irregularity results, but the yearly cycles persist. The scenario is quite different for pertussis. There is a long transient before the system settles back on the attractor of annual dynamics, and during these transients, 3-4 year cycles appear.

Taking into account that demographic stochasticity continually introduces perturbations of the system away from the attractor, it is conceivable that we observe 3-4 year cycles in pertussis as long-time transients to the annual cycle attractor. In particular in the troughs between outbreaks, the system is sensitive to stochastic events.

Another approach to capture the difference between measles and pertussis is to look at the way an invasion of the disease takes until it settles on the attractor. This is done in Figure 13. Starting from a very small number of infectives, the system is simulated, and the number of susceptibles and infectives is taken at one fixed time every year. (The first ten years are discarded as initial transients.) This process is repeated with different numbers of initial infectives. Plotting these points in phase space gives the orbit of an invasion of the disease, or, in more mathematical terms, the unstable manifold of the disease free equilibrium.

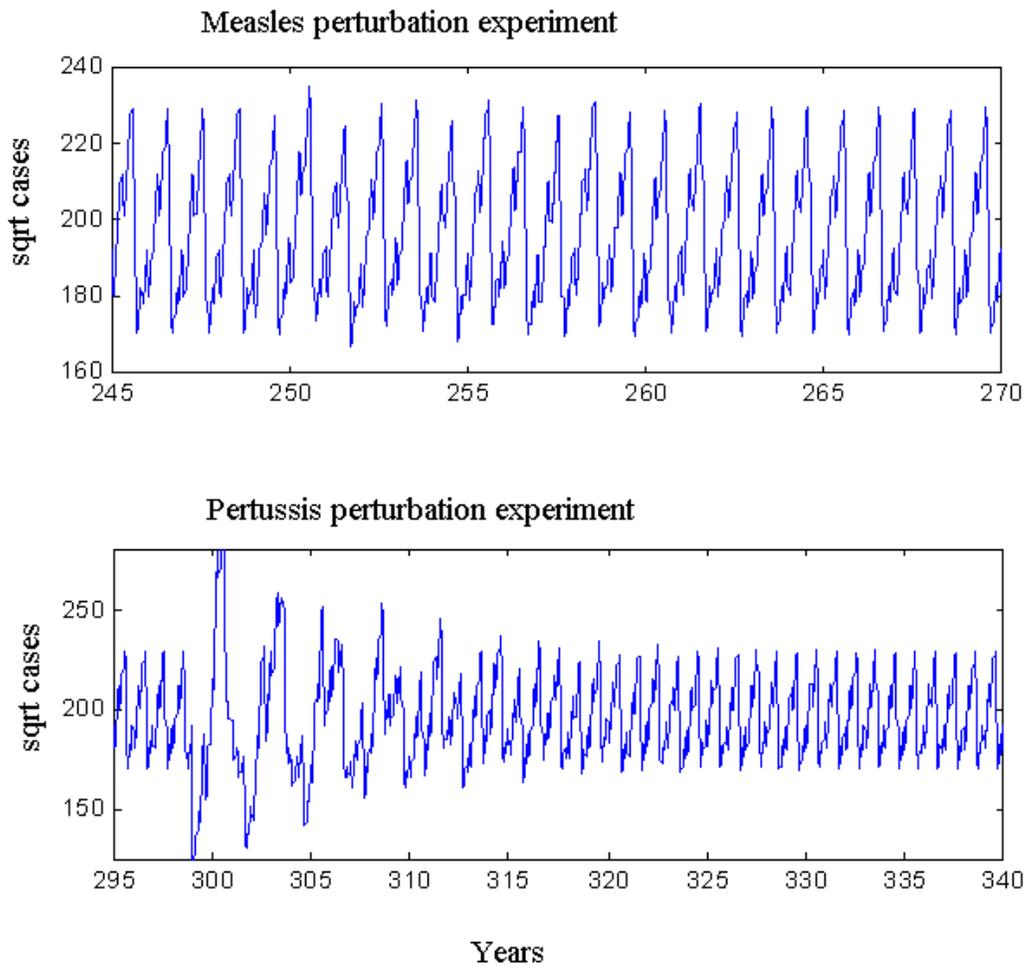


Figure 12: Perturbation experiments for measles and pertussis. Compare to Figures 4,5 in Rohani et al. (2002).

The following can be observed from Figure 13. For measles, the time series of infectives on the right shows that the system eventually exhibits biannual cycles. These correspond to two fixed points in the phase-space plot on the left. For pertussis, the system reaches an equilibrium that corresponds to a single fixed point in phase space. For a more detailed discussion of the invasion orbit and clearly visible four-years cycles, we refer to Rohani et al. (2002).

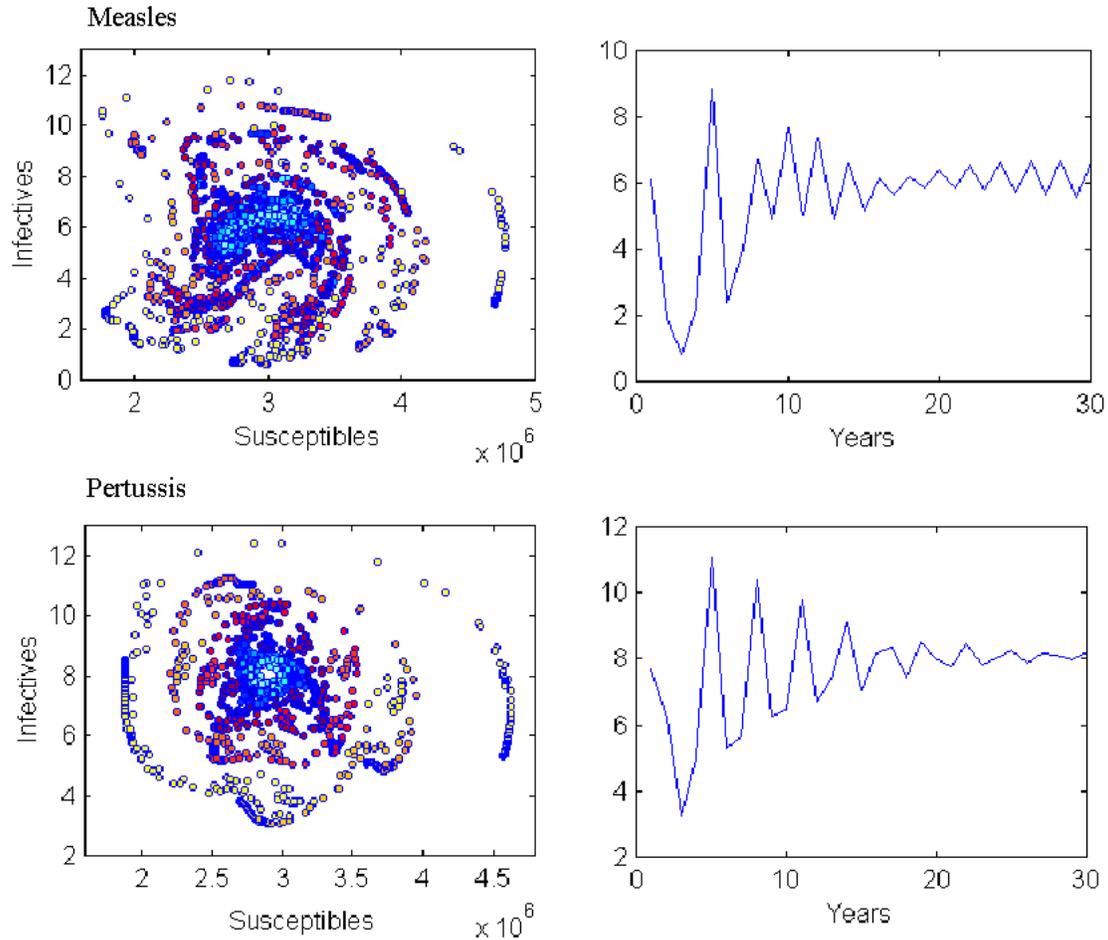


Figure 13: Invasion orbits for measles and pertussis. Compare to Figures 7,8 in Rohani et al. (2002).

To summarize, measles and pertussis are childhood diseases with similar characteristics and similar basic reproductive number R_0 but with quite different observed dynamics. From a mathematical perspective, the duration of the infectious period in measles, the build-up of susceptible individuals in the population and the seasonal forcing form a system with strong resonance such that solutions are tightly locked onto the attractor, and even large perturbations do not change the behavior significantly. In contrast, the longer infectious period in pertussis leads to a much weaker resonance, and perturbations can move solutions far off the attractor. Even though the attractor may consist of annual cycles, we observe 3-4 year cycles as transients to that attractor. From a more ecological perspective, we note that in the simple SIR system, the basic reproductive number R_0 is approximately given by the product of the average transmission probability β and the average infective period $1/\gamma$. Since the infective period for measles is shorter, the transmission probability must be higher to achieve the same R_0 . Therefore, during school term the transmission probability is very high, but the two week break over Christmas corresponds to a full generation of the disease during which the transmission probability is low. Pertussis, on the other hand, has a lower transmission probability all along, but the

short break over Christmas does not interrupt its transmission as severely since the break is shorter than the disease generation time.

