*Appl. Statist.* (2000)
**49**, *Part* 2, *pp.* 187–205

# Time series modelling of childhood diseases: a dynamical systems approach

Bärbel F. Finkenstädt and Bryan T. Grenfell

*University of Cambridge, UK*

[Received June 1998. Revised July 1999]

**Summary.** A key issue in the dynamical modelling of epidemics is the synthesis of complex mathematical models and data by means of time series analysis. We report such an approach, focusing on the particularly well-documented case of measles. We propose the use of a discrete time epidemic model comprising the infected and susceptible class as state variables. The model uses a discrete time version of the susceptible–exposed–infected–recovered type epidemic models, which can be fitted to observed disease incidence time series. We describe a method for reconstructing the dynamics of the susceptible class, which is an unobserved state variable of the dynamical system. The model provides a remarkable fit to the data on case reports of measles in England and Wales from 1944 to 1964. Morever, its systematic part explains the well-documented predominant biennial cyclic pattern. We study the dynamic behaviour of the time series model and show that episodes of annual cyclicity, which have not previously been explained quantitatively, arise as a response to a quicker replenishment of the susceptible class during the baby boom, around 1947.

*Keywords*: Cyclicity and seasonality in measles populations; Dynamical models in epidemics; England and Wales measles data; Population dynamics of childhood diseases; Susceptible–exposed–infected–recovered model

## 1. Introduction

The cyclic recurrence of childhood epidemics in unvaccinated populations provides one of the best-documented phenomena in population dynamics. Before the advent of vaccination, measles was a major source of mortality in developed countries. This, along with its ease of diagnosis, prompted reliable case notification schemes. The resulting rich epidemiological database, coupled with the relatively simple natural history of measles infection, make it the paradigm for studies of epidemic dynamics.

The advent of mass vaccination in the late 1960s reduced both average measles incidence and the relative amplitude and regularity of major epidemics (Anderson and May, 1991; Bolker and Grenfell, 1996). Measles is now at a very low level in the UK (Miller, 1996). However, the disease causes considerable mortality in malnourished young children in developing countries (McLean and Anderson, 1988a, b), and epidemics are still a significant public health problem in developed countries (Morse *et al.*, 1994). Understanding the dynamics of the disease is therefore undoubtedly important from a public health point of view (Black, 1984).

*Address for correspondence*: Bärbel F. Finkenstädt, Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK.
E-mail: B.F.Finkenstadt@Warwick.ac.uk

 0035–9254/00/49187

The spectacular dynamic epidemic pattern of measles is also intrinsically interesting and has attracted the attention of many researchers. The simple natural history of the disease has led to a family of relatively realistic mathematical models. The simplest description of measles infected dynamics is the susceptible–exposed–infected–recovered (SEIR) model which is expressed as a set of four ordinary differential equations. The underlying mechanism is fundamentally non-linear owing to the structure of the contact process between susceptible and infected individuals (Kermack and McKendrick, 1933; Bartlett, 1957, 1960; Bailey, 1975; Dietz, 1976; May, 1986; Anderson and May, 1991; Grenfell and Dobson, 1995). Measles infection is also characterized by marked heterogeneities in infection, related to the spatial and social aggregation of hosts (Anderson and May, 1991). In terms of models for large scale dynamics, the major heterogeneity is a strong concentration of infection in children during school-terms (Fine and Clarkson, 1982). Schenzle (1984) addressed this issue by developing a realistic age-structured (RAS) model. This accurately captures many features of measles dynamics, notably a tendency for biennial epidemics. The RAS model (Schenzle, 1984; Grenfell *et al*., 1994a; Grenfell, Kleczkowski, Gilligan and Bolker, 1995; Keeling and Grenfell, 1997) assumes different contact rates for the various age classes, each of which exhibits SEIR dynamics.

However, it is very difficult to forge a direct statistical link between measles time series and the SEIR or RAS models because of the present state of complexity of these models. The gap between the data analysis and the theoretical developments arises from the following problems.

(a) Only one state variable, namely the number of infected individuals, is observable. Moreover, this variable is under-reported by the number of notified cases (Fine and Clarkson, 1982).
(b) Relatively little is known about the age structure of the infected population through time.
(c) Model parameters such as birth-rates and vaccination uptake rates change with time (Grenfell *et al*., 1994a).
(d) The SEIR or RAS models are continuous dynamical systems, whereas the data are discrete time measurements.

Sometimes, pre-vaccination measles epidemic cycles in large cities were irregular in duration or amplitude. This has prompted an extensive search for non-linearity and chaos in measles dynamics (Schaffer and Kot, 1985; Olsen and Schaffer, 1990; Sugihara *et al*., 1990; Ellner, 1991; Nychka *et al*., 1992; Grenfell, 1992; Tidd *et al*., 1993; Ellner and Turchin, 1995). Many of the model- and data-oriented studies were stimulated by Takens's (1981) embedding theorem, which enables a reconstruction of the characteristic properties of the attractor from a single observed time path of the dynamical system. From the reconstruction one can, in principle, extract measures such as Lyapunov exponents and study the predictability on or near the attractor (Casdagli, 1989). Although stimulating discussions between population dynamicists about whether measles dynamics are chaotic or not, the methods have little to say about the structure of the data-generating mechanism.

More recently, researchers have tackled the measles data from a more mechanistic point of view. Ellner *et al*. (1998) have approached the problem of reconstructing the dynamics of the unobserved susceptible state variable and have followed an interesting semiparametric approach to model the time series of cases. The work presented here is based on an analysