Seasonality and dynamics of whooping cough

by

Hanh T.H. Nguyen

(Under the direction of Pejman Rohani)

Abstract

Whooping cough (pertussis) dynamics provide an interesting disease ecology case study. Unlike other childhood diseases, the observed patterns of pertussis dynamics are found very diverse and are not easily captured by simple deterministic models. This has led to the current understanding that the disease dynamics can only be explained by adding stochasticity into the models. In this work, we demonstrate that an appropriate deterministic model can explain pertussis dynamics. The consequences of using the model in making public health decisions are also discussed.

Whooping cough dynamics also exhibits strong seasonality, which is thought to result from variation in contact rates. In this research, seasonal change in disease incidence and the timing of outbreaks are also analyzed using case report data from several UK cities. We show that birth rates in prior years have a positive correlation with current outbreak sizes and a negative correlation with current outbreak peak time. A modeling approach is also used to understand and explain the patterns found in the data.

INDEX WORDS: Whooping Cough, Modeling, Seasonality, Gamma Distribution

Seasonality and dynamics of whooping cough

by

Hanh T.H. Nguyen

B.S., Hanoi National University, Vietnam, 2003

A Thesis Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Master of Science

Athens, Georgia

2007

Seasonality and dynamics of whooping cough

by

Hanh T.H. Nguyen

Approved:

Major Professor: Pejman Rohani

Committee: Sonia Altizer

Daniel Promislow

Electronic Version Approved:

Maureen Grasso

Dean of the Graduate School

The University of Georgia

August 2007

Acknowledgments

I would like to thank Pej Rohani, my advisor, for his numerous wonderful advices, constant support and for being patient and understanding. My thanks also go to Sonia Altizer and Daniel Promislow for their great advices on doing the project and doing research in general.

I am very thankful to Helen Wearing, Marc Choisy and Dan Vasco for their comments and encouragement.

I would also like to thank John Pickering for his support during my first year in the US and in the graduate school.

My parents deserve special thanks for their love and support. And I also wish to thank my older sister, chi Binh, my two younger brothers, Duc and and Nghia, for being the most wonderful supporters and inspirators ever. Finally, I am thankful to my husband, Ha, for whatever's left and because he would get disappointed if his name did not appear here.

Table of Contents

List of Figures

Chapter 1

Introduction

Population ecology aims to understand the mechanisms underlying the temporal and spatial distribution of species. Disease ecology, another subfield in Ecology applies ecological principles to problems in epidemiology. In population disease ecology, the patterns of most interest are the fluctuations of disease incidence in time and space. Thus, many studies in this subject can be found focusing on understanding the dynamics of disease incidences.

In epidemiology and disease ecology, the management of infectious diseases is among the most important issues. There are not only concerns about newly emerging infectious diseases, such as SARS (Lipsitch *et al.* 2003; Riley *et al.* 2003) and the H5N1 strain of avian influenza (Obenauer *et al.* 2006; Olsen 2006), but also "re-emergent" diseases such as whooping cough (Crowcroft & Pebody 2006). Whooping cough, an infectious disease caused by the bacterium *Bordetella pertussis* causes an estimated 20-40 million cases and 200,000-400,000 deaths every year (WHO 1999), mostly in developing countries. In some highly-vaccinated countries like Australia, Canada and the US, there are also reports that whooping cough incidence has been increasing in recent years (WHO 1999). These facts highlight that we are still far from eradicating the disease.

To help predict outbreaks, plan successful vaccination programs and eventually eradicate infectious childhood diseases, one must understand the nature of disease dynamics. This involves answering several related questions about disease dynamics in time and space. One of the main issues is explain temporal patterns of infectious disease dynamics in historical data.

For example, what are the mechanisms responsible for the periodicity of epidemic outbreaks? Other important questions involves understanding why spatial patterns of epidemics are more synchronous and regular in the pre-vaccination era than in the vaccine era for some childhood diseases and vice versa for some others (Rohani *et al.* 1999). In addition, because vaccination changes the patterns of epidemic dynamics, it is also essential to understand the dynamical effects of different vaccination strategies. As direct transmission of infection depends on the aggregation of individuals, which in reality is not random, there is also a question about the role of social networks in temporal and spatial patterns of epidemic dynamics. These questions are linked to one another so answering one question will give insights into the other.

Mathematical modeling is a powerful tool to help answer those questions. Modeling has proven to be helpful in conceptualizing and quantifying our ideas about the behavior of a particular system (Keeling 2005). In epidemiology, the use of models has a long history initiated by Bernoulli's work on smallpox in 1760 (Bernoulli 1760) and followed by other important studies (Hamer 1906; Kermarck & McKendrick 1927; Soper 1929). The development of computers has allowed models to be more helpful regarding their ability to quantitatively solve various problems as well as give detailed predictions (Levin *et al.* 1997). Although there are limits to the capacity and precision of models used, models can still provide insights into questions aimed at understanding and predicting disease outbreaks or testing different control strategies (Keeling 2005).

In this thesis, I aim to understand the dynamics and seasonality of whooping cough using a modeling approach. Whooping cough is a good case study because of its different patterns of fluctuating incidence observed at different times and in different regions (countries and even cities). In many UK cities, the temporal dynamics in the pre-vaccination era were irregular with annual outbreaks punctuated by periods of 2, 2.5 and 3-years. After vaccination was introduced in 1957, not only did the number of infectious cases change but so did the dynam-

ical patterns. Longer multi-year cycles have been observed including dynamics with periods of 3, 3.5 and 4-years (Hethcote 1998; Gomes *et al.* 1999; Rohani *et al.* 2000; Broutin *et al.* 2004). It is then natural to ask what are the underlying mechanisms of these behaviors? Can the same mechanism give rise to different patterns? Or are different mechanisms involved? The use of mathematical models has proved invaluable in giving insights into such epidemic fluctuations. The two main goals of my work presented here include capturing whooping cough dynamics as observed in real data (focusing on periodicity) and explaining the seasonality of whooping cough dynamics (focusing on seasonality in outbreak size and phase). These two problems are presented as Chapter two and three of the thesis.

In Chapter two, I focus on explaining and capturing whooping cough dynamics as seen in historical data. In general, the current understanding is that the main factors determining patterns of disease dynamics are inherent properties (transmissibility, duration of exposed/infectious period, etc), and external forces: amplitude of seasonality (changes in contact rate during school term and vacation) and rate of recruitment (birth and vaccination) (Anderson and May 1991; Earn *et al.* 2000; Lloyd 2001a, 2001b; Rohani *et al.* 1999; Rohani *et al.* 2002). Nevertheless, there is not always a general answer for different diseases. The success of deterministic models in capturing dynamics of several infectious diseases including measles is obvious (Schenzle 1984; Anderson and May 1991; Bolker and Grenfell 1993; Earn *et al.* 2000). For pertussis, the story is different. Simple deterministic models always predict annual cycles of the disease (Rohani *et al.* 1999, 2002), whereas observed whooping cough dynamics are very diverse at different times and in different regions. For example, a mixture of annual cycles and cycles of 2, 2.5 and 3-years are observed in England and Wales; and periods of 3, 3.5 and 4-years in some other countries (Hethcote 1998; Gomes *et al.* 1999; Rohani *et al.* 2000; Rohani *et al.* 2002; Broutin *et al.* 2005a). Interestingly, in the vaccine era, whooping cough exhibits more regular and spatially synchronized patterns with periods of 3.5-4 years (Rohani *et al.* 1999, 2000; Broutin *et al.* 2005b). The failure of deterministic models in capturing whooping cough dynamics is believed to be due to the

fragility of whooping cough attractors. As whooping cough dynamics is too sensitive to be understood by a purely deterministic model, adding the interplay between stochasticity and inherent non-linearity is argued to be the reason for the irregular dynamics of whooping cough.

The argument about the role of stochasticity in whooping cough dynamics, in fact, is just a part of a larger debate that concerning the relative importance of deterministic and stochastic factors in shaping ecological dynamics. While the deterministic school (Nicholson, 1957) claimed that ecological dynamics are determined by deterministic density-dependent factors, Adrewathar and Birch (1957) argued that we need to rely on stochastic forces to understand these dynamical patterns. Although the emphasis in the last century was mostly on the deterministic approach, there is a growing trend toward uncovering the importance of stochasticity and the interaction between noise and deterministic nonlinearities in many ecological systems. These systems include laboratory flour beetle populations (Cushing *et al.* 1998), sheep populations (Coulson *et al.* 2001; Grenfell *et al.* 2001) and infectious diseases (Schwartz and Smith 1983, Rand and Wilson 1991, Bjørnstad and Grenfell 2001; Keeling *et al.* 2001; Rohani *et al.* 2002; Coulson *et al.* 2004). As for the case of whooping cough, Rohani *et al.* (1999, 2000, 2002) proposed that because whooping cough has a longer infectious period in comparison with measles it is more prone to the effect of noise, making it impossible to understand the dynamics by fully deterministic models. This finding has greatly supported the role of stochasticity in regulating ecological dynamics.

However, the argument about the role of noise in shaping whooping cough can be challenged. Much of the argument about the role of stochasticity in shaping infectious disease dynamics, like those of measles and whooping cough, is based on several assumptions. One of them is that the infection rate and recovery rate are constant which gives rise to the use of exponential distributions to describe the time for which individuals remain exposed or infectious. This is a mathematically convenient but biologically unrealistic assumption (Sartwell 1950; Simpson

1952; Bailey 1954, 1975; Gough 1977; Lloyd 2001a; Lloyd 2001b; Wearing *et al.* 2005). In fact, the probability of staying in a class is more likely to depend on the time already spent in that class. Thus, the distribution of the latent or infectious period is less dispersed around the mean than expected from the exponential (Lloyd 2001a, 2001b; Wearing *et al.* 2005), suggesting that some other distributions should be used to describe exposed and infectious periods. Will the addition of a more realistic distribution affect the current understanding of whooping cough dynamics? Chapter 2 will discuss the use of an alternative and more appropriate model in explaining whooping cough dynamics.

The next chapter (Chapter 3) centers on understanding seasonality of whooping cough. I chose to study seasonal whooping cough dynamics firstly because seasonality is a very common phenomenon in epidemiology. Seasonal infections have been observed in a wide range of diseases not only childhood diseases (measles, chickenpox) but also fecal-oral infections (cholera, rotavirus) and vector-borne diseases (malaria) (Grassly & Fraser 2006). The second reason is that understanding seasonal disease dynamics is a critical step toward understanding and forecasting how long-term climate changes affect human health and other species in nature (Altizer *et al.* 2006; Pascual & Dobson 2005). Lastly, as Chapter 2 focuses on whooping cough dynamics at large scale (tens of years), Chapter 3 can be considered as a way to explore the dynamics at smaller scale (yearly).

My goal in exploring the seasonality of whooping cough is to explain the differences in outbreak sizes (the peak number of infected cases) from one outbreak to another of the same population using UK data. In addition, I also attempt to understand the variation in outbreak phases (the month in which outbreaks reach their peak) among outbreaks of the same population and among different populations. Results from data analysis and modeling of whooping cough seasonality in several UK cities are also compared. The results show that for whooping cough, there are correlations between birth rates in prior years and current

outbreak sizes and phases. Although these are just initial results, they can give some insights into predicting future outbreaks.

Chapter 2

The dynamics of whooping cough

2.1 Introduction

In recent years, there has been an increased awareness of the threats posed by newly emerging and high profile infectious diseases, such as SARS (Lipsitch *et al.* 2003; Riley *et al.* 2003), the H5N1 strain of avian influenza (Obenauer *et al.* 2006; Olsen 2006), HIV (Walker *et al.* 2003; Quinn & Overbaugh 2005), and Ebola (Frankish, 2003). In addition to novel pathogens, however, public health practitioners are concerned about a number of well-established infectious diseases that are re-emerging, defined as pathogens that have been around for a long time but exhibit increasing incidence and geographic range. These include whooping cough, multi-drug-resistant TB, dengue fever, West Nile virus and cholera (Daszak *et al.* 2000; Morens *et al.* 2004; Crowcroft & Pebody 2006). Understanding the mechanisms underlying disease transmission and spread is, therefore, clearly important from a public health perspective. Additionally, it has been argued by some that the rare combination of reasonably well understood natural history, a suite of appropriate mathematical models and abundant data make infectious diseases an important testbed for ecological theory (Anderson & May 1979; Earn *et al.* 1998; Keeling & Rohani 2007).

A particularly successful avenue for the study of population ecological questions has been via the examination of case notifications for the great micro-parasitic infections of childhood,

[0]This chapter is under review for J. R. Soc. Interface, titled "Noise, Nonlinearity and Seasonality: the Epidemics of Whooping Cough Revisited" by Hanh TH Nguyen and Pejman Rohani.

including measles, whooping cough, rubella and chickenpox. This has led to a substantial and important body of work addressing the role of nonlinearity and chaos (Schwartz & Smith 1983; Sugihara & May 1990; Grenfell 1992; Rand & Wilson 1993), stochastic extinction dynamics (Bartlett 1957; Anderson & May 1982; Keeling & Grenfell 1997; Nåsell 2005), and the consequences of various sources of heterogeneity, be they temporal (Soper 1929; London & York 1973; Fine & Clarkson 1982; Edmunds *et al.* 2000; Keeling *et al.* 2001), spatial (May & Anderson 1984; Rohani *et al.* 1999; Grenfell *et al.* 2001; Bjørnstad *et al.* 2002; Xia *et al.* 2004; Broutin *et al.* 2005a) or pertaining to the pattern of contacts (Schenzle 1984; Bolker & Grenfell 1993; Ferguson *et al.* 1996; Lloyd-Smith 2005; de Gama & Nunes 2006).