So far, all the models considered have been non-spatial, and truly spatial models will be discussed below in Section 2. However, the data in Figure 11, cases sorted by city size, suggest to try a semi-spatial approach in form of a metapopulation-patch model. Every city is considered as a patch, in which the SIR model is used to describe the disease dynamics (see Rohani et al, 1999, for details). The simulation in Figure 14 shows that such models are able to capture the difference between measles and whooping cough dynamics quite well. For measles, the dynamics change from regular periodic, spatially correlated outbreaks to irregular uncorrelated occurrences after vaccination. In pertussis, the situation is reversed.

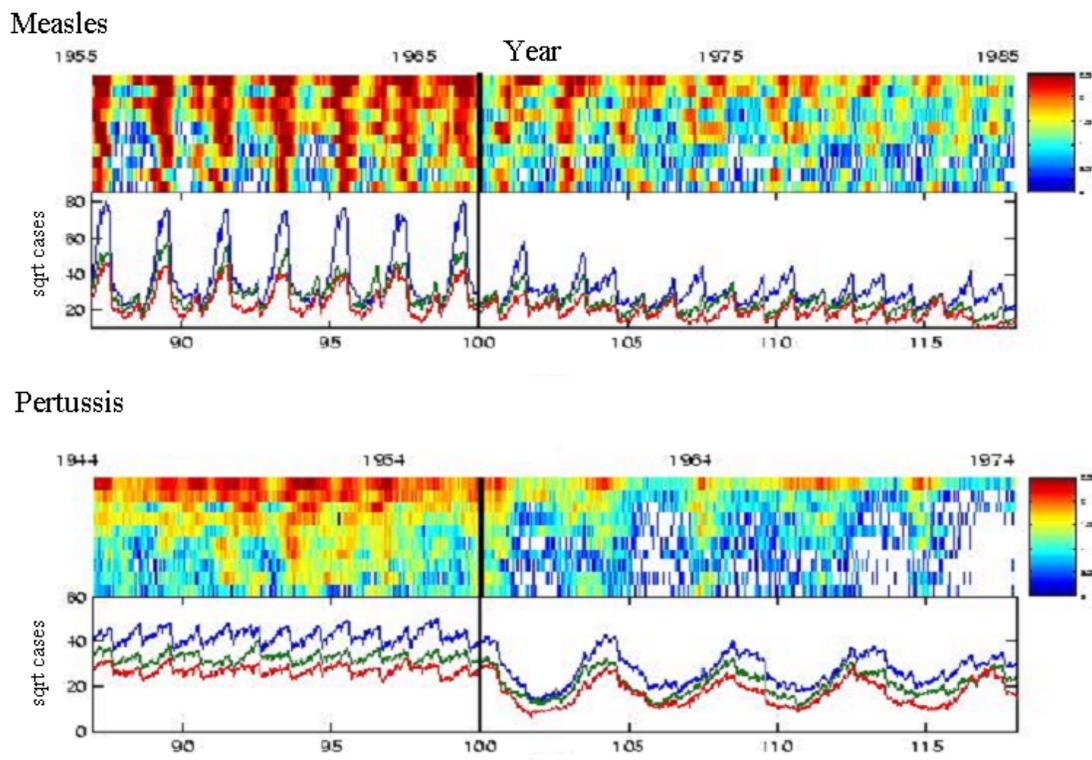


Figure 14: Metapopulation-patch model simulation (First published in: Rohani et al., 1999, Science)

1.6 Control by Vaccination

In this section, we explore how and under which circumstances outbreaks of diseases can be prevented by vaccination, see Adnerson and May (1991) for more details. We illustrate the ideas using the simple SIR model from section 1.1 above and assume that all parameters are constant. The equation for infectives is given by

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \alpha I,$$

where β is the transmission rate, and $\alpha = \gamma + \mu$ is the combined removal rate from the infectious class given by recovery and natural mortality. The basic reproductive rate R_0 is the number of secondary infectives that a single infective produces in a totally susceptible population. In a totally susceptible population, we have $S = N$, and therefore, a single infective produces β new infectives per unit time. The average time spent in the infective compartment is $1/\alpha$, and therefore the basic reproductive rate is given by

$$R_0 = \beta / \alpha.$$

The epidemic can invade if R_0 is greater than one but not if R_0 is less than one. When a population is not totally susceptible, for example due to acquired immunity from a previous epidemic or due to vaccination, then the *effective value of R_0* , denoted by R_{eff} , is reduced by the fraction of susceptibles in the population, i.e.,

$$R_{eff} = R_0 \frac{S}{N} = \frac{\beta}{\alpha} \frac{S}{N}.$$

The number of infectives grows if R_{eff} is greater than one and declines if R_{eff} is less than one. Schematically, this is given in Figure 15 below. The critical proportion of susceptibles, sc , is hence given by the condition that R_{eff} equals one. If immunity is to be achieved by vaccination only, then the critical level of vaccination, p_c , is given by

$$1 = R_0(sc) = R_0(1 - p_c) \quad \text{or} \quad p_c = 1 - 1/R_0.$$

Therefore, not everybody has to be vaccinated in order to prevent an outbreak of a disease. This phenomenon is called *herd immunity*. In Figure 16, p_c is plotted as a function of R_0 . It is clear that with increasing R_0 it is more and more difficult to achieve the critical level of vaccination. Whereas eradication has been successful for smallpox, the required vaccination rate for measles is much higher.

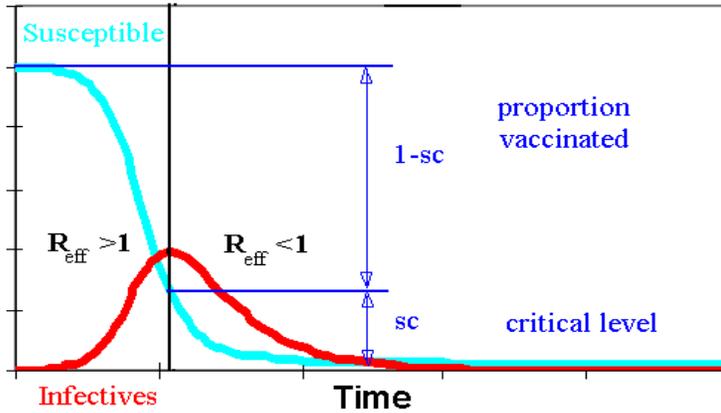


Figure 15: Schematic plot of the effective reproductive number and the critical level of susceptibles

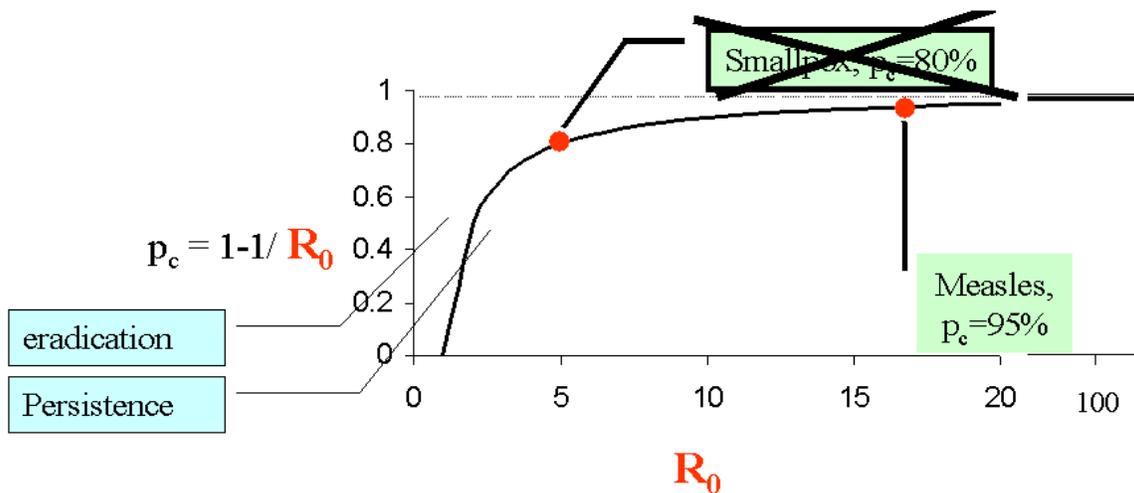


Figure 16: The critical level of vaccination as a function of the basic reproductive number. Whereas smallpox requires that only roughly 80% of the population are vaccinated in order to obtain herd immunity (and has been eradicated), measles would require a much higher level.

We have derived exact conditions for the critical level of vaccination to eradicate a disease that follows a simple deterministic SIR model. We expect the analysis and the resulting conditions to be more difficult in a complex stochastic setting. The most detailed analysis of spatio-temporal vaccination patterns in stochastic models was carried out during the outbreak of foot and mouth disease in the UK in 2001, during which the farming industry in the UK was devastated (Keeling et al., 2001). Surprisingly, simple extensions of the above theory work quite well to predict the impact of vaccination in this complex, heterogeneous spatio-temporal epidemic. This is because mass vaccination

effectively ‘damps down’ stochastic fluctuations at all spatial scales, so that the R_0 threshold (in this case for two hosts, sheep and cattle) works well.

2. Spatio-temporal patterns in measles

Most of the theory in the first section did not take into account the spatial nature of the measles epidemics. In this section, we focus on inherently spatial phenomena, characterize spatial patterns, suggest some mechanisms for these patterns to occur and mention some implications for vaccination.

Going back to the bar graph in Figure 11, we see that the regular temporal dynamics of measles before the vaccination era are accompanied by strong spatial synchronization across the country from big to small cities. After the onset of vaccination, not only the temporal cycles disappear but also the spatial synchrony seems to vanish. Note also that stochastic fade out is much more common after the onset of vaccination.

The reported cases of measles from 60 cities in England and Wales from 1944 to 1990 have been made into a movie that was shown during the talk. It is available from the [PIMS website](#) as a [Streaming Video](#) of Bryan Grenfell’s second lecture at the university of Alberta. In that movie, it can be seen that

- measles outbreaks are highly spatially synchronized in England and Wales before the onset of vaccination
- within each outbreak there are spatial waves away from London and away from the northwest, and there are waves along the coast
- the region Norwich/East Anglia is consistently out of phase with the rest of the country from the mid 1940s until the late 1960s.

2.1 Characterizing spatial patterns

When we try to characterize the observed spatial patterns, we need to isolate the major epidemics and look at the variations around these, thereby allowing for changes in cyclicity. As mentioned above, wavelet analysis is preferable to (windowed) Fourier analysis for our purposes. Note that we chose a complex wavelet (Morlet wavelet, see above), which allows us to get information about the amplitude *and* phase of the epidemic. To study spatial synchrony and waves, information about their phase is essential. How can information from wavelet analysis be used to describe the observed patterns?

Just as with Fourier series, one can reconstruct the observational data by integrating the computed wavelets appropriately (full details are given in Grenfell et al., 2001). This is shown in the first plot in Figure 17. But since we know that the measles epidemic is driven by the annual cycles of seasonal forcing and, at least over a long time, a biennial cycle needed to build up the critical number of susceptibles, we focus on these components. The second and third plot in Figure 17 show the reconstruction of cases given by integrating the annual and biennial components, only. We see that prior to 1950, the biennial component is fairly weak, but then dominates the observations until the late 1960s when it starts to decline. Compare also Figures 4 and 8.

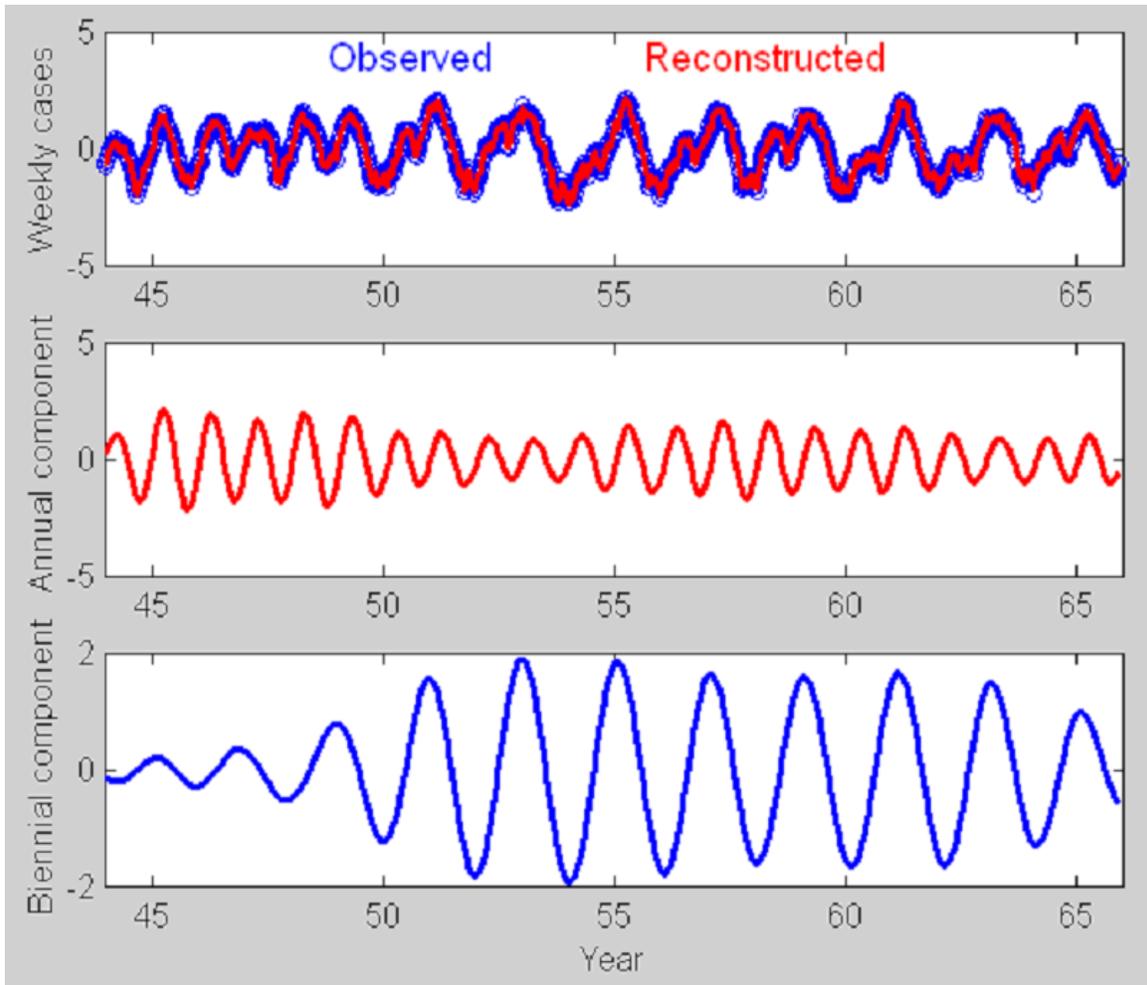


Figure 17: Reconstruction of observations from the wavelet spectrum. Top plot: Integrating the whole wavelet spectrum. Middle plot: Integrating only the annual component. Bottom plot: integrating only the biennial component.

2.2 Two cities out of phase

We now apply the reconstruction of the biennial component described above to the observational data from two cities, Cambridge and Norwich, between 1944 and 1966. During this (pre-vaccination) period, both cities show clear biennial measles epidemics, but until the mid 1950s, the phases of the epidemics differ by 180 degree, i.e., the outbreak in the one city happens in a year when there is no outbreak in the other. In the late 1950s, a switch occurs and the epidemics are in phase thereafter. The observed data are plotted in Figure 18.

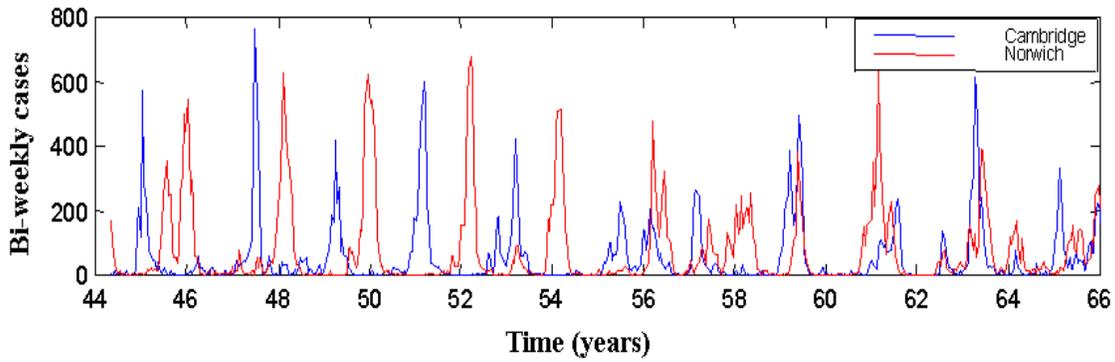


Figure 18: Bi-weekly recordings of measles in Cambridge and Norwich.