The study of childhood disease dynamics has also contributed to the perennial debate concerning the relative importance of deterministic versus stochastic forces in shaping observed patterns (Bjørnstad & Grenfell 2001; Coulson *et al.* 2004). For example, using a simple deterministic model with seasonality in the transmission rate –to mimic the aggregation of children in schools– Earn *et al.* (2000) successfully captured the dynamical transitions in measles epidemics in different cities and eras. They demonstrated that the observed dramatic shifts in dynamics are driven by changes in the recruitment rate of susceptibles, as determined by demographic trends and widespread paediatric vaccination programs. However, the application of this general approach to explaining the epidemics of whooping cough has been spectacularly unsuccessful. Classical deterministic models, in realistic regions of parameter space, always predict annual epidemics (Hethcote 1998; Rohani *et al.* 1998, 1999, 2002; Bauch & Earn 2003; Greenman *et al.* 2005), in direct contrast to the variable inter-epidemic periods documented in case notification data (Fine & Clarkson 1982). Prior to the onset of mass vaccination campaigns in England and Wales, for example, pertussis dynamics in different cities contained a significant multi-year signature, which gave way to regular 3.5-4 year epidemics in the vaccine era (Rohani *et al.* 1999, 2000). These findings are echoed in studies of pertussis dynamics in different countries (Hethcote 1998; Gomes *et al.* 1999; Bauch & Earn 2003; Broutin *et al.* 2005b).

To understand the stark contradiction between the dynamics predicted by deterministic models (rigidly annual epidemics) and those observed in data (a complex mixture of annual and multi-year outbreaks), Rohani *et al.* (1999) relaxed the assumption of determinism and examined event-driven stochastic models. These models exhibited spatio-temporal patterns that were broadly consistent with patterns in the England & Wales data. Since then, a number of authors have focused on this question and the general consensus appears to be that pertussis epidemics result from the interaction between seasonality, nonlinearity and, importantly, stochasticity (Hethcote 1998; Keeling *et al.* 2001; Rohani *et al.* 2002; Bauch & Earn 2003). This body of work has highlighted the importance of understanding whether transient dynamics following stochastic perturbations are sustained (if the deterministic attractor is very weakly stable) or are very short-lived (due to rapid contraction of trajectories onto the attractor)(Hastings & Higgins 1994; Hastings 2004; Coulson *et al.* 2004; Noonburg & Abrams 2005; Caswell & Neubert 2005). Other ecological case studies where the dynamical importance of the interaction between stochasticity and nonlinearity has been documented include laboratory populations of flour beetles (Cushing *et al.* 1998; Reuman *et al.* 2006), outbreaks of forest insects (Dwyer *et al.* 2004), island populations of soay sheep (Coulson *et al.* 2001; Grenfell *et al.* 1998; Brenton *et al.* 2006), insect host-pathogen interactions (Bjørnstad *et al.* 2001) and dungeness crab (Higgins *et al.* 1997).

To get a handle on the stability properties of the deterministic whooping cough model, detailed numerical studies have been carried out, concentrating on the fates of perturbations made to trajectories on the pertussis attractor (Keeling *et al.* 2001; Bauch & Earn 2003). It has been demonstrated that small perturbations generate transient dynamics that are multi-year, with a period of approximately three years. Bauch & Earn (2003) explained these effects can be detected in time-series data when there are *non-resonant peaks* in the power spectral density (as opposed to *resonant peaks* which are annual and are seasonally driven). Further examination of these dynamics revealed the presence of an unstable structure (Rohani *et al.* 2002; Greenman *et al.* 2005), which was termed the "invasion orbit" by Rohani *et al.* (2002),

primarily because its presence was demonstrated by examining the invasion of the disease into a wholly susceptible population. What has remained unclear since then is precisely what generated the invasion orbit.

In this chapter, we revisit this problem and examine the dynamical consequences of key model assumptions, other than determinism. Specifically, the models described above all assume that the instantaneous probability of leaving the latent and infectious class is constant, giving rise to latent and infectious periods that are exponentially distributed. This assumption is mathematically convenient but biologically unrealistic. In fact, the probability of staying in a class depends on the time already spent in that class, with waiting times that have a strong central tendency (Sartwell 1950; Simpson 1952; Bailey 1954; 1975; Gough 1977). The inclusion of appropriate distributions for the latent and infectious periods has been shown to be important in other contexts (Keeling & Grenfell 1997, 2002; Lloyd 2001a, 2001b; Wearing *et al.* 2005; Heffernan & Wahl 2006). We examine household data for pertussis incubation periods and find that a gamma distribution represents a significantly better fit than the exponential distribution. We then demonstrate that simple $SEIR$ models with more realistic distributions of latent and infectious period can explain the qualitative pattern of whooping cough epidemics. Furthermore, this framework sheds light on the genesis of the invasion orbit and its dynamical implications.

2.2 The model

We start with the classical $SEIR$ framework in which individuals are grouped into 4 epidemic classes: Susceptible, Exposed (latent), Infectious and Recovered. For a population of size N, disease dynamics are given by

$$\frac{dS}{dt} = \mu N - \left(\frac{\beta(t)I}{N} + \mu\right)S \tag{2.1a}$$

$$\frac{dE}{dt} = \frac{\beta(t)I}{N}S - (\sigma + \mu)E \tag{2.1b}$$

$$\frac{dI}{dt} = \sigma E - (\gamma + \mu)I \tag{2.1c}$$

$$\frac{dR}{dt} = \gamma I - \mu R. \tag{2.1d}$$

Here, μ gives the *per capita* birth and death rates. The average exposed and infectious periods are given by $1/\sigma$ and $1/\Gamma$, respectively. The contact rate, $\beta(t)$, is a function of time representing the aggregation of children in schools. We use term-time forcing (Schenzle 1984), which simply means that transmission rate is high during school term ($\beta(t) = b_0(1 + b_1)$), and low during the holidays ($\beta(t) = b_0(1 - b_1)$).

As mentioned above, this model explicitly assumes exponentially distributed latent and infectious periods. We will refer to the model given by equations (1-4) as $SEIR^e$. While in principle it is possible to incorporate any distribution into the model, from a computational perspective, the gamma distribution is especially convenient. Specifically, we can use the method of stages –also called the linear chain trick– whereby the latent and infectious periods consist of m and n sequential stages, respectively (Cox & Miller 1965; MacDonald 1978). The number of stages m and n affect the relative variation of the distribution with the coefficient of variation for the latent and infectious periods given by $1/\sqrt{m}$ and $1/\sqrt{n}$, respectively. When these shape parameters are equal to 1, we recover the classical exponentially distributed models and as they approach infinity, waiting times in each class become fixed.

The coupled ordinary differential equations describing the $SEIR$ model with gamma distributed latent and infectious periods are given by:

$$
\begin{aligned}
\frac{dS}{dt} &= \mu N - (\beta(t)\frac{I}{N} + \mu)S && \text{(2.2a)}\\
\frac{dE_1}{dt} &= \frac{\beta(t)I}{N}S - (m\sigma + \mu)E_1 && \text{(2.2b)}\\
\frac{dE_2}{dt} &= m\sigma E_1 - (m\sigma + \mu)E_2 && \text{(2.2c)}\\
&\vdots \\
\frac{dE_m}{dt} &= m\sigma E_{m-1} - (m\sigma + \mu)E_m && \text{(2.2d)}\\
\frac{dI_1}{dt} &= m\sigma E_m - (n\gamma + \mu)I_1 && \text{(2.2e)}\\
\frac{dI_2}{dt} &= n\gamma I_1 - (n\gamma + \mu)I_2 && \text{(2.2f)}\\
&\vdots \\
\frac{dI_n}{dt} &= n\gamma I_{n-1} - (n\gamma + \mu)I_n && \text{(2.2g)}\\
\frac{dR}{dt} &= n\gamma I_n - \mu R. && \text{(2.2h)}
\end{aligned}
$$

Henceforth, we will refer to the model given by equations (5-12) as $SEIR^{\Gamma}$

To estimate the parameters m and n, we examined household data on the incubation periods of whooping cough (Heininger *et al.* 1998). The data are plotted in figure 2.1, together with best fit exponential and gamma distributions, estimated using maximum likelihood. As evidenced by the maximum log-likelihood values, the gamma distribution (with $m + n = 5$ and a log-likelihood value of -636.42) represents a better fit to the data than the exponential distribution (with $m = n = 1$: -706.815). In the next section, we examine the dynamical consequences of different values of m and n, with a view to resolving the discussion concerning pertussis epidemics.

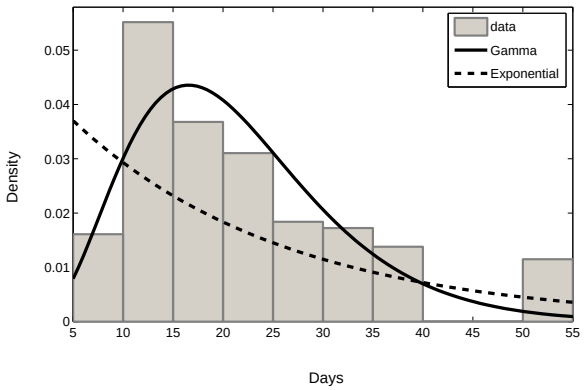


Figure 2.1: Frequency histogram showing household incubation data for whooping cough (data from Heininger *et al.* 1998). The dotted and solid lines represent best fits to exponential (m = 1) and gamma (m + n = 5) distributions, respectively. The log-likelihood score for $SEIR^e$ (-706.815) is smaller than $SEIR^\Gamma$ (-636.420)

2.3 Results

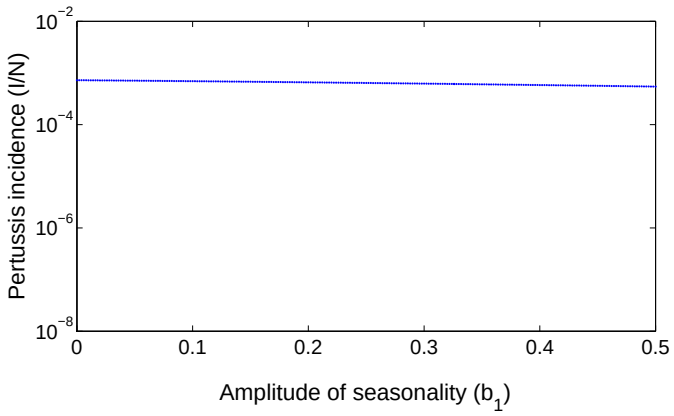
In figure 2.2a, we present a bifurcation diagram describing the dynamics of the seasonally forced $SEIR^e$ model as the amplitude of seasonality is varied. The ordinate shows pertussis incidence on January 1 of each year, hence annual cycles are represented by a single curve in the diagram. As noted in the Introduction, annual epidemics are predicted for all values of b_1 in the range (0, 0.5]. In contrast, as shown in Figure 2.2b, the gamma distributed model ($SEIR^\Gamma$) exhibits a range of dynamical behaviors, with the different colors corresponding to different stable solutions. When the amplitude of seasonality is small ($b_1 < 0.2$), annual epidemics are predicted. For b_1 in the range (0.2, 0.333], the annual cycle coexists with a three year cycle. Further increases in b_1 result in the coexistence of three attractors, with periods of 1, 3 and 4 years. Eventually, as b_1 exceeds approximately 0.44, the triennial cycle undergoes a cascade of period-doubling bifurcations, leading to a small window of chaotic epidemics. In instances where multiple stable attractors coexist with finely structured basins

of attraction, stochasticity can result in jumps between attractors (Earn *et al.* 2000). We plot the basins of attraction for $b_1 = 0.25$ and $b_1 = 0.4$ in the inset panels of Figure 2.2b.

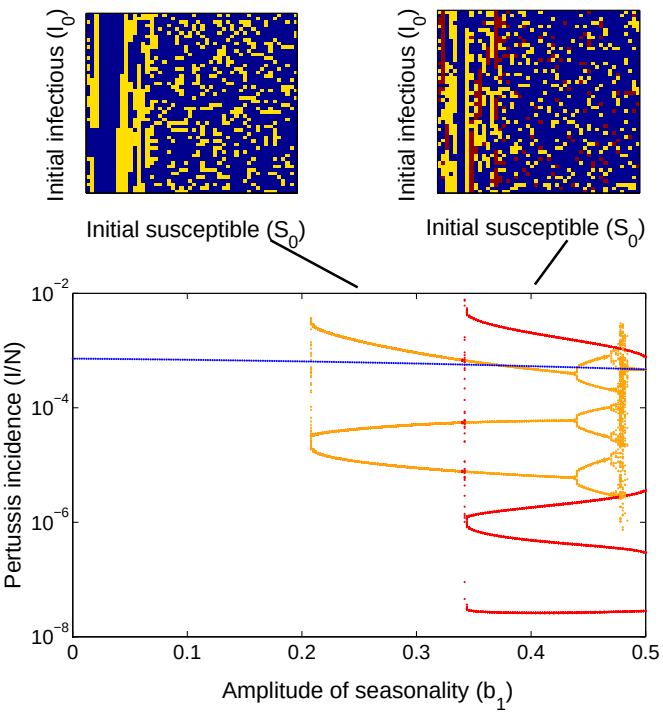
To investigate in more detail the relationship between the seasonal amplitude, the infectious period and the distribution of the latent and infectious periods, we carried out a series of bifurcation analyzes (Figure 2.3). Within the context of pertussis dynamics and the studies by Rohani *et al.* (2002) and Bauch & Earn (2003), perhaps the key findings are presented in figure 2.3a. In this figure, we show that the annual cycle is stable throughout the range of $b_1 - m$ parameter space we explored. For sufficiently large amplitude of seasonality, however, values of m, n exceeding unity give rise to the coexistence of the annual attractor with a triennial cycle. As the variance in the distribution of the latent/infectious periods decreases (with increasing m, n), the triennial cycle is observed with smaller levels of seasonality.

Figure 2.3a also demonstrates that for very high levels of seasonality and large values of the shape parameters, attractors with periods of 1, 2, 3 and 4 years coexist. The findings from this figure raise a key point: starting from the exponentially distributed model, the qualitative dynamics are highly sensitive to increases in the shape parameters (Lloyd 2001a). Once m, n exceed approximately 5, however, further increases yield incremental changes in the bifurcation structure of the model. Therefore, our results are largely insensitive to the precise values of m and n, as long as they exceed unity. Note that for simplicity, we set $m = n$ while generating the figure. However, our preliminary findings suggest very strongly that the distribution of the infectious period is overwhelmingly the key determinant of the dynamics, as might be expected intuitively (also see Blythe & Anderson 1988).

We also examined how the periodicity of epidemics is influenced by changes in the recovery rate (Γ) and the shape parameters, when the seasonal amplitude is small ($b_1 = 0.15$; Figure 2.3b) or large ($b_1 = 0.25$; Figure 2.3c). For reference purposes, it is worthwhile to note that for whooping cough, the recovery rate (Γ) is approximately 26 $year^{-1}$ and for measles 73 $year^{-1}$.



(a) Exponential SEIR.



(b) Gamma SEIR.

Figure 2.2: Bifurcation diagram for the seasonally forced (a) exponentially ($SEIR^e$) and (b) gamma-distributed ($SEIR^\Gamma$) models, showing whooping cough dynamics as a function of the amplitude of seasonality (b_1). In (a), an annual attractor is predicted over the entire interval $0 < b_1 < 0.5$. In (b), predominantly attractors with period 1, 3 and 4 years are observed. Above this bifurcation diagram, we plot the basins of attraction for $b_1 = 0.25$ and $b_1 = 0.4$. Here, the initial number of susceptibles varies from 10^3 to 10^5 and initial infectious numbers from 10^1 to 10^3. The annual attractor is depicted in blue, orange represents the 3-year cycle and red the four-year attractor. Seasonality was incorporated using term-time forcing (for details see Keeling *et al.* 2001). Other model parameters are: $\mu = 0.02yr^{-1}$, $N = 5 \times 10^6$, $1/\sigma = 8d$, $1/\gamma = 14d$, with $R_0 = 17$.

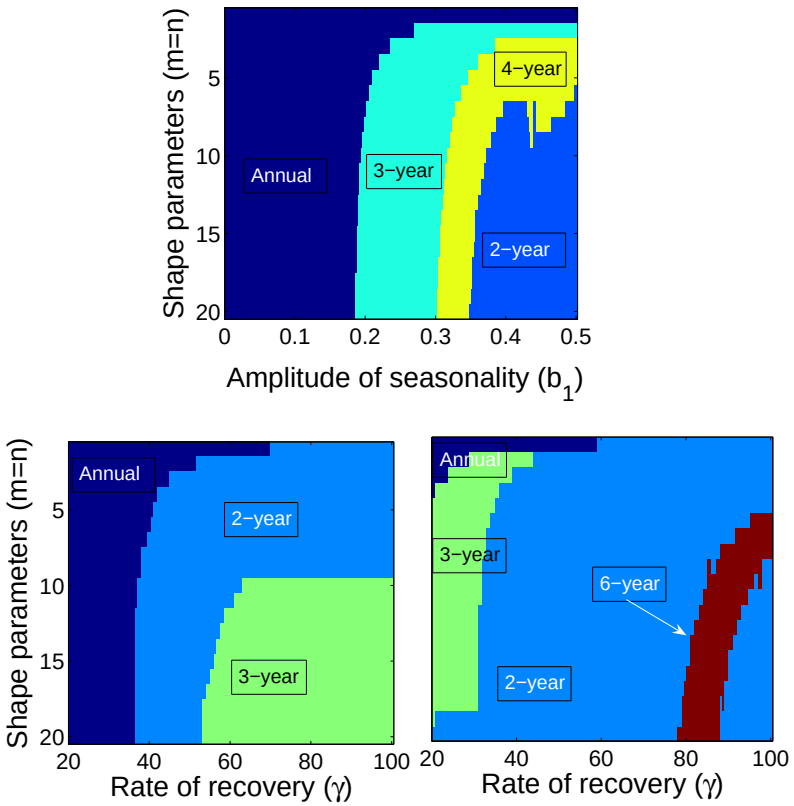
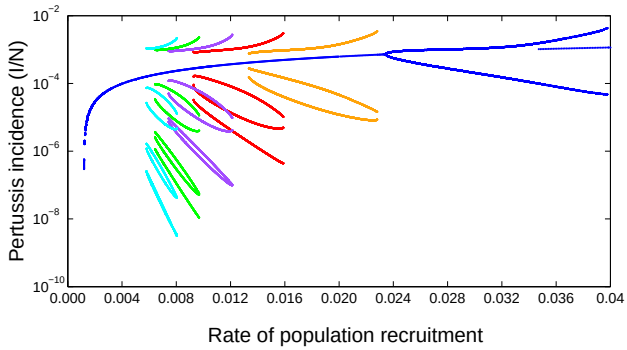


Figure 2.3: Periodicity of whooping cough dynamics depending on shape parameters ($m = n$) and amplitude of seasonality (top); shape parameters and recovery rate (Γ) when seasonality is $b_1 = 0.15$ (bottom left); and $b_1 = 0.25$ (bottom right). Except for the annual and 2-year regions in bottom left panel in which there is no overlap between them, the colored regions overlap, showing coexistence of different stable solutions as the control parameters are varied. Different initial conditions (susceptible and infectious) are used to capture the coexistence of different attractors.

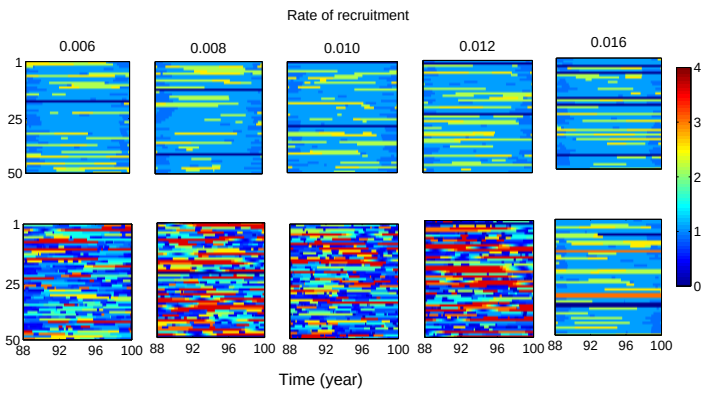
When seasonal forcing is relatively weak (Figure 2.3b), dynamics are always annual as long as the mean infectious period exceeds 10 days. For mean infectious periods of between 7 and 10 days, biennial outbreaks are predicted unless the exponential distribution is assumed, in which case annual cycles are observed. As the average infectious period becomes increasingly shorter (< 7 days), a stable triennial cycle coexists with the biennial cycle when m, n exceed 10. When the amplitude of seasonality is increased (Figure 2.3c), the region of stability for the triennial cycle expands considerably, coexisting with biennial ($30 < \Gamma < 100$) and annual ($20 < \Gamma < 30$) cycles. At the extremities, these boundaries are influenced by the shape parameters, though the most dramatic shifts occur once m, n exceed one. The figure also demonstrates that when mean infectious periods are very short (of the order of 3-4 days), windows of longer period oscillations are observed.

In order to understand how mass vaccination programs and systematic demographic trends affect pertussis epidemics, in Figure 2.4a we present a bifurcation diagram describing the dynamics of $SEIR^{\Gamma}$ with the susceptible recruitment rate as the control parameter (cf Earn *et al.* . 2000). To do so, we replaced the μN term in Equation (5) with $\mu^{'}(1-p)N$, where p represents the fraction of newborns immunized and $\mu^{'}$ denotes the modified *per capita* birth rate. The figure bears a striking resemblance to Figure 1 of Earn *et al.* (2000), which was produced with parameter values chosen to correspond to *measles*! Here, the default parameter values for pertussis in England & Wales in the 1950s correspond to the coexistence of the annual and triennial attractors ($\mu \sim 0.02$). Increases in the vaccination fraction or decreases in the per capita birth rates give rise to a cascade of bifurcations, resulting in longer epidemic periods. Similarly, "baby booms" result in biennial dynamics, which at first coexist with and eventually give way to annual cycles.

We examine the interaction between the dynamical complexity observed in Figure 2.4a and demographic stochasticity by formulating an exact stochastic analogue of $SEIR^{\Gamma}$ using Gillespies direct method (Gillespie 1977; see also Keeling & Rohani 2007). Figure 2.4b



(a)



(b)

Figure 2.4: (a) Bifurcation diagram depicting pertussis dynamics as the susceptible recruitment rate ($\mu^{'}(1-p)$) is varied in $SEIR^{\Gamma}$, using $m = 1, n = 5$ and $b_1 = 0.25$. For each value of the recruitment rate, annual samples of I/N from years 151 to 200 are plotted. (b) Analysis of periodicity using an event-driven stochastic $SEIR^{\Gamma}$ model, as the recruitment rate is varied from 0.006 to 0.016. The upper panel depicts results when simulations are started with initial conditions on the deterministic annual attractor (dark blue line in panel (a)). The lower panel presents comparable periodicities when simulations were started with initial conditions on the multi-year deterministic attractor (orange, red, purple, green and cyan lines in panel (a)). For each panel, the 50 rows represent the periodicity detected using wavelet analysis in a replicate stochastic time series from year 88 to 100. Inter-epidemic periods with highest spectral density at any instance are depicted, showing frequent switching between attractors, especially in the lower panels. Stochastic simulations included an immigration rate of 10 per million per year.

demonstrates observed periodicity through time in replicated simulated time-series for different recruitment rates and 50 stochastic replicates. In each instance, for comparison with case notifications data, we generate 12 years of weekly case notifications, after the first 88 years of simulations are discarded. The dominant period through time was then determined using wavelet analysis (Torrence & Compo 1998). The figure demonstrates that for a fixed recruitment rate, there is substantial dynamical variability across realizations. Some runs exhibit a number of switches between attractors over the 12-year time series (as depicted by abruptly changing dominant periods through time), while others show a constant period. It is important to note that the largest inter-epidemic period detected in these data is 4 years, even when the susceptible recruitment is very small and the deterministic model predicts 58 year cycles (Figure 2.4a). This may be in part due to the shortness of the time-series data or may, in fact, be due to the instability of the longer-period solutions when a small rate of immigration is incorporated into the stochastic simulations (cf, Alonso *et al.* . 2007).

Another approach to studying the intricate whooping cough epidemics is to examine the topology of the system in the vicinity of the deterministic attractors. This may be achieved by plotting invasion orbits (*sensu* Rohani *et al.* . 2002), which are generated by observing the trajectories of disease invasion as they approach stable points. We plot the annual samples of susceptibles and infectives after a single infectious individual is introduced into the population, starting simulations at different initial times ($t_0 \in [0, 1]$). Figure 2.5 shows pertussis invasion orbits in the absence of vaccination, using $SEIR^e$ and $SEIR^\Gamma$ with $b_1 = 0.25$. For each sub-figure, the green points represent orbits during the transient approach towards asymptotic dynamics. The large dots represent stable fixed points, which are color-coded as in previous figures (blue: annual cycle, orange: 3-year cycle). It is worth noting that the structures observed by carrying out this kind of invasion analysis are very similar to those obtained by simply studying the consequences of different initial conditions, as proposed by Rand & Wilson (1991).