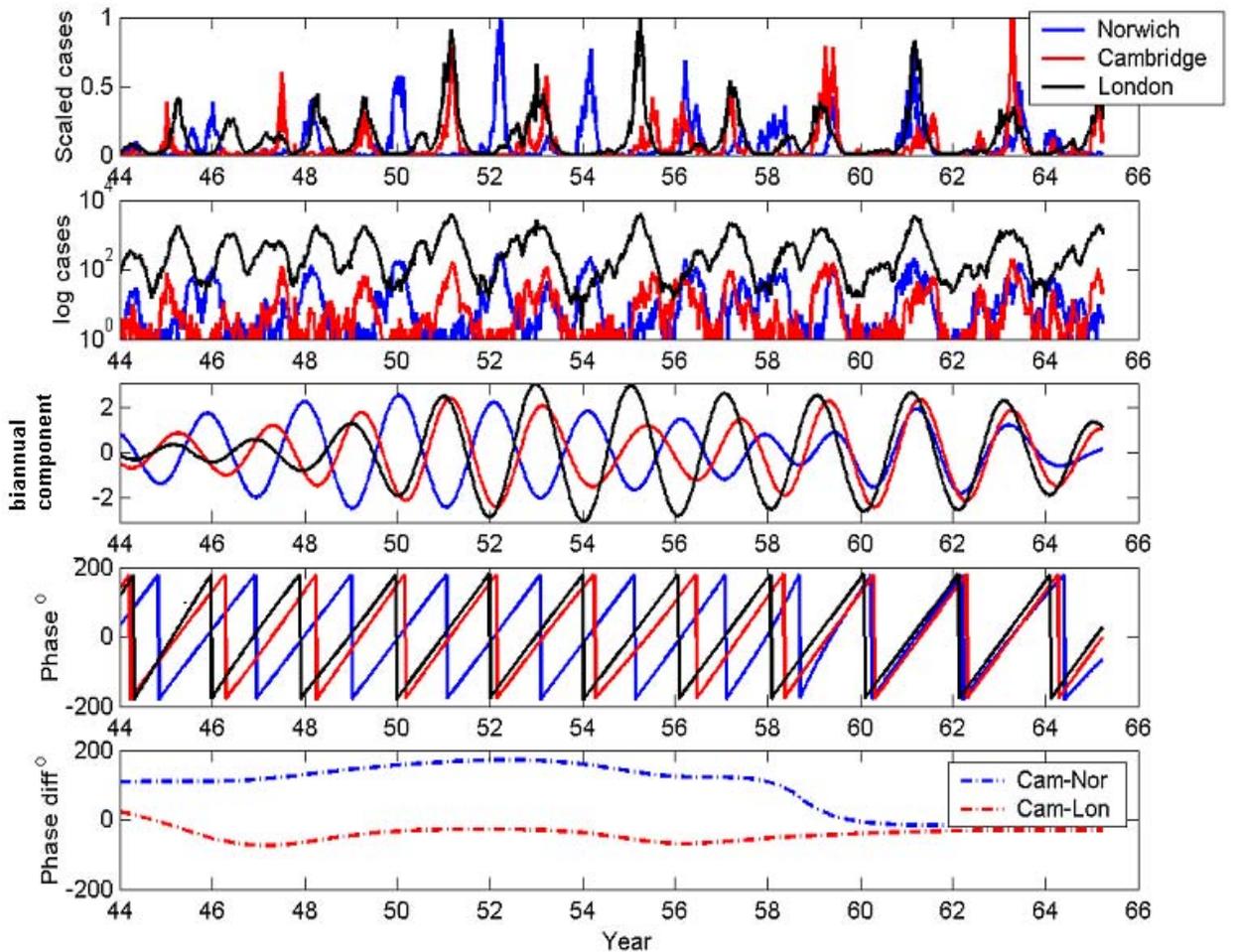


Figure 19: Comparing the phase of Cambridge, Norwich and London (see text for explanation), from: Grenfell et al. (2001). Copyright permission from Nature Publishing Group, <http://www.nature.com>

In Figure 19, we investigate the phase differences of the measles epidemics in Cambridge, Norwich, and London (for comparison). The top plot gives the number of cases, scaled to fit on the same axes, the second plot shows the same data but log transformed. The middle plot contains the reconstruction of the biennial component only, as described above. We see clearly that the phase difference observable in the first two plots is well captured in the biennial component. Whereas Cambridge and London are fairly well in phase for the whole time, Cambridge and Norwich are completely out of phase until the late 1950s. The same is apparent when we plot the phases for the epidemics in the three cities, see the fourth plot in Figure 19. Finally, the last plot shows the phase differences between Cambridge and London (red), which is close to zero and negative most of the time, and Cambridge and Norwich (blue), which is close to 180 degrees until the late 1950s when it decreases to small values around zero.

We note that the analysis of the biennial component of the wavelet spectrum gives a good description of the phase and phase difference of the epidemics in different cities. As we are interested in spatio-temporal patterns of measles, we ask two related questions: Is the phase difference between London, Cambridge and Norwich caused by the spatial distance between them, and more generally, can we observe spatial waves?

It turns out that the phase difference between London and Cambridge is due to the distance between them and we do see waves, but that distance alone does not explain the situation in Norwich. In fact, it seems that isolation or inaccessibility are the dominant factors in this case, not distance. Spatial waves can be observed, but the question of how to describe or model them is still only partially answered.

2.3 Phase differences from London

To look at the effect of distance from London on the phase of the epidemic, we apply the above analysis of the biennial wavelet component to all the cities in which measles were reported. In the streaming video (<http://www.pims.math.ca/>), the phases and phase differences are shown as a movie. One observes that the initial spark of an epidemic seems to come from the northwest, then trigger the epidemic outbreak in London, from where subsequently epidemic waves spread. In the movie of the phase *differences* to London, the situation is close to stationary. Therefore, we plot these phase differences, averaged over a period of 16 years (Figure 20).

We observe the following. In the plot for the urban areas (left), there is a little red triangle in the northwest, indicating that this city is ahead of London in terms of the epidemic. London, of course, is green since it is in phase with itself. Close to London, there are a few green and light blue triangles, indicating that the phase is slightly lagging behind, further away from London, the colors are predominantly darker blue, which indicates that there the epidemic clearly breaks out later. Cambridge (between London and Norwich) is lagging slightly behind London, as we already saw in Figure 19. Norwich and the whole area of East Anglia, however, is, on average, ahead of London. This is in stark contrast to the adjacent rural areas. A similar phenomenon on a smaller scale can be found at the coast of Wales.

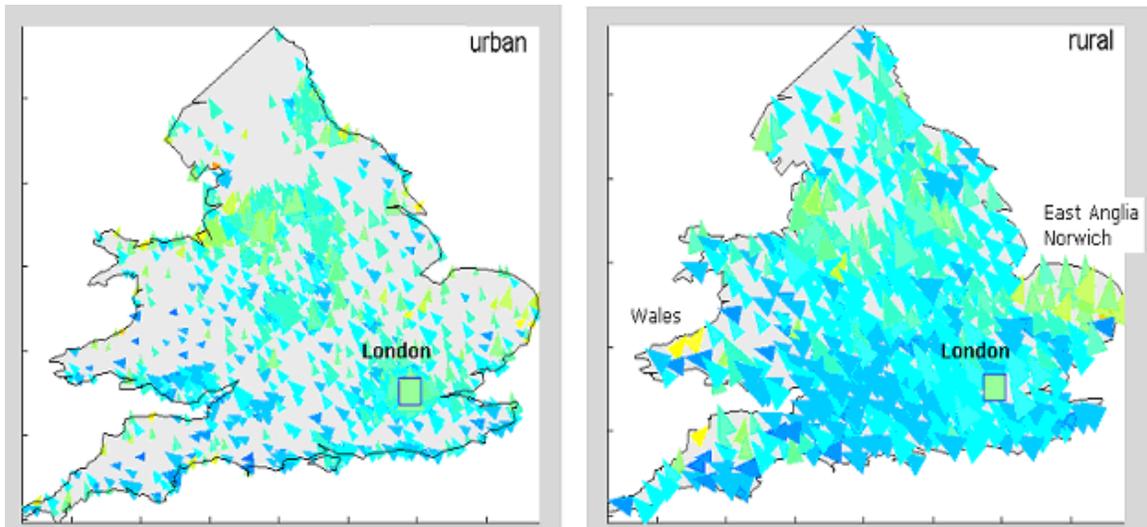


Figure 20: Phase differences with respect to London in the biennial component of the measles epidemic, averaged from 1950 to 1966. Green indicates in phase with London, blue is lagging behind, yellow and red is ahead (from: Grenfell et al., 2001). Copyright permission from Nature Publishing Group, <http://www.nature.com>

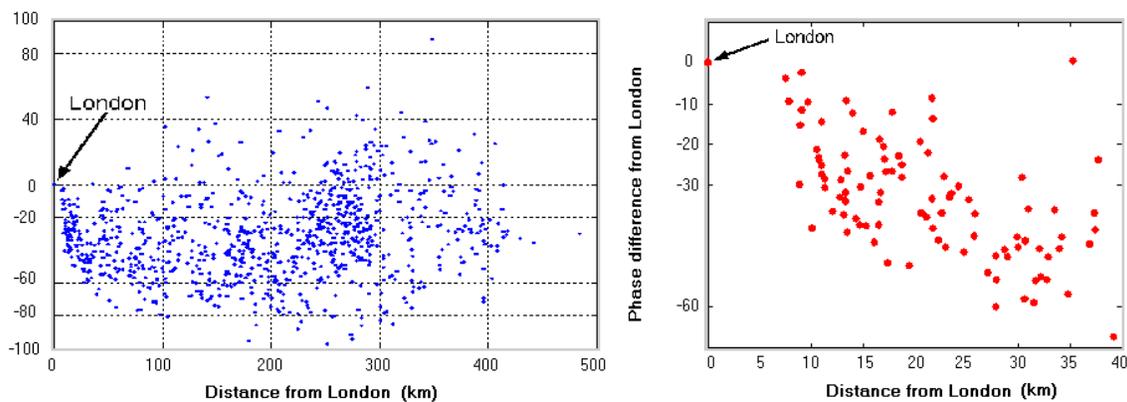


Figure 21: Phase difference with respect to London, plotted as a function of the distance to London. The plot on the right is zooming in on the cities that are less than 40km from London (compare to Figure 3c in Grenfell et al., 2001). Copyright permission from Nature Publishing Group, <http://www.nature.com>

As a first step to analyzing this situation more closely, the phase difference to London is plotted as a function of the distance to London in Figure 21. It becomes apparent from the plot on the left that, compared to London, the epidemic is lagging behind in most cities (negative phase difference) and only a few are ahead of London. By zooming in on the cities closest to London (the plot on the right), we see clear indication of a spatial wave of the epidemic, starting in London and moving away.