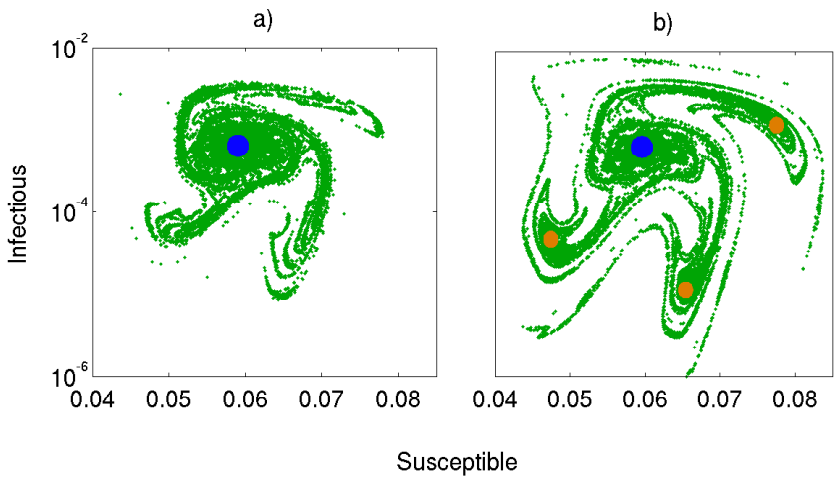


Figure 2.5: Invasion orbits of whooping cough in the pre-vaccine era. In (a) and (b), we depict the orbits generated by the $SEIR^e$ and $SEIR^\Gamma$ models, both with amplitude of seasonality $b_1 = 0.25$. The invasion orbits are captured by starting simulations at 1000 different initial times ($t_0 \in [0:1]$), with a single infective in a population of susceptibles (ie, $S(t_0) = 4,999,999$ and $I(t_0) = 1$). The proportion susceptible and infected at a fixed time are recorded every year and plotted, ignoring the first 20 years (due to transients). Model parameters are the same as in previous figures

The invasion orbits of $SEIR^e$ and $SEIR^\Gamma$ exhibit broadly similar structures, once we bear in mind the fact that changes in the infectious period distribution affect the amplitude of oscillations and therefore the size of the structures depicted. The key message from this figure can be gleaned by considering the panel on the right (Figure 2.5b). There is an aggregation of points near the annual attractor at the center of the invasion orbit. Additionally, there is a pronounced star shape with 3 prominent branches, corresponding to trajectories near the stable triennial cycle. Reductions in the amplitude of seasonality or increases in the variance of the infectious period (Figure 2.2a; Rohani *et al.* 2002) result in the loss of stability of the triennial solution. Importantly, however, the star shape is preserved. The clear implication of this observation is that the "invasion orbit" documented by Rohani *et al.* (2002) using $SEIR^e$ was simply the "ghost of a departed attractor" (as coined by Earn *et al.* 2000). The long transient dynamics documented in Rohani *et al.* 's stochastic simulations were due to the dynamical influence of destabilized attractors. These findings remain qualitatively unaffected when dynamics in the vaccine era are considered.

2.4 Discussion

For well over a century, epidemiologists have been working towards understanding the periodicities observed in the case notification data for childhood infections (Ransome 1880; Hamer 1897). It was the work of Soper (1929) on measles epidemics in Glasgow that, as far as we know, first demonstrated seasonal variation in transmission rates. The first systematic examination of seasonality in mathematical models arrived almost half a century later in Dietz's seminal paper in 1976 (Dietz 1976). Dietz examined conditions under which the periodic changes in contact rates can interact with the inherent oscillatory properties of the $SEIR^e$ system to produce either simple or subharmonic "resonance". One of the key questions that he raised in that paper concerned "whether the shape of the distribution of the latent or infectious period affects the resonance behavior" (Dietz 1976). In this paper, we

have returned to this question with the specific intention of examining whether changes to the assumed distribution of the infectious period can account for the observed epidemics of whooping cough.

Previous attempts at explaining long-term pertussis epidemics have argued for a significant role of stochasticity (Hethcote 1998; Rohani *et al.* 1999, 2002; Bauch & Earn 2003). Here, we have studied the dynamics of the $SEIR$ model, with waiting times in the latent and infectious classes determined by a gamma distribution, with distribution parameters estimated from household data. Our key finding is that a reduction in the variance in the infectious period gives rise to stable multi-year solutions. The implication of these results is that appropriately formulated deterministic $SEIR$ models are indeed capable of providing a qualitative explanation for observed pertussis dynamics.

This work also places into context the numerical observations of previous authors (Keeling *et al.* 2001; Rohani *et al.* 2002; Bauch & Earn 2003). The long and multi-year transients documented in exponentially distributed models of pertussis provided a compelling explanation for the patterns in case notifications data, but their origins remained unexplained. The results shown in Figures 2.5a & 2.5b show that the underlying cause of longer-period transient oscillations lies in the destabilization of the triennial attractor as shape parameters approach unity. This affects our interpretation of the role of stochasticity in this system. Using the classification of Millonas (1995), Coulson *et al.* (2004) suggested that epidemics of pertussis represent an example of "active" noise, where stochasticity interacts with the nonlinearity in the deterministic clockwork producing patterns that cannot result from either factor alone. Our findings suggest that possibly whooping cough dynamics may be the result of the less exciting "passive" treatment of noise, where stochasticity influences the transition among different deterministic states. Ultimately, the precise interpretation of this question relies on the accurate estimation of model parameters. This is especially true for the amplitude of seasonality. Some authors have suggested that for pertussis $b_1 \sim 0.15$ (Rohani *et al.* 2002;

Bauch & Earn 2003), while others have used age-structured arguments to propose a value closer to 0.25 (Keeling *et al.* 2001). Unbiased and confident estimation of this parameter is clearly a significant issue and we return to it below.

An intriguing aspect of this work is the bifurcation diagram shown in Figure 2.4a. We were surprised by the remarkable similarity between the bifurcation structure in this figure and that presented by Earn *et al.* (2000) using $SEIR^e$ in the context of explaining dynamical transitions in measles epidemics. While a detailed analysis of $SEIR^\Gamma$ parameterized for measles is lacking, the findings of Glass *et al.* (2003) and our own preliminary results suggest that the bifurcation diagram in Earn *et al.* (2000) is altered in significant ways when constant infectious periods are assumed. The subharmonic resonances resulting from the interaction between seasonality, the nonlinearity in transmission, and the distribution of the infectious period may be crucially determined by the epidemiological parameters (Greenman *et al.* 2005; Choisy *et al.* 2006). A systematic analysis of this issue remains a priority for future research.

The recent elegant work by Alonso *et al.* (2007) has argued for an alternative perspective on the epidemiology of childhood infectious diseases. These authors point out that the dynamics of such host-pathogen systems are determined both by the amplitude of seasonality in transmission and the tendency of the endogenous "clockwork" to amplify fluctuations. Focusing on the SIR^e framework, they derived an analytical expression for the power spectral density of the number of infectious and susceptible individuals, reaching the interesting conclusion that childhood infectious diseases are in regions of parameter space corresponding to high noise amplification. It would be interesting to re-examine Alonso *et al.* (2007)'s proposed endogenous stochastic resonance idea (or the similar concept of coherence resonance put forward by Kuske *et al.* 2007) when more realistic latent and infectious period distributions are assumed.

Finally, we have, thus far, side-stepped two potentially important aspects of whooping cough epidemiology and its modeling. The first is concerned with the ongoing debate about the frequency and consequences of loss of immunity acquired from natural infection and vaccination (Grenfell& Anderson 1989;Wirsing von Köenig *et al.* 1995; Broutin *et al.* 2004; Crowcroft & Pebody 2006). Clearly, of central importance is the question of the duration of immunity to pertussis, both derived naturally and as a result of vaccination. In the absence of unambiguous empirical information, parallel work by Wearing & Rohani (submitted) has attempted to address this question using an extended $SEIR$ model with waning immunity. The aim is to arrive at the most parsimonious estimate of the duration of immunity by matching global measures such as extinction thresholds and inter-epidemic periods with those estimated from the England & Wales case notifications data. Model predictions were found to be most consistent with incidence data for durations of immunity between 25 and 70 years, suggesting that models assuming long-term immunity (eg, $SEIR$ models) can still be useful in explaining pertussis epidemics. The second aspect relates to the robust estimation of pertussis model parameters. To address both of these topical and important questions, we are currently in the process of applying the "maximum likelihood via iterated filtering" methodology proposed by Ionides *et al.* (2006) to the waning immunity model of Wearing & Rohani, as well as the simpler $SEIR^e$ and $SEIR^\Gamma$ models discussed here. A better understanding of pertussis epidemics will be greatly facilitated by linking mechanistic transmission models with appropriate inferential methodologies.

Chapter 3

Seasonality of whooping cough: peak sizes and peak timing

3.1 Introduction

Seasonality is observed in a variety of natural systems. Many populations fluctuate with certain periods driven by different mechanisms, such as delayed density-dependence (Bjørnstad & Grenfell 2001), trophic interactions (Hanski *et al.* 1993), periodic environmental forcing (Bjørnstad & Grenfell 2001; Koelle *et al.* 2005; Hosseini *et al.* 2004), or parasites (Hudson *et al.* 1999; Tompkins *et al.* 2003). In epidemiology, cyclic outbreaks of childhood diseases are found in historical data providing ideal case studies to investigate the underlying mechanisms of disease dynamics. Whooping cough, for instance, exhibits outbreaks in many cities every one, two, three or more years (Hethcote 1998; Gomes *et al.* 1999; Rohani *et al.* 2000; Broutin *et al.* 2005). The current understanding is that seasonality via changes in contact rate due to school terms, plays an essential role in driving periodic disease dynamics (Bolker & Grenfell 1995; Dushoff *et al.* 2004).

As seasonality may be an important driver of cyclic outbreaks, understanding how seasonality affects periodicity is an important step toward understanding childhood disease dynamics. Furthermore, investigating how seasonality shapes patterns of infectious disease dynamics may aid in explaining the different patterns of disease dynamics observed in different regions because academic calendars are often region-specific. Studying seasonality is also of interest because seasonality may influence pathogen evolution (Ferguson *et al.* 2003; Kamo & Sasaki

2005). Ultimately, clarifying the links between seasonality and disease dynamics can give insights into planning vaccination strategies with appropriate vaccination timing (Altizer *et al.* 2006).

It is therefore not surprising that the effects of seasonal contact rates on patterns of disease dynamics have been explored in several studies (London & Yorke 1973; Fine & Clarkson 1982; Finkenstadt & Grenfell 2000; Keeling *et al.* 2001). One of the fundamental findings is that increases in the amplitude of seasonal forcing can generate a cascade of bifurcations, with large cycle associated with stronger seasonal forcing (Dietz 1976; Greenman *et al.* 2004; Altizer *et al.* 2006). Most of these studies involve models that use unrealistic sinusoidal forcing functions with a period of 1 year (Bolker & Grenfell 1993; Kamo & Sasaki 2005; Altizer *et al.* 2006). For childhood diseases like measles or whooping cough seasonal forcing arises from changes in the contact rate between the school terms and vacations which is more likely to have the form of a square wave with two uniform values: high transmission during school term and low otherwise (Keeling *et al.* 2001; Rohani *et al.* 2002).

Factors causing seasonal disease dynamics can be divided into intrinsic and extrinsic groups. One of the main intrinsic causes of seasonality in disease incidence is the variation in transmission term, as in the case of measles (Soper 1929; Fine &Clarkson 1982) and other diseases such as mumps, chickenpox and influenza (London & Yorke 1973; Bolker & Grenfell 1995; Dushoff *et al.* 2004). The seasonal transmission term is in turn caused by periodic changes of environmental condition, host behavior, host immune competence and abundance of vectors and non-human hosts (Altizer *et al.* 2006; Grassly & Fraser 2006). However, to a more detailed level, the reason why variation in outbreak sizes and the timing of a particular disease in a particular population differ from year to year is not always clear. For infections which are greatly affected by external factors, changes in seasonal environmental conditions are responsible for the changes in yearly disease outbreaks. For example, there are strong associations between epidemic cycles and cycles of temperature (Chew *et al.* 1998; Checkley *et al.*

2000), humidity (Nathason & Martin 1979; Chew *et al.* 1998), wind (Sultan *et al.* 2005) and so on. In contrast, for diseases such as childhood diseases which seasonality is more affected by intrinsic factors, the answer to the question why there is variation in outbreak sizes and timing between years is less obvious.

In addition, although there are an increasing number of studies on seasonality, only a small proportion can give a predictive insight into the disease dynamics under investigation. Among the studies that have yielded potentially predictive tools is a recent work on measles by Stone *et al.* (2007). The authors found that outbreak phase of a year has a correlation with outbreak size of the following year.

Data on whooping cough in England and Wales present various patterns of seasonality. In addition to the variety of inter-annual periodicity exhibited by the infection in different populations (cities), disease outbreaks within each population also vary from year to year. For example, historical data on whooping cough in Liverpool shows a great variation in outbreak amplitude and outbreak timing. Given that the population size and the seasonal transmission term can be assumed to be the same for every year during this period of time, the question is what is the cause of this variation? This is the driving question of my work on seasonality.

In this chapter, I analyze data on 3 representative cities in the UK (London, Liverpool and Sheffield) to seek an explanation for the differences in outbreak sizes and phases. As recruitment into the population which includes births and immigration, is thought to be the main force maintaining recurrent epidemics (Anderson & May 1991; Finkenstadt *et al.* Grenfell 1998; Earn *et al.* 2000), it is possible that variation in recruitment could be responsible for variation in outbreak properties. Thus, my hypothesis is that peak sizes and peak timing of whooping cough outbreaks are determined by population birth rates. Here

and henceforth, peak sizes are defined as the peak number of infected cases in an outbreak; and peak phase is the month in which outbreaks reach their peak.

3.2 Data analysis

Data on whooping cough dynamics in three UK cities: London, Liverpool and Sheffield (figure 3.1) are analyzed. These cities are chosen for several reasons. First, their population sizes are higher than the critical community size for whooping cough which is 200,000-300,000. This ensures that the effect of immigration on these cities is weak enough to be ignored in the analysis and modeling. Second, these three cities are a representative sample of all the cities with large population sizes with different outbreak timing. We also focus on disease dynamics before vaccination so that we can study the effects of birth rates without the complication of instant vaccination.

The average age at infection of whooping cough is 4-5 years (Anderson & May 1991). Thus, the characteristics of an outbreak (size and phase) in a particular year are not only influenced by the birth rates of that year but also by those 1, 2, 3, 4 years earlier. In order to take into account the effect of age at infection on the relationships, these birth rates are averaged for every 5 continuous years to be considered "corrected birth rates". A better way to correct birth rates is to use the age distribution of the disease. This will take into account different effects of different ages at infection. Incorporating age at infection requires more data on age distribution for the cities under investigation. Henceforth, we will ignore the annual birth rates and use the averaged birth rates of every 5 years in analysis.

Data filtering is essential to uncovering patterns in noisy data such as whooping cough case reports. This is partly due to the fact that stochasticity plays a large role in whooping cough dynamics (Rohani *et al.* 1999, 2002) and partly due to errors in reporting infected cases (Anderson & May 1991). Due to these factors, the real peaks might not exactly be the peek

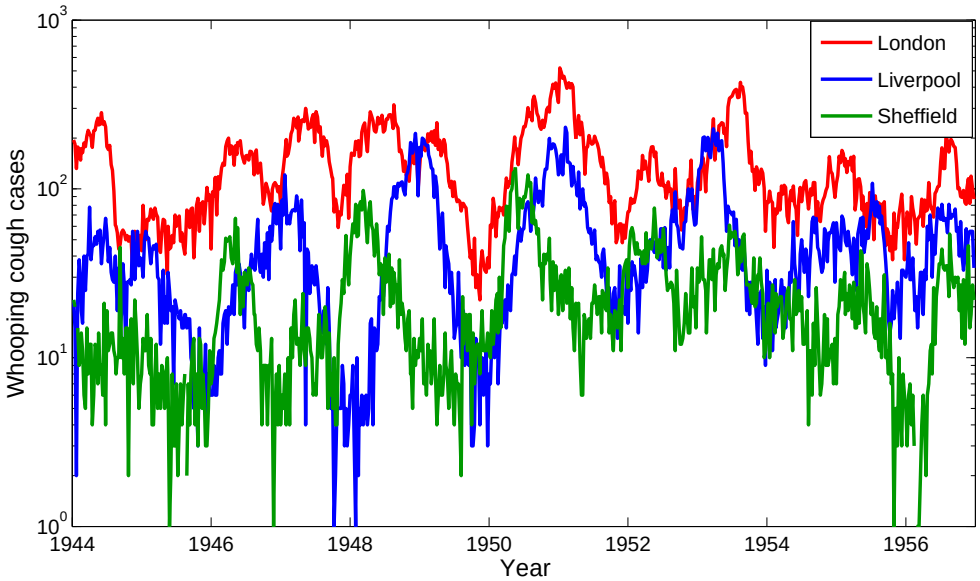


Figure 3.1: Weekly case reports for whooping cough in London (red), Liverpool (blue) and Sheffield (green) in the pre-vaccine era (1944-1956).

with the highest number of infected cases. To resolve this problem, in this chapter, we use two methods of filtering data: Band-pass filter and Fast Fourier filter.

A band-pass filter is a way to filter data with an appropriate range of frequencies. In other words, it will pass desired frequencies within a certain range and reject those outside that range. Figure 3.2 demonstrates this method for whooping cough weekly cases in Liverpool. The time series of whooping cough case reports in Liverpool is presented in blue and the corresponding time series after being filtered is in red. The range of frequencies kept in this particular time series depends on the periodicity of the time series. For example, as shown in figure 3.2, in the data regions of one-year and 2-year period, one-year and 2-year band-pass filters are respectively used. The arrows show the peak of outbreaks detected for the time-series presented.

Fast fourier filter (FFT filter) is one of low- and high-pass filtering devices in which the data, after being converted to frequencies using the Fast Fourier Transform, are filtered appropriately and then converted back using an inverse FFT. Here I design the FFT filter as a low-pass filter which will keep low frequencies (monthly, annual, biennial, etc) and reject high frequencies (weekly). Moreover, unlike the band-pass filter mentioned earlier, this filter is not specific to the time series or different intervals of the time series. Figure 3.3 is an example of data filtering using FFT filtering method for whooping cough in Liverpool. The blue time series shows the data before filtering and the red one after being filtered. In this figure, arrows are also used to indicate peaks of disease outbreaks.

After being filtered, the peaks of outbreaks in different years were analyzed focusing on 2 main characteristics: peak sizes and peak phase. Figure 3.4 presents results of using band-pass filter with a) relationship between peak sizes and birth rates; b) peak months across different birth rates; and c) the time interval between every 2 adjacent peaks versus the difference between the 2 corresponding birth rates.

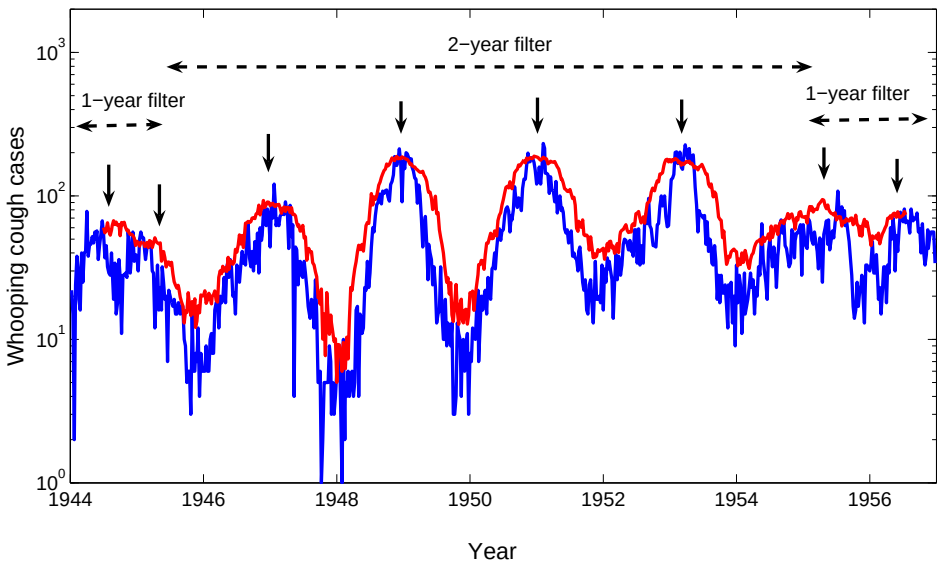


Figure 3.2: Data on whooping cough in Liverpool before (blue) and after being filtered (red) by band-pass filter. Specific band-pass filters are used for data with specific periodicity. The arrows show the peaks of outbreaks in different years which later were used in analysis.

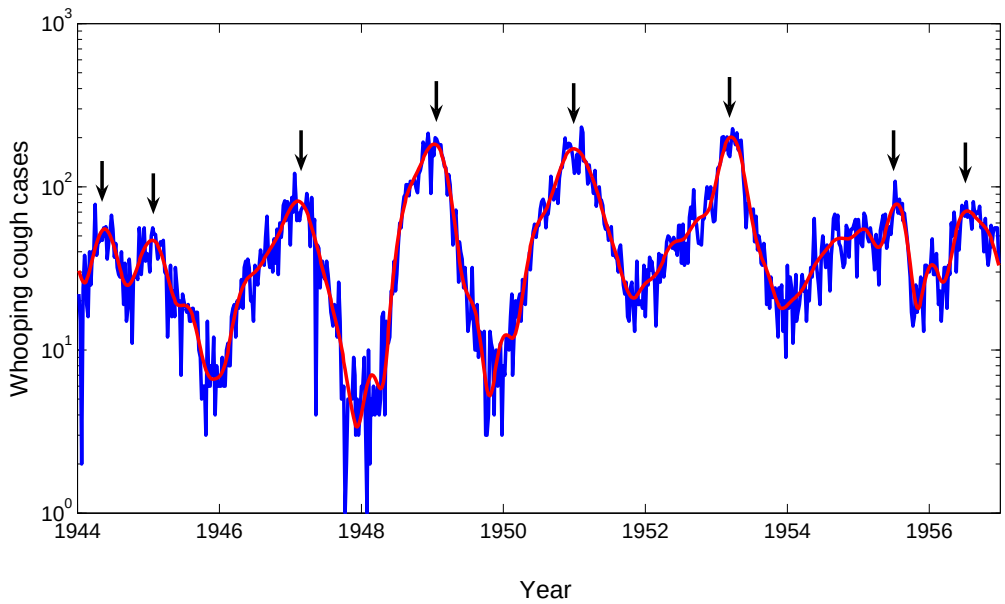


Figure 3.3: Data on whooping cough in Liverpool before (blue) and after being filtered (red) by FFT method. The arrows show the peaks of outbreaks in different years.

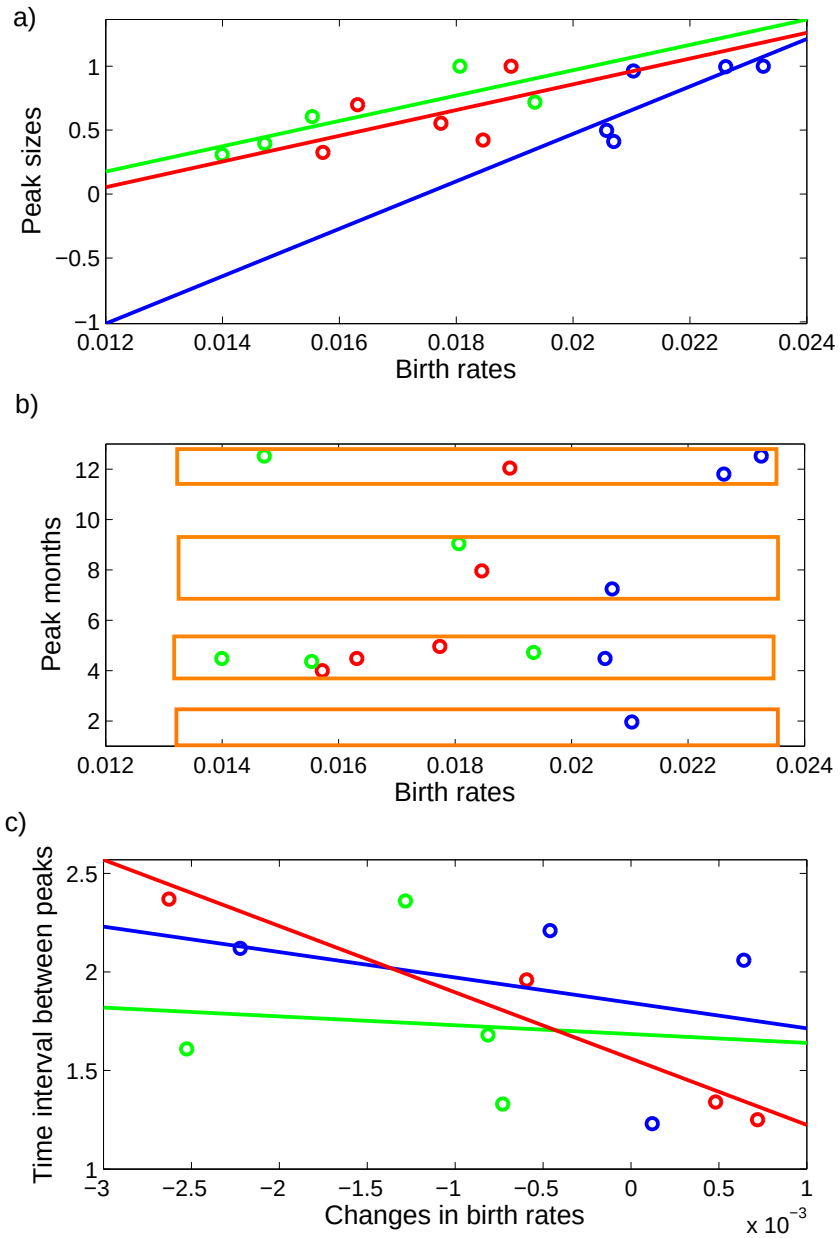


Figure 3.4: Data analysis of seasonality using band-pass filter for London (red), Liverpool (blue) and Sheffield (green). a) Relationship between peak sizes and birth rates. The p-values for the linear regression of London, Liverpool and Sheffield are 0.36; 0.13; 0.09, respectively. b) Peak months for different birth rates. c) The relationship between peak time intervals and changes in birth rates. The p-values for London, Liverpool and Sheffield are 0.03; 0.65; 0.91, respectively.