2.4 Spatial Correlograms

In the last section, we looked at the phase differences as a function of distance in order to detect spatial waves. A slightly different approach is to plot the correlation of the epidemic in different cities as a function of the distance between these cities. The resulting nonparametric spatial correlograms in Figure 22 below are work by Ottar Bjørnstad, see Grenfell et al. (2001).

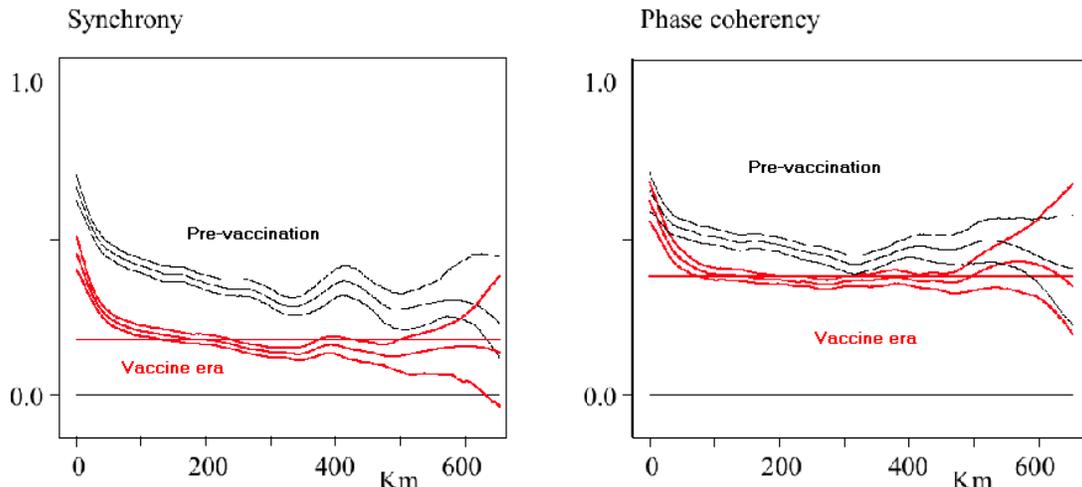


Figure 22: Spatial correlograms. The plot on the right contains information about the phase only, irrespective of the amplitude of the epidemics. The plot on the left incorporates both, phase and amplitude of the epidemics to compute correlations (from: Grenfell et al., 2001). Copyright permission from Nature Publishing Group, <http://www.nature.com>

The plot on the right compares the phase correlations only and ignores amplitudes of the epidemics. The phase coherency drops as the distance from London increases. However, this trend seems to stop at a distance of about 300km. This suggests a regional scale of influence of the large city of London. In the vaccine era, phases are much less correlated than in the pre-vaccination era. The plot on the left includes both, amplitude and phase. We see the same overall pattern that the correlation drops significantly with distance and that the correlation is lower in the vaccine era. Upon comparing the plot on the left with the plot on the right, we see that the overall correlation is lower, when amplitudes are included. This leads us to conjecture that the amplitude is more sensitive to stochasticity than the phase is. The difference between the black and red graphs in the plots also leads to suggest that the correlation in the amplitudes is more severely affected by vaccination than the phase is.

2.5 What causes spatial waves?

The work on this question is still in its infancy. It is known that deterministic spatial models for infectious diseases synchronize quickly (Lloyd and May, 1996). Hence, these models cannot account for the observed data. On the other hand, in a stochastic model with populations of equal size, there are no waves (Lloyd and May, 1996). Therefore, we imagine a simple system of a big and a small city between two outbreaks of an epidemic (see Figure 23).

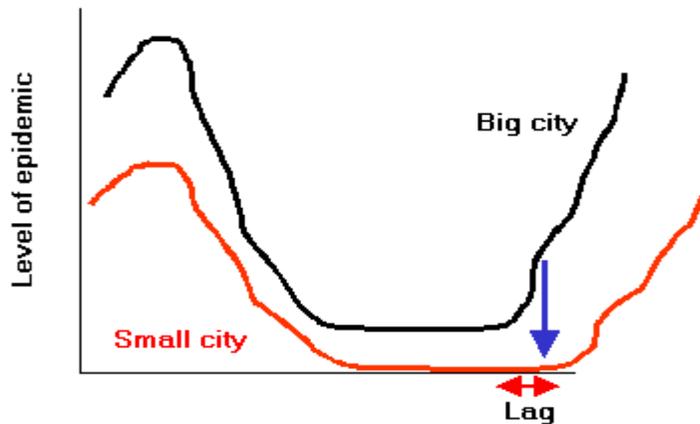


Figure 23: Hypothetic levels of epidemic for a big city (top curve, black) and a small city (bottom curve, red).

At the end of an epidemic outbreak, the disease vanishes from the small city due to stochastic fade out, but persists in the big city whose population is above the threshold. As soon as the critical number of susceptibles is reached again, the epidemic can take off in the big city. In the smaller city, a spark of the epidemic is needed from outside the small city in order to trigger the outbreak. This spark may come through individuals moving between the cities. Therefore, there will be a time lag between the next outbreak in the two cities. This resembles the dynamics of a core-satellite metapopulation.

To summarize, three ingredients are needed to produce the behavior observed in the recurrent measles epidemics in England and Wales.

1. The epidemic oscillator with a seasonally forced infection rate drives the course of a fast epidemic with long intervals between two successive outbreaks.
2. Stochasticity is needed to obtain a demographic threshold for extinction of the disease in small places.
3. A hierarchy in the population numbers is needed to obtain the spatial waves observed in the data and described by the thought experiment above.

3. Dynamics and evolution of pathogens

In the first two sections, we studied epidemic diseases from the viewpoint of the population that is affected rather than from the point of view of the pathogen that infects the individual. Whereas modeling the whole population yields some important insight in the dynamics of these diseases and gives management advice for immunization strategies, it captures only one part of the picture. Some important questions can only be tackled when the virus itself, its strain structure, mutations, and the selective forces acting upon it are the focus of our model. The novel aspect here is to link the evolutionary processes on the parasite level to the ecological processes at the host population level. The combined study of these effects has been termed *phylodynamics* (Grenfell et al., 2004).

In this last section, we consider two exemplary studies that link the evolution of the pathogen with the dynamics of the disease on the population level. In the case of measles discussed above, the strain structure and possible mutations seem mostly irrelevant because after the disease, complete immunity for all strains of measles is obtained. In flu-like diseases, however, we see evolutionary processes in action as viruses mutate into new strains for which previous infections yield only partial cross-immunity. Another example is the emergence of drug resistant strains of pathogens that are becoming a major concern in hospitals and large animal aggregations.

3.1 Selection and partial cross-immunity

Influenza is a small RNA virus of 8 genes that is highly variable. At any point in time, the strain structure of the existing viruses can be recorded, so that over long time periods changes in the strain structure can be detected. A plot of existing strains as a function of time is called a *phylogenetic tree*. The question now is, how mutation on the level of the individual virus turns into the observed phylogenetic tree. To answer that question, the epidemic dynamics have to be taken into account, most notably the duration of the infection and the conferred partial cross-immunity to other strains. Some models have been proposed for this but they usually dealt with only a few different strains. Since cross-immunity depends on previously contracted strains, the history of infection is taken into account in these models, and therefore the number of parameters increases exponentially with the number of strains. We present a simplified generic model here in which the number of parameters grows only linearly with the number of strains, and which, therefore, is tractable even when 100 and more strains are included. For details and further background, see Gog and Grenfell (2002).

We apply the same compartmental modeling idea on the host population level as above for measles, except that this time susceptible and infectious refers to a particular strain, indexed by j . Hence, if there was no interaction between strains, the model would simply consist of many uncoupled equations of the form

$$\begin{aligned}\frac{d}{dt}I_j &= \beta_j S_j I_j - \nu_j I_j - \mu I_j, \\ \frac{d}{dt}S_j &= \mu - \beta_j S_j I_j - \mu S_j,\end{aligned}$$

where μ is the population birth and death rate (the population is held at a constant level, which is scaled to unity), and β_j and ν_j are the strength of infection and the recovery rate for strain j . In this model, a single host individual appears as many individuals, namely as one for every strain to which it is susceptible or by which it is infected.

There are several ways to think about and consequently model partial cross-immunity. The approach followed here is that of *reduced transmission* that assumes that a host can still be infected but is less likely to be infectious, i.e., cannot transmit that strain to which it had acquired immunity. The other modeling assumption is that of *polarized immunity*, which means that *partial* cross-immunity renders *some* of the hosts completely immune. This polarized immunity is introduced into the model through a probability σ_{jk} that describes the immunization effect of an individual susceptible to strain j , becoming infected by strain k . Hence, individuals are removed from S_j by becoming infected with strain k , but they do not appear in I_j , i.e., they cannot transmit strain j . The equation for S_j then becomes

$$\frac{d}{dt} S_j = \mu - \sum_k \beta_k S_j \sigma_{jk} I_k - \mu S_j.$$

We choose the simplest possible strain space, namely a linear arrangement of strains, and leave the discussion of possible alternatives to the literature. Cross-immunity is assumed to be of the symmetric form

$$\sigma_{jk} = \exp\left(-\frac{(j-k)^2}{d}\right),$$

where d is the typical length scale involved in cross-immunity. Mutation is modeled as a simple random walk in strain space, modeled by discrete diffusion with parameter m , so that the equation for infectives becomes

$$\frac{d}{dt} I_j = \beta_j S_j I_j - \nu_j I_j - \mu I_j + m(I_{j+1} - 2I_j + I_{j-1}).$$

A schematic of the linear strain space with mutation and cross-immunity superimposed is given in Figure 24 below.

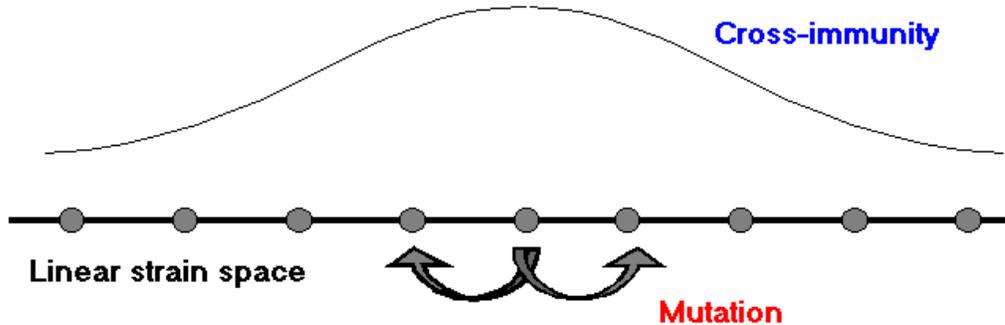


Figure 24: Linear strain space, each point corresponds to a strain. Schematically, mutation is modeled by a random walk to an adjacent site, and cross-immunity is given by a Gaussian curve (adapted from Gog and Grenfell, 2002).

We focus on the question of how the duration of the infection influences the pattern of possibly coexisting strains. To achieve this goal, we assume that all strains have equal parameters (β , ν , m) and plot the simulation results for two different infectious periods in Figure 25 below. In the case of a long infectious period, we see that a first cluster of strains emerges. Even though mutations to nearby sites occur, these strains cannot multiply because their hosts are depleted due to cross-immunity.

Once a mutation occurs far enough away from the initial cluster, this new strain finds a large reservoir of hosts and it can subsequently multiply to form a second cluster. The process continues through strain space. Each cluster of strains remains prevalent in the population since the infection is long with respect to the lifetime of the host, whereby new susceptibles are supplied before the infection with that strain cluster has died out.

The picture for short infections is different in that the emerging clusters do not persist in time. This is because for each strain cluster it takes too long to build up the level of susceptibles to the threshold above which the disease could take off a second time (compare the measles dynamics above). By the time enough individuals are susceptible to a certain strain again, this strain is not present in the population any more. Cross-immunity and only localized mutations ensure that the strain cannot invade any more.

Since the level of strains between clusters is very low, the second scenario would be perceived as a series of jumps in the strain structure rather than a continuous process. Therefore, the purely deterministic model presented above, based on overly simplifying assumptions, was able to generate clumps in strain space. Referring back to the originally considered Influenza infection, we suggest that the observed drift in Influenza strains occurs from a long lived immune memory in the host, i.e., the human immune system.

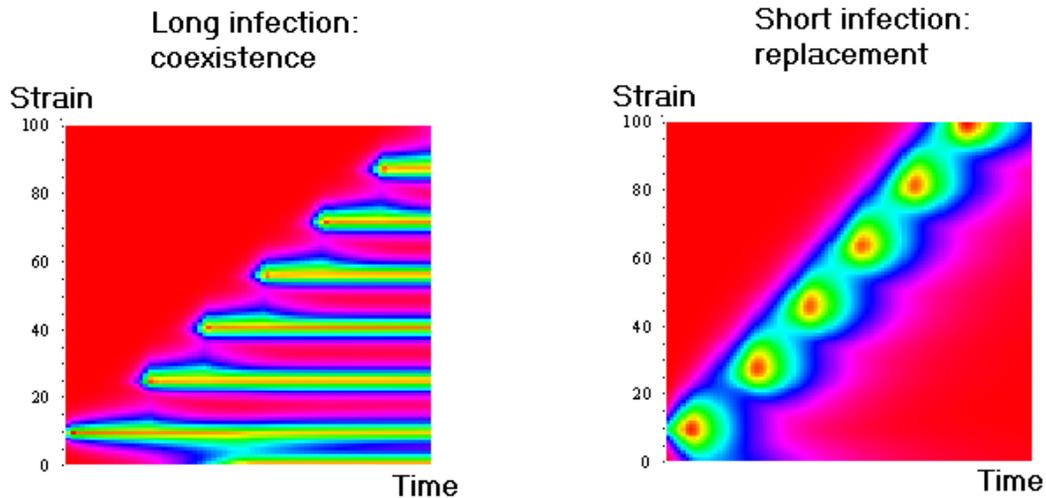


Figure 25: Simulation results of the strain structure in a model with 100 strains. Strain prevalence increases from purple via blue and green to yellow and red. The big red areas above the diagonal (and in the lower right corner in the second plot) indicate absence of strains (compare Figure 2 in Gog and Grenfell, 2002).

Recent work also underlines the importance of stochastic persistence in the invasion of human influenza A escape mutants against cellular immunity (Gog et al. 2003).

3.2 Spread of rare recessives

In this last section, we study macroparasite (mainly helminth worms) dynamics and the invasion of drug resistant escape mutants. In this case, the parasite is chosen as the unit of study because the actual burden of parasites per host individual determines the level of the disease as well as the build-up of immunity. Hence, we leave the framework of SIR models and turn to metapopulation models for the parasites. Since adult parasites don't generally multiply directly in hosts, we need to model the transmission between hosts more carefully and take larval stages of the parasite into account, see Figure 26. During a season, parasitism will increase and build up in the host. If the host is not killed by the parasite load then the burden will decrease eventually. This decrease is partly due to resistance of the host and partly to drugs. The following observation about parasite levels in actual sheep population stands out. Worms are distributed in patches so that a minority of hosts carries the majority of worms. This patchiness is quite important for the evolution of the worm. The following is based on Cornell et al. (2003)

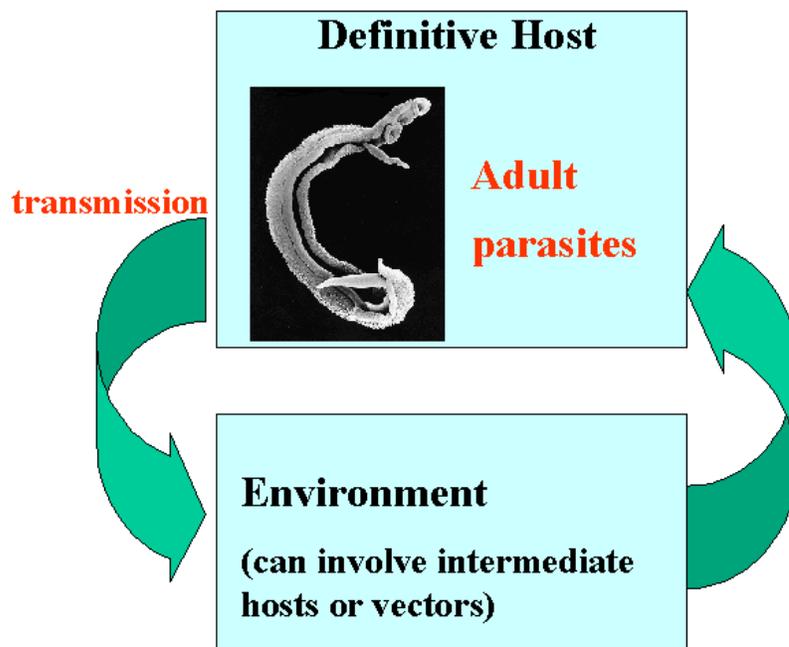


Figure 26: Schematic of the transmission cycle for the parasite worm between adult stage in the host and larval stage in the environment.

As indicated in the section heading, we want to study why and how rare recessive traits can spread in such a worm population. The background of this question is that drug resistance of worms seems to be spreading fast in sheep populations in New Zealand. In some cases, this resistance a single-locus recessive trait. According to the law of Hardy-

Weinberg, such a trait should be spreading in a population with rate proportional to q^2 , provided that its initial proportion in the population is q . Given that drug resistance initially develops in a small number of worms, spreading of resistance in the Hardy-Weinberg scenario is highly unlikely. However, the Hardy-Weinberg law holds only for well-mixed populations, and we have already seen that parasite load is very patchy. This patchy distribution between the hosts leads to spatial clumping of larvae in the environment (see Figure 27), which in turn leads to “clumped infection” (Figure 28). Larvae from one host aggregate and are taken up by the next host, which promotes inbreeding in the worms. The simple metapopulation model described below shows that the genetic variability together with the inbreeding predicts a spread rate of the rare recessive gene that is much higher than the well-mixed Hardy-Weinberg prediction.