As high birth rates provide the population with a large number of susceptibles, the transmission rate should also be high. This, in turn should causes the outbreak to occur early and rapidly reach its peak. Thus, our hypothesis is that a high birth rate will associate with early peak phase. As we expected the positive correlation between peak sizes and birth rates was found (figure 3.4 a). Figure 3.4 b shows that the months in which outbreaks reach their peak are not randomly distributed. Peak months can be organized into 3 groups: winter, spring and fall as presented in the figure. The reason for this phase aggregation could be the seasonality of contact rate which is raised high at the beginning of two school terms (Fall and Spring) and low otherwise. It is also due to the difference between the number of students admitted in the Fall semester (larger) and that in the Spring semester (smaller).

In addition to the aggregated distribution of peak phases, a negative relationship was also detected between peak phase intervals (the time interval between the peaks of 2 adjacent outbreaks) and changes in birth rates (the difference in birth rates between 2 adjacent peaks). Although this is not statistically significant, the trends are obvious and consistent for these three cities and other large cities in the UK (not shown here). The correlation between peak interval and changes in birth rates can be explained using the same argument as used in explaining the relationship between peak sizes and birth rates. In other words, as birth rates increase, the rate at which susceptibles are exposed to infection is higher leading the outbreak to occur earlier. Thus, the outbreak will reach its peak sooner.

With the FFT filtering method, we also found the same patterns as found when using band-pass filter (figure 3.5). Although they are consistent, patterns found using FFT filters are not as strong as those found using band-pass filters. This is reasonable as the band-pass filters are designed individually for each time series and thus are more specific and precise.

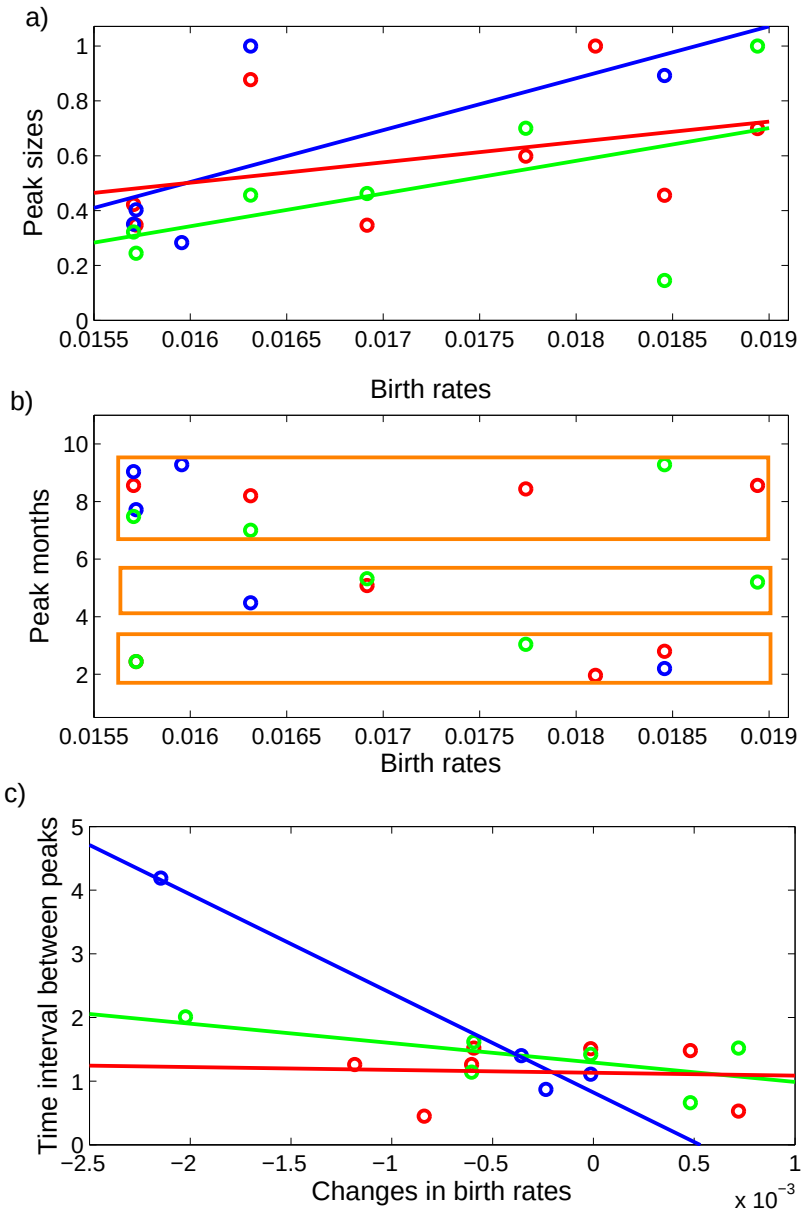


Figure 3.5: Data analysis of seasonality using FFT filter for London (red), Liverpool (blue) and Sheffield (green). a) Relationship between peak sizes and birth rates. The p-values for the linear regression of London, Liverpool and Sheffield are 0.36; 0.23; 0.22, respectively. b) Peak months for different birth rates. c) The relationship between peak time intervals and changes in birth rates. P-values are 0.88; 0.01; 0.16.

3.3 Model predictions

Using stochastic versions of the exponential $SEIR^e$ and the $SEIR^\Gamma$ model (described in detail in Chapter 2), we want to compare the patterns observed in the data with model predictions and compare between models. As seasonality is the important factor in this analysis and modeling, we use both term-time forcing and sinewave forcing to describe the shape of the contact rate function. The term-time forcing is as described before in Chapter 2 which basically resembles the real school terms and vacations in the UK and assigns a high contact rate for school days and low otherwise. The sinewave forcing contact rate assumes contact rate changes as a cosine function:

$$\beta = \beta_0(1 \pm \beta_1 \cos(2\pi t + w)).$$

Although a sinewave contact rate is not realistic, it is still used in many studies to describe the seasonal aggregation of children. The main reasons are its simplicity and mathematical convenience.

We run these models for 80 years using the average birth rate of $\mu = 0.02$ and then change the birth rate according to the birth rate data of the three cities London, Liverpool and Sheffield. The interval of these time series that represents the real birth rate changes will be analyzed and compared with the patterns observed in real whooping cough data.

Figure 3.6 shows the relationship between peak sizes and the corresponding birth rates from the different models. Similar to the real data, these models also show the positive correlation between peak sizes and birth rates. The correlations are not statistically significant but show a consistent trend, which is confirmed for nearly all repeated simulations of other cities in England and Wales. In addition, as stochasticity can greatly affect whooping cough dynamics, it is understandable that there is not a strong correlation. The differences between different models are small.

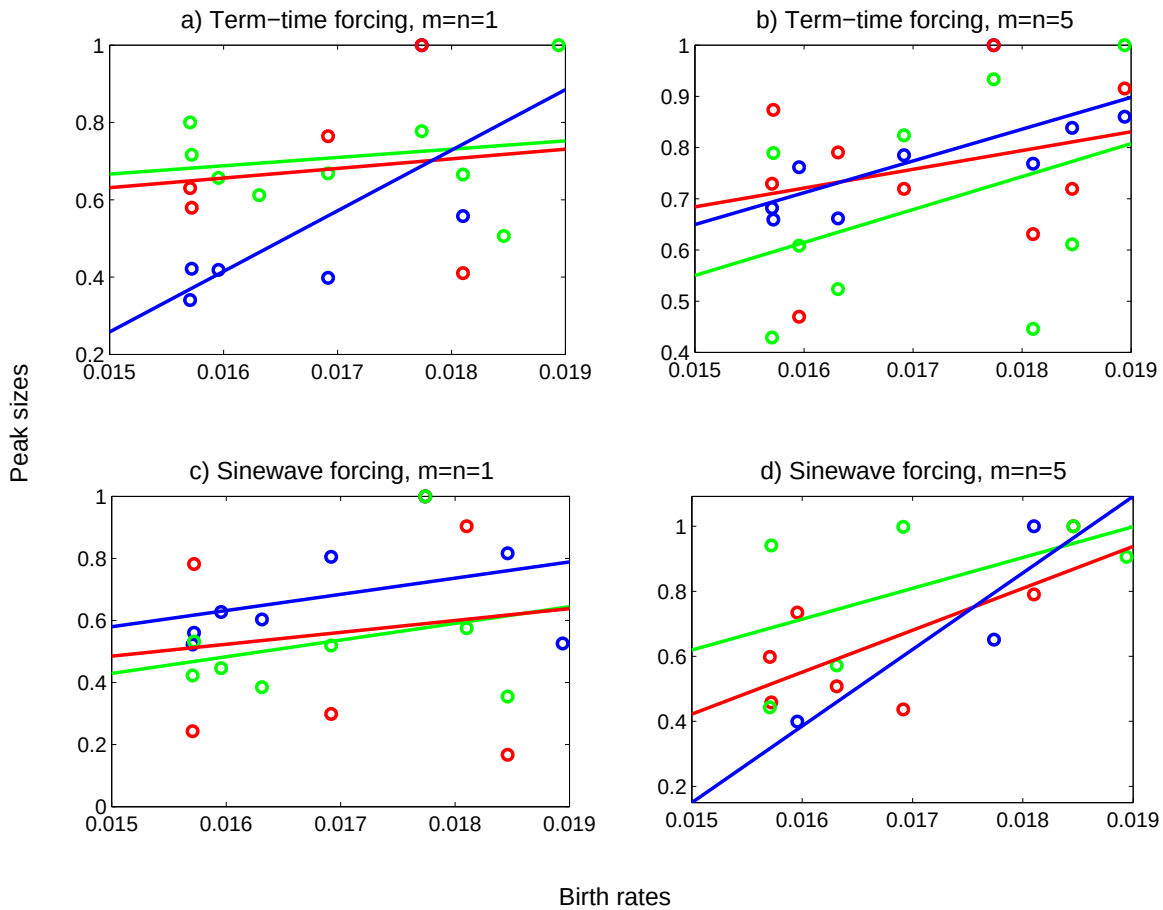


Figure 3.6: Relationship between peak sizes and birth rates predicted by different models. Each subfigure represents 3 stochastic realizations using London parameters. a) exponential $SEIR$ ($SEIR^e$) with term-time forcing contact rate (p values for red, blue and green line fittings are 0.87, 0.14 and 0.56, respectively). b) gamma $SEIR$ ($SEIR^\Gamma$) model with term-time forcing (p values are 0.45, 0.03 and 0.31). c) $SEIR^e$ with sinvewave forcing (0.82, 0.35 and 0.48) and d) $SEIR^\Gamma$ with sinewave forcing (0.07, 0.29 and 0.22)

The phases at which the outbreaks reach their peaks are also plotted across the range of birth rates for all the models (figure 3.7). These peak months can also be categorized into several season groups. Whereas models using term-time forcing predicts the three month groups winter, spring and fall as seen in the data, the models with sinewave forcing can only predict the spring and fall months. This result is not surprising given that the term-time forcing function is more realistic than the sinewave function. Compared to the $SEIR^e$, the $SEIR^\Gamma$ model achieves better results in capturing the patterns found in the data (figure 3.7a and 3.7 b) although this improvement is small.

In figure 3.8, we plot the time interval between every two adjacent peaks against the corresponding changes in birth rates. The same patterns as seen in the data are found. However, it should be noted that not all the stochastic replications of these models give the same results. In a few cases, negative relationships between these peak intervals and changes in birth rates are found. However, because in most of these cases the numbers of peaks are small and their dynamics are subject to stochasticity, it is possible that the reverse results could be found.

In addition to the models described above in which birth rates varied every year, we also investigate the effect of having constant birth rates on the sizes and phases of disease outbreaks. The term-time forcing $SEIR$ models are run for 200 years with the same birth rates using a Monte-Carlo method. Peak characteristics are recorded for the last year of the time series. The simulation is repeated for different birth rates. The peak characteristics including peak sizes and phases are plotted against the corresponding birth rates (figure 3.9).

There is clearly a strong positive correlation between birth rates and peak sizes of disease outbreaks. As the birth rate is higher, the recruitment into the population every year is higher providing the disease with more susceptibles and later producing more infected cases. It is almost obvious why the trend in this model is much stronger than that in the previous ones in which birth rates are changed every year. Firstly, the whole time series have either

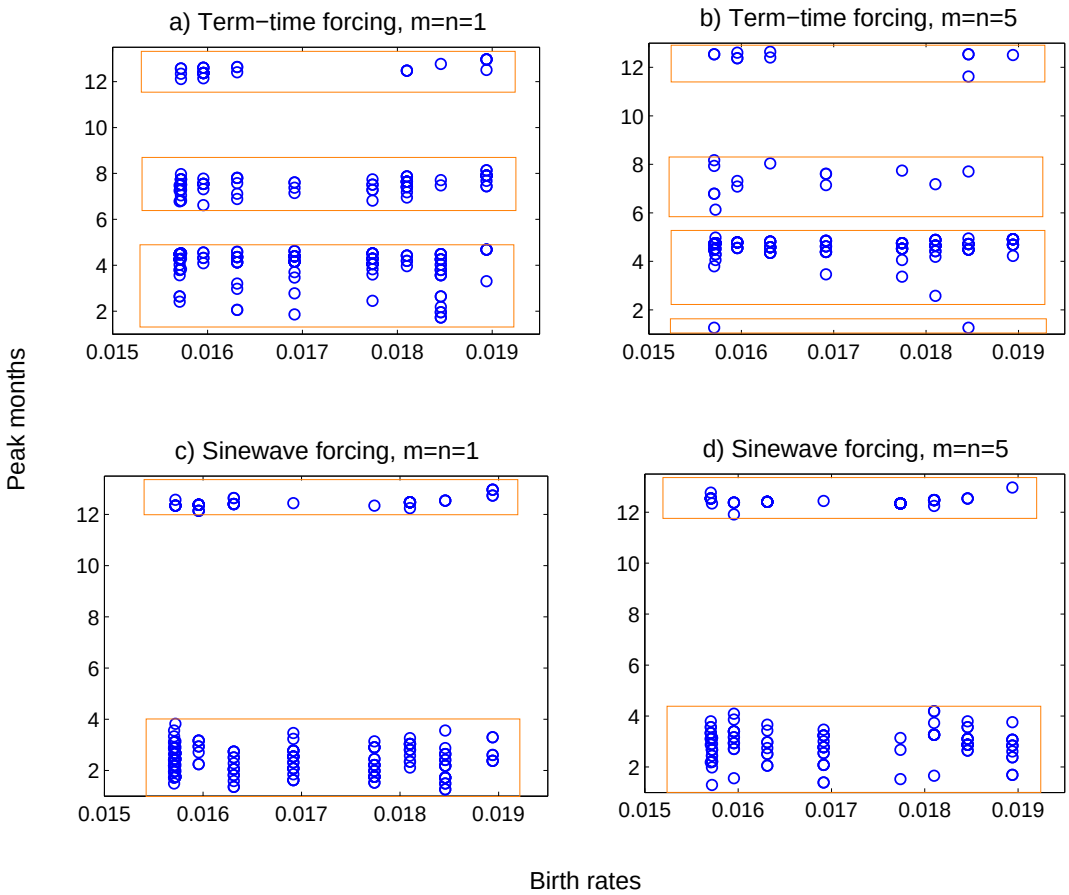


Figure 3.7: Peak months in which the outbreaks reach their peak across a range of birth rates predicted by different models. a) $SEIR^e$ with term-time forcing contact rate. b) $SEIR^\Gamma$ model with term-time forcing. c) $SEIR^e$ with sinewave forcing and d) $SEIR^\Gamma$ with sinewave forcing

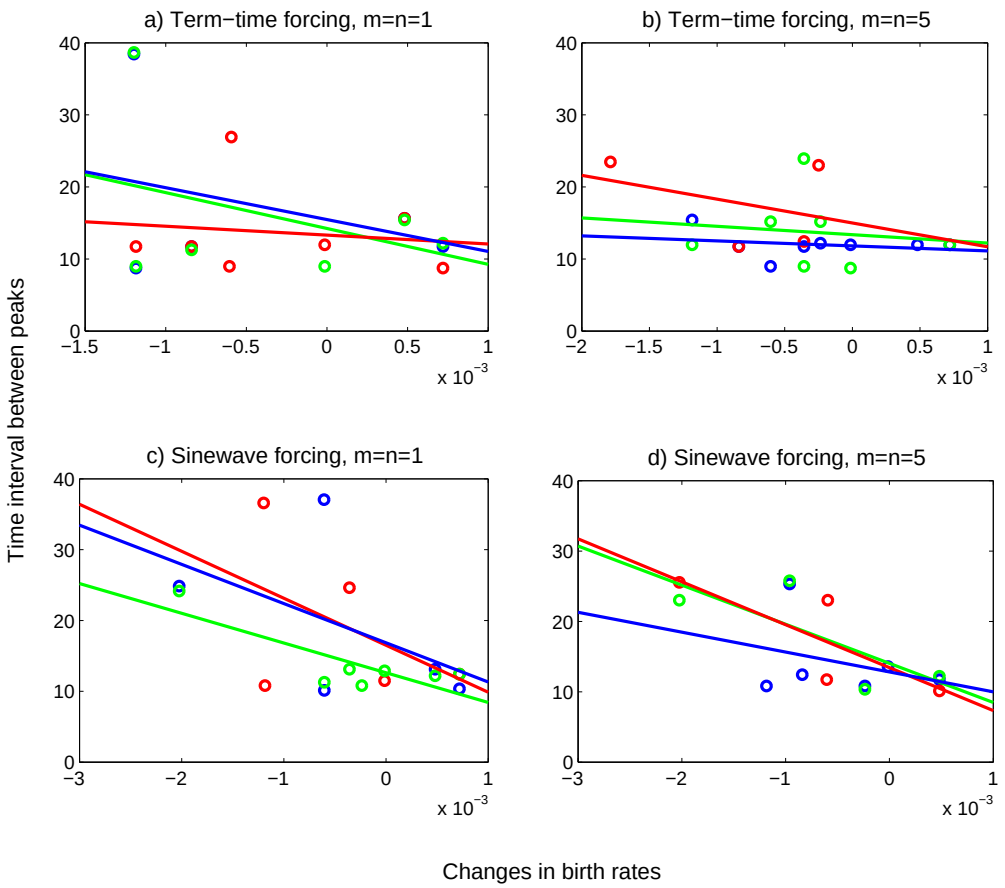


Figure 3.8: Relationship between peak time intervals and changes in birth rates predicted by different models: a) $SEIR^e$ with term-time forcing contact rate (p values for red, blue and green fittings are 0.54, 0.0.25 and 0.21, respectively). b) $SEIR^\Gamma$ model with term-time forcing (0.64, 0.54 and 0.78). c) $SEIR^e$ with sinewave forcing (0.41, 0.37 and 0.03) and d) $SEIR^\Gamma$ with sinewave forcing (0.20, 0.53 and 0.23)

high or low birth rates for all the years, not switching from one to another. Thus, the difference between having high birth rates and having low birth rates in this model is much larger. Secondly, unlike the previous models in which changing birth rates every year can be considered a perturbation to the system, the system here is more stable and thus the effect of stochasticity is less severe, which in turn leads to a more consistent pattern.

3.4 Discussion

Predictive methods can greatly help reducing the consequences and costs in preventing and controlling disease outbreaks. The outbreak characteristics that need to be predicted include its behavior (how the number of cases change over time); its magnitude (how big the outbreak will be); and timing: when it occurs or when it reaches the maximum number of cases (peaks). Answering these questions requires uncovering the driving mechanisms that cause seasonality of the disease dynamics.

Understanding the mechanisms affecting outbreak sizes and phases is critical in predicting future outbreaks. There is a variety of factors that can determine how large the outbreak will be and when it reaches its peak including the peak phases of the previous years, as in the case of measles (Stone *et al.* 2007), the level of population immunity and climate changes in cholera outbreaks (Koelle & Pascual 2004; Koelle *et al.* 2005), recruitment rates into the population (birth and immigration), population size, age-structure of the population and so on. It is important to have an assessment of how these factors influence the dynamics and shape the characteristics of outbreaks in the coming years. The complete assessment requires taking into account the effects of all these factors together.

In this chapter, using available data on birth rates and whooping cough infected cases in the pre-vaccine era, we show the trends in which birth rates in prior years affect peak sizes and phases in current years. The size of an outbreak increases as the birth rates of the previous

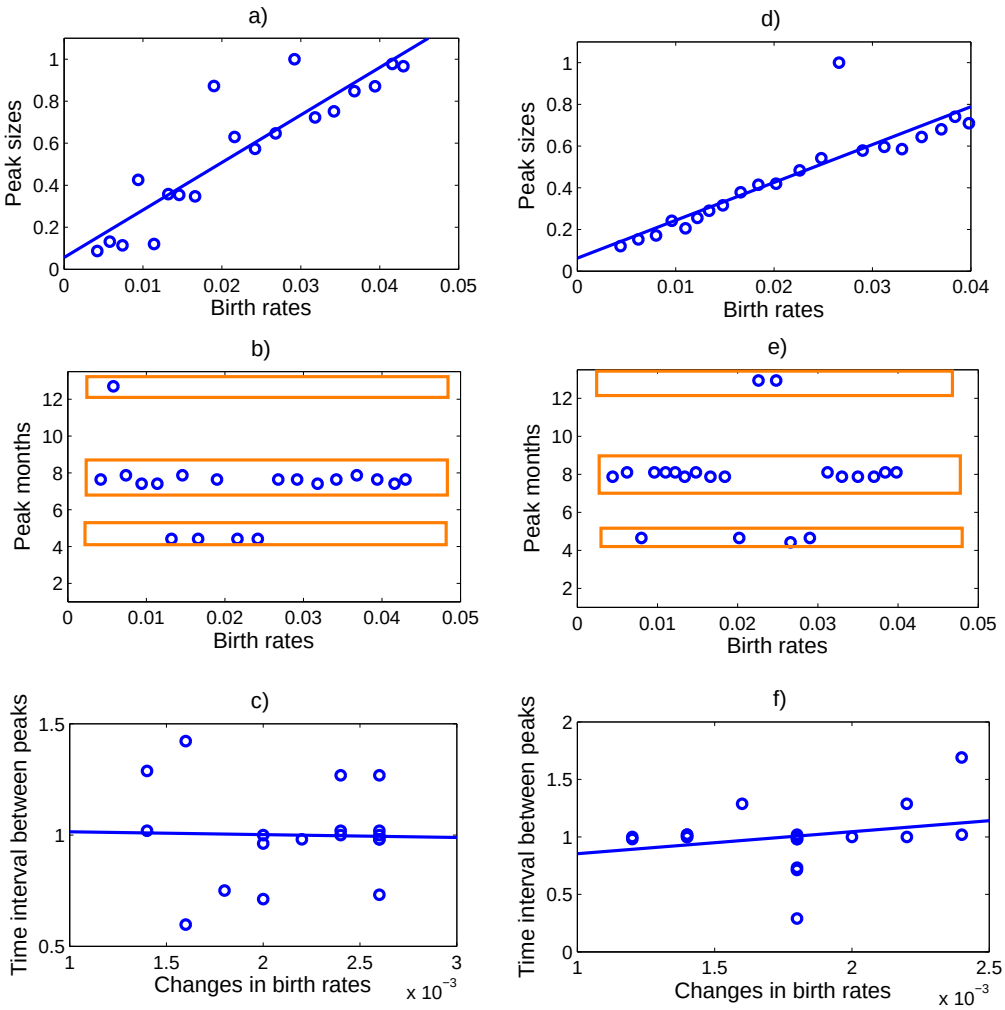


Figure 3.9: Seasonality predicted by models of constant birth rates. The left panels (a, b, c) are from $SEIR^e$ and the right panels (c, d, f) are from $SEIR^\Gamma$ model. a) and d) Relationship between peak sizes and birth rates (p values are 0.00 and 0.05); b) and e) Peak phases for different birth rates; c) and f) Relationship between peak time interval and changes in birth rate.

5 years increase. The phase of an outbreak is also dependent on these birth rates. Moreover, the phases can be categorized into time groups which are determined by the sudden influx of children into 2 school terms (Fall and Spring semester) after vacations.

Although in most of the cases, the linear regressions are not statistically significant, the trend is still clear and consistent. As there are only a few real peaks for each time series (regarding whooping cough data before vaccination), it is almost impossible to have a complete and strong analysis and understanding of the data. The consistency of the analyzes however suggest that there is a positive relationship between birth rates and peak size of outbreaks and a negative relationship between birth rates and outbreak phases in the range of birth rates available. However, further investigation is needed in order to have a complete and clear picture of these relationships.

Bibliography

[1] Alonso D, McKane AJ & Pascual M 2007 Stochastic amplification in epidemics. J. R. Soc. Interface. doi:10.1098/rsif.2006.0192.

[2] Altizer A, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P 2006 Seasonality and the dynamics of infectious diseases. Ecology Letters 9: 467-484.

[3] Anderson RM & May RM 1979 Population Biology of Infectious Diseases 1. Nature 280: 361-367.

[4] Anderson RM, May RM 1991 Infectious diseases of humans: Dynamics and control. Oxford: Oxford University press. 757p.

[5] Andrewathar HG and Birch LC 1954 The distribution and abundance of animals. University of Chicago Press, Chicago.

[6] Águas R, Gonçalves G and Gomes DGM 2006 Pertussis: increasing disease as a consequence of reducing transmission. Lancet infectious diseases 6 (2): 112-117.

[7] Bailey NTJ 1954 A statistical method of estimating the periods of incubation and infection of an infectious disease. Nature (London) 174: 139-140.

[8] Bailey NTJ 1975 "The mathematical theory of infectious diseases" 2nd ed., Griffin, London.

[9] Bauch CT and Earn DJD 2003 Transients and attractors in epidemics. Proceedings of the Royal Society of London Series B- Biological Sciences, 270: 1573-1578.

[10] Benton TG, Plaistow SJ, Coulson TN 2006 Complex population dynamics and complex causation: devils, details and demography. Proceedings of the Royal Society of London B, Biological Sciences 273: 1173-1181.

[11] Bernoulli D 1760 Mém. Math. Phys. Acad. R. Sci. Paris, 1-45.

[12] Bjørnstad ON and Grenfell BT 2001 Noisy clockwork:. time series analysis of population fluctuations in animals. Science 293, 638643.

[13] Bjørnstad ON, Grenfell BT and Finkenstadt BF 2002 Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. Ecological Monographs 72:169-184.