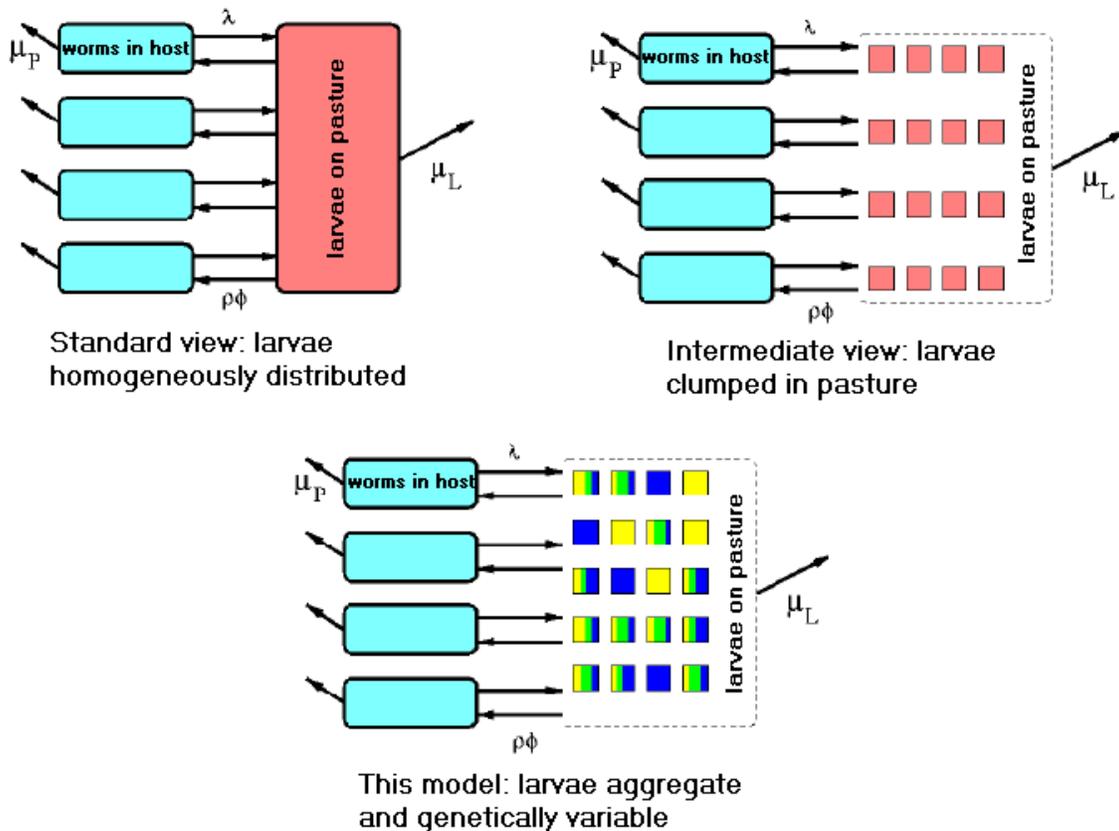


Figure 27: Standard views assume that larvae are homogeneously distributed over a pasture. This model assumes spatial clumping and genetic.

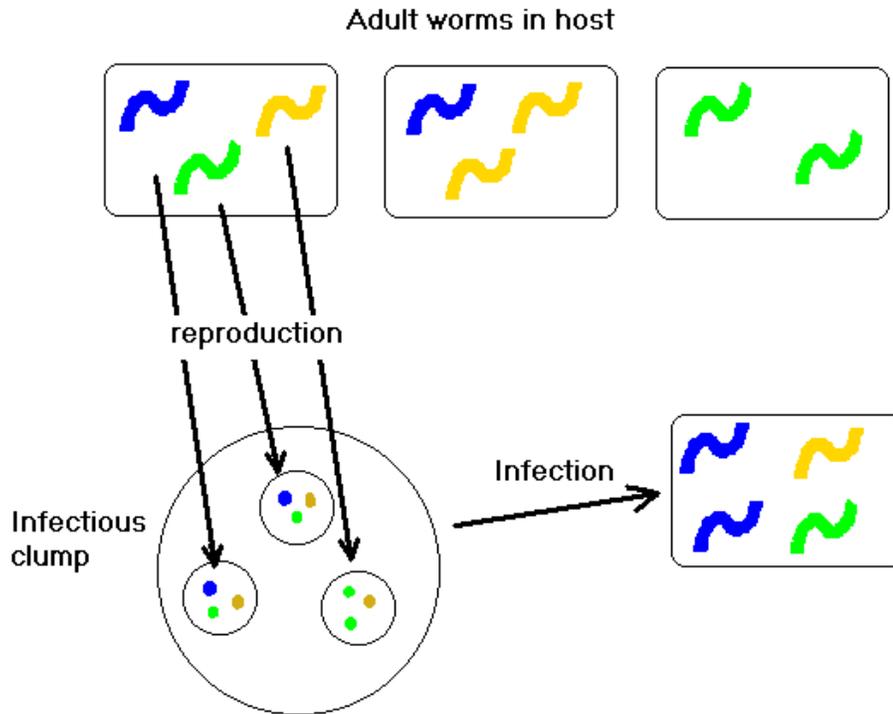


Figure 28: Clumping of larvae promotes inbreeding (adapted from Cornell et al., 2003).

The first metapopulation model to explore this effect of clumping is a continuous time Markov process. Its state variable is $W_{ij}(t)$, the number of parasites of genotype j in host i . In this case, i ranges from 1 to N , and j is either 1, 2, or 3, corresponding to genotypes ss , sS , and SS . The rare allele is denoted by s , the wild type by S . Infections occur with rate λ/N , i.e., λ is the rate at which a host encounters a clump of larvae. For the infection, the worms in the original host multiply (if there are at least two worms present), and the number of offspring of each genotype is computed according to the assumption of well-mixing within the host. Then this clump is taken up by the new host.

In Figure 29, we plot the results from this simulation model in a doubly logarithmic plot. The prevalence of the rare homozygote “ ss ” is plotted against the initial frequency of the resistant gene. If the system was well mixed, the Hady-Weinberg prediction would give that the prevalence is the square of the initial frequency, which leads to a line of slope 2 in the plot. The model with clumping, however, produces a line of slope 1 for all time. We only plot the distributions for four different instances in time. At $t=0$, the distribution naturally begins at the Hardy-Weinberg equilibrium, but then it increases drastically until it reaches a maximum from which it settles to an asymptote at $t=6$. The difference in the model compared to Hardy-Weinberg is particularly strong for small initial frequencies. With the much higher prevalence, the rare recessive gene is much more likely to spread due to clumping and inbreeding than in a well-mixed population.

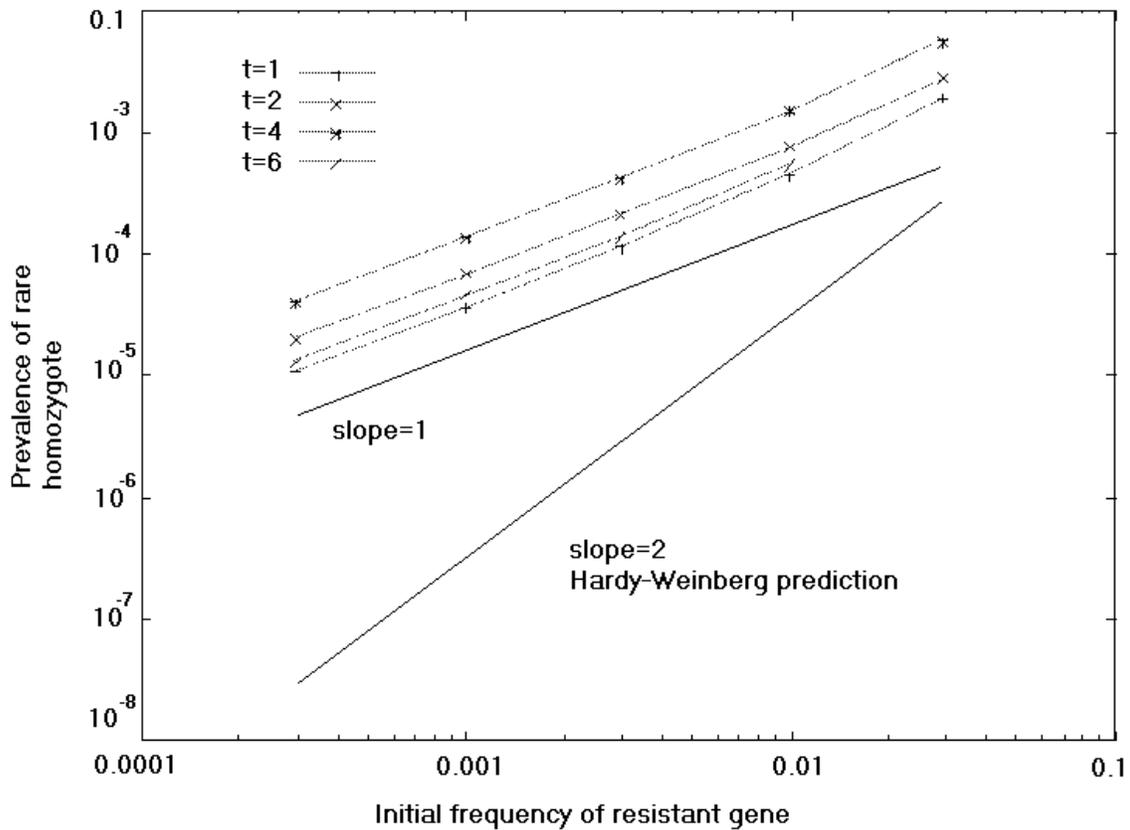


Figure 29: The prevalence of the rare homozygote in the simple metapopulation model (from: Cornell et al, 2003). For explanation, see text above.