[14] Bolker BM and Grenfell BT 1993 Chaos and biological complexity in measles dynamics. Proceedings of the Royal Society of London B, Biological Sciences 251: 75-81.

[15] Broutin H, Rohani P, Guégan JF, Grenfell BT and Simondon F 2004 Loss of Immunity to Pertussis in a Rural Community. Vaccine 22: 594-596.

[16] Broutin H, Guégan JF, Elguero E, Simondon F, Cazelles B 2005a Large-scale comparative analysis of pertussis population dynamics: Periodicity, Synchrony, and Impact of vaccination. American journal of epidemiology, 161: 1159-1167.

[17] Broutin H, Mantilla-Beniers NB, Simondon F, Aaby P, Grenfell BT, Guégan JF, Rohani P 2005b Epidemiological impact of vaccination on the dynamics of two childhood diseases in rural Senegal Microbes and Infection 7: 593-599.

[18] Caswell H, Neubert MG 2005 Reactivity and transient dynamics of discrete-time ecological systems. Journal of difference equations and applications 11 (4-5): 295-310.

[19] Chan KS and Tong H 2001 Chaos: a statistical perspective. New York: Springer.

[20] Checkley W, Epstein L, Gilman R, Figueroa D, Cama RI, Patz JA Effects of El Niño and ambient temperature on hospital admissions for diarrheal diseases in Peruvian children. Lancet 2000; 355: 442-50.

[21] Chew F, Doraisingham S, Ling A, Kumarasinghe, G, Lee B 1998 Seasonal trends of viral respiratory tract infections in the tropics. Epidemiol Infect 121: 121-8.

[22] Choisy M, Guegan JF and Rohani P 2006 Resonance effects and the dynamics of infectious diseases. Physica D 223: 26-35.

[23] Coulson T, Rohani P, Pascual M 2004 Skeletons, noise and the population growth: the end of an old debate? Trends in Ecology and Evolution 19(7): 359-364.

[24] Cox DR and Miller HD 1965 "The theory of stochastic processes" Chapman and Hall, London.

[25] Crowcroft NS, Stein C, Duclos P, Birmingham M 2003 How best to estimate the global burden of pertussis? Lancet Infect Dis 3: 413-18.

[26] Crowcroft, NS, and Pebody RG 2006 Recent Developments in Pertussis Lancet 367:1926-1936.

[27] Daszak P, Cunningham AA, Hyatt AD 2000 Emerging infectious diseases of wildlife-Threats to biodiversity and human health. Science 287: 443-449.

[28] Dietz K 1976 The incidence of infectious diseases under the influence of seasonal fluctuations. Lect. Notes Biomath., 11: 1-15.

[29] Dushoff J, Plotkin JB, Levin SA and Earn DJD 2004 Dynamical resonance can account for seasonality of influenza epidemics. Proc. Natl. Acad. Sci. USA, 101: 16915-16916.

[30] Dwyer G, Dushoff J and Yee SH 2004 The Combined Effects of Pathogens and Predators on Insect Outbreaks. Nature 430: 341-345.

[31] Dye C & Williams BG 2000 Criteria for the control of drug-resistant tuberculosis. PNAS 97: 8180-8185.

[32] Earn DJD, Rohani P & Grenfell BT 1998 Persistence, Chaos and Synchrony in Ecology and Epidemiology. Proc. R. Soc. Lond. B. 265: 7-10.

[33] Earn DJD, Rohani P, Bolker BM, Grenfell BT 2000 A simple model for complex dynamical transitions in epidemics. Science 287: 667-670.

[34] Edmunds WJ, Van de Heijden OG, Eerola M and Gay NJ 2000 Modelling rubella in Europe. Epidemiology and Infection 125: 617-634.

[35] Ferguson NM, Galvani AP & Bush RM 2003 Ecological and immunological determinants of influenza evolution. Nature 422: 428-433.

[36] Fine PME, Clarkson J 1982 Measles in England and Wales- 1: an analysis of factors underlying seasonal patterns. Int. J. Epidemiol. 11: 5-14.

[37] Frankish H 2003 death tolls continue to climb in Congo Ebola outbreak. Lancet 361: 1020.

[38] Glass K, Xia Y & Grenfell BT 2003 Interpreting time-series analyses for continuoustime biological models-measles as a case study. Journal of Theoretical Biology 223: 19-25.

[39] Gomes MC, Gomes JJ and Paulo AC 1999 Ditheria, pertussis, and measles in Portugal before and after mass vaccination: a time series analysis. European Journal of Epidemiology, 15:791-798

[40] Gough KJ 1977 The estimation of latent and infectious periods. Biometrika 64: 559-565.

[41] Greenman J, Kamo M & Boots M 2004 External forcing of ecological and epidemiological systems: a resonance approach. Physica D, 190: 136-151.

[42] Grenfell B & Anderson R 1989 Pertussis in England and Wales: An Investigation of Transmission Dynamics and Control by Mass Vaccination. Proc. R. Soc. Lond. B. 236: 213-252.

[43] Grassly NC & Fraser C 2006 Seasonal Infectious Disease Epidemiology. Proceedings of the Royal Society of London - Biological Sciences 273: 2541-2550.

[44] Guris D., Strebel PM, Bardenheier, *et al.* 1999 Changing epidemiology of pertussis in the United States: increaseing reported incidence among adolescents and adults, 1990-1996 1999 Cin Infect Dis 28: 1230-37.

[45] Guzman MG & Kouri G 2002 Dengue: an update. Lancet Infectious Diseases 2: 33-42.

[46] Gzyl A, Augustynowicz E, Rabczenko D 2004 Pertussis in Poland. Int J Epidemio 33: 358-65.

[47] Hanski I., Turchin P., Korpimaki E., Henttonen H. 1993 Population oscillations of boreal rodents- Regulation by mustelid predators leads to chaos. Nature 364: 232-235

[48] Hastings A, Higgins K 1994 Persistence of transients in spatially structured ecological models. Science 263: 1133-1136.

[49] Hastings A 2004 Transients: the key to long-term ecological understanding? TRENDS in Ecology and Evolution 19, 39-45.

[50] Heffernan JM & Wahl LM 2006 Improving Estimates of the Basic Reproductive Ratio: Using Both the Mean and the Dispersal of Transition Times. Theoretical Population Biology 70: 135-145.

[51] Heininger U *et al.* 1998 Comparative Efficiency of the Lederle/Takeda 1998 Acellular Pertussis Component DTP (DTaP) Vaccine and Lederle Whole-cell component DTP vaccine in German children after household exposure. Pediatrics 102: 546-553.

[52] Heathcote HW 1998 Oscillations in an endemic model for pertussis. Canadian Applied Mathematics Quarterly, 6:61-88.

[53] Henson SM, King AA, Costantino RF, Cushing JM, Dennis B and Desharnais RA 2003 Explaining and predicting patterns in stochastic population systems. Proc. R. Soc. Lond B 2003 207: 1549-1553.

[54] Higgins K, Hastings A, Sarvela JN and Botsford LW 1997 Stochastic dynamics and deterministic skeletons: Population behavior of Dungeness crab. Science 276: 1431-1435.

[55] Hope Simpson RE 1952 Infectiousness of communicable diseases in the household, Lancet 263: 549-554.

[56] Hudson PJ, Dobson AP, Newborn D 1999 Prevention of population cycles by parasites removal. Science 282: 2256-2258.

[57] Ionides EL, Breto C & King AA 2006 Inference for Nonlinear Dynamical Systems. Proceedings of the National Academy of Sciences of the USA 103: 18438- 18443.

[58] Kamo M, Sasaki A 2005 Evolution toward multi-year periodicity in epidemics. Ecology letters 8:378-385.

[59] Keeling MJ, Grenfell BT 1997 Disease extinction and community size: Modeling the persistence of measles. Science 275: 65-67.

[60] Keeling MJ & Rohani P 2007 Modelling Infectious Diseases: In Humans and Animals. Princeton University Press: Princeton.

[61] Keeling MJ, Rohani P, Grenfell BT 2001 Seasonally forced disease dynamics explained by switching between attractors. Physica D 148: 317-335.

[62] Koelle K and Pascual M 2004 Disentangling extrinsic from intrinsic factors in disease dynamics: a nonliear time series approach with an application to cholera. Am. Nat., 163: 901-913.

[63] Koelle K, Pascual M, Yunus M 2005 Pathogen adaptation to seasonal forcing and climate change. Proc. R. Soc. B 272: 971-977.

[64] Kuske R, Gordillo LF & Greenwood G 2007 Sustained Oscillations via coherence resonance in SIR. J. theor. Biol. 245: 459-469.

[65] Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D, Murray M 2003 Transmission dynamics and control of severe acute respiratory syndrome. Science 300: 1966-1970.

[66] Lloyd AL 2001a Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods.Proc R Soc Lond Biol Sci 268: 985-993.

[67] Lloyd AL 2001b Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics. Theoretical Population Biology 60: 59-71.

[68] Lloyd AL & Jansen VAA 2004 Spatiotemporal Dynamics of Epidemics: Synchrony in Metapopulation Models. Mathematical Biosciences 188:1-16.

[69] Lloyd AL and May RM 1996 Spatial heterogeneity in epidemic models. J. theor. Biol. 179: 1-11.

[70] de Melker HE, Coyn-van Spaendonck MAE, Rumeke HC, *et al.* 1997 Pertussis in the Netherlands: an outbreak despite high levels of immunization with whole cell vaccine. Emerg Infect Dis 3: 175-78.

[71] Millonas M 1995 Fluctuations and Order. The New Synthesis, Springer Verlag.

[72] Nåsell I 2005 A New Look at the Critical Community Size for Childhood Infections. Theoretical Population Biology 67: 203-216.

[73] Nicholson AJ 1957 The self-adjustment of populations to change. Cold Spring Harbor Symposium on quantitative Biology 22:153-172.

[74] Noonburg EG, Abrams PA 2005 Transient dynamics limit the effectiveness of keystone predation in bringing about coexistence. American naturalist 165; 3: 322-335.

[75] Ntezayabo B, De serres G, Duval B 2003 Pertussis resurgence in Canada largely caused by a cohort effect. Pediatr Infect Dis J 22: 22-27.

[76] Obenauer JC, Denson J, Mehta PK, Su XP, Mukatira S, Finkelstein DB, Xu XQ, Wang JH, Ma J, Fan YP, Rakestraw KM, Webster RG, Hoffmann E, Krauss S, Zheng J, Zhang ZW, Naeve CW 2006 Large-scale sequence analysis of avian influenza isolates. Science 311: 1576-1580.

[77] Olsen B, Munster VJ, Wallensten A, Waldenstrom J, Osterhaus ADME, Fouchier RAM 2006 Global patterns of influenza A virus in wild birds. Science 312: 384-388.

[78] Quinn TC, Overbaugh J 2005 HIV/AIDS in women: An expanding epidemic. Science 308: 1582-1583.

[79] Rand DA and Wilson HB 1991 Chaotic Stochasticity: A Ubiquitous Source of Unpredictability in Epidemics. Proceedings of the Royal Society of London Series B - Biological Sciences 246:179-184.

[80] Reuman DC, Desharnais RA, Costantino RF, Ahmad OS & Cohen JE 2006 Power spectra reveal the influence of stochasticity on nonlinear population dynamics. PNAS 103: 18860-18865.

[81] Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, Leung GM, Ho LM, Lam TH, Thach TQ, Chau P, Chan KP, Leung PY, Tsang T, Ho W, Lee KH, Lau EMC, Ferguson NM, Anderson RM 2003 Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. Science 300: 1961-1966.

[82] Rohani P, Earn DJD and Grenfell BT 1999 Opposite Patterns of Synchrony in Sympatric Disease Metapopulations. Science 286: 968-971.

[83] Rohani P, Earn DJD and Grenfell BT 2000 Impact of immunisation on pertussis transmission in England and Wales. Lancet 355: 285-286.

[84] Rohani P, Keeling MJ, Grenfell BT 2002 The interplay between Noise and Determinism in Childhood Diseases. American Naturalist 159: 469-481.

[85] Sartwell PE 1950 The distribution of incubation periods of infectious disease. Am J Hyg 51: 310-318.

[86] Schenzle D 1984 An age-structured model of pre- and post-vaccination measles transmission. IMA Journal of Mathematics applied in Medicine and Biology. 1:169:191.

[87] Simpson REH 1952 Infectiousness of communicable diseases in

[88] Tompkins D. M., White A. R., Boots M. 2003 Ecological replacement of native red squirrels by invasive greys driven by disease. Ecology Letters 6: 189-196 the household (measles, chikenpox and mumps). Lancet 263:549-554.

[89] Torrence C and Compo GP 1998 A Practical Guide to Wavelet Analysis. Bulleting of the American Meteorological Society. 79: 61-78.

[90] Walker PR, Worobey M, Rambaut A, Holmes EC & Pybus OG 2003 Epidemiology: Sexual transmission of HIV in Africa. Nature 422: 679.

[91] Wearing HJ, Rohani P, Keeling MJ 2005 Appropriate models for the management of infectious diseases. PloS Med 2(7): e174.

[92] WHO 1999 Weekly epidemiological record relevé epidémiologique hebdomadaire. WHO report 74: 137-144.