A more detailed model was made (Cornell et al, 2003) that incorporates the number of male and female worms, keeps track of the number of different clumps on the pasture, and models the dynamics over several seasons. We will not describe the details here, but simply show the schematics in Figure 30 and the results in Figure 31. For two typical runs, the worm load per sheep is plotted. In one case, the resistant strain takes over and the worm population persists, in the other, the worm population dies out. Finally, we show how the invasion probability depends on the initial allele frequency for in the case of randomness and clumping. The shape of the graph is similar to the one in Figure 29, hence, the prediction from the simple model is confirmed by the more complex model.

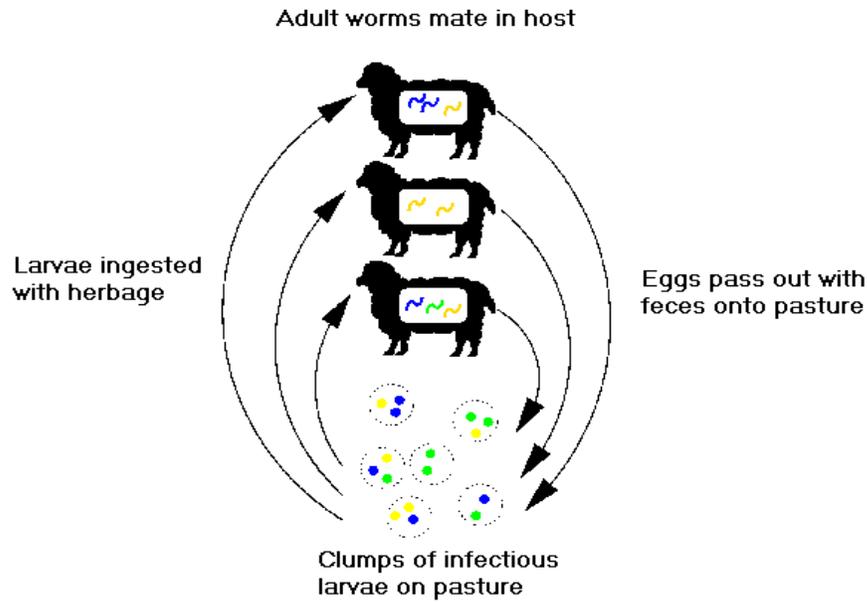


Figure 30: Schematic depiction of the processes modeled in the complex model, from: Cornell et al. (2003).

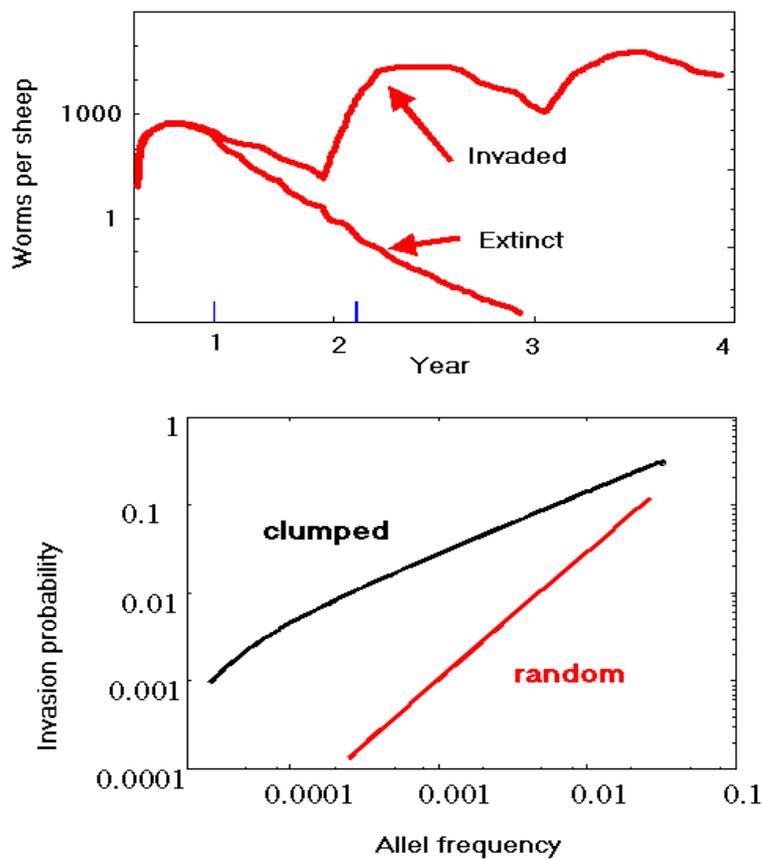


Figure 31: Outcome of the complex worm model. See also Figure 3 B,C in Cornell et al. (2003)

References

For extensive references, please see the [website](#) of the Grenfell lab.

General Reference

Anderson, R.M., and May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press

Grenfell, B.T., and Dobson, A.P. (1995) *Ecology of Infectious Diseases in Natural Populations*. Cambridge University Press

Childhood disease dynamics

Earn, D.J.D., Rohani, P., Bolker, B.M. and Grenfell, B.T. (2000) A Simple Model for Complex Dynamical Transitions in Epidemics. *Science* **287** 667-670

Hethcote, H. (2002) New Vaccination Strategies for Pertussis. In: Castillo-Chavez, C. (ed.) et al., *Mathematical approaches for emerging and reemerging infectious diseases: An introduction*. IMA Vol. Math. Appl. **125**, 97-118

Grenfell, B.T., Bjørnstad, O.N., and Kappey J. (2001) Travelling Waves and Spatial Hierarchies in Measles Epidemics. *Nature* **414** 6865 716-723

Keeling, M.J., Rohani, P. and Grenfell, B.T. (2001) Seasonally-forced Disease Dynamics Explored as Switching Between Attractors. *Physica D* **148** 317-335

Lloyd, A.L., and May, R.M. (1996) Spatial Heterogeneity in Epidemic Models. *J. theor. Biol.* **179** 1-11

Rohani, P., Earn, D.J.D. and Grenfell, B.T. (1999) Opposite Patterns of Synchrony in Sympatric Disease Metapopulations. *Science* **286** 968-971

Rohani, P., Earn, D.J.D. and Grenfell, B.T. (2000) The Impact of Immunisation on Pertussis Transmission in England and Wales. *The Lancet* **355** 285-286

Rohani, P., Keeling, M.J. and Grenfell, B.T. (2002) The Interplay between Determinism and Stochasticity in Childhood Diseases. *Am. Nat.* **159** 469-481

Schaffer, W.M. and Kot, M. (1985) Nearly One Dimensional Dynamics in an Epidemic. *J. theor. Biol.* **112** 403-427

Schenzle, D. An Age-Structured Model of Pre- and Post-Vaccination Measles Transmission. *IMA J.Math. Appl. Med. Biol.* **1** 169-191

Torrence, C. and Compo, G.P.A. (1998) A Practical Guide to Wavelet Analysis. *Bull. Am. Meteorol. Soc.* **79**, 61-78

TSIR Models

Bjørnstad, O.N., Finkenstädt, B., and Grenfell, B.T. (2002) Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates using a Time Series SIR Model *Ecological Monographs* **72** 169-184

Finkenstädt, B. and Grenfell, B. (2000) Time Series Modelling of Childhood Diseases: a Dynamical Systems Approach. *J. Roy. Stat. Soc. C* **49** 187-205

Grenfell, B.T., Bjørnstad, O.N., and Finkenstädt, B. (2002) Dynamics of Measles Epidemics: Scaling Noise, Determinism and Predictability with the TSIR Model. *Ecological Monographs* **72** 185-202

Foot and mouth disease

Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J., and Grenfell, B.T. (2001) Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. *Science* **294** 813-817

Keeling, M. J., Woolhouse, M. E. J., May, R. M., Davies, G., and Grenfell B. T. (2003) Modelling Vaccination Strategies against Foot-and-Mouth Disease. *Nature* **421** 136-142

Evolution of pathogens

Boni, M.F., Gog, J.R., Andreasen, V., and Christiansen, F.B. (2004) Influenza Drift and Epidemic Size: The Race between Generating and Escaping Immunity. *Theor. Pop. Biol.* **65** 179-191

Cornell, S.J., Isham, V.S., Smith, G., and Grenfell, B.T. (2003). Spatial Parasite Transmission, Drug Resistance, and the Spread of Rare Genes. *Proc. Nat. Acad. Sci., USA* **100** 7401-7405

Gog, J. R., Rimmelzwaan, G.F., Osterhouse, D.M.E., and Grenfell, B.T. (2003). Population Dynamics of Rapid Fixation in Cytotoxic T Lymphocyte Escape Mutants of influenza A. *Proc. Nat. Acad. Sci., USA* **100** 11143-11147

Gog, J. R., and Grenfell, B.T. (2002). Dynamics and Selection of Many-Strain Pathogens. *Proc. Nat. Acad. Sci., USA* **99** 17209-17214

Gog, J. R., and Swinton, J. (2002). A Status-Based Approach to Multiple Strain Dynamics. *J. Math. Biol.* **44** 169-184

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L.N., Daly, J., Mumford, J.A., and Holmes, E.C. (2004) Unifying the Ecological and Evolutionary Dynamics of Pathogens. *Science* **303** 327-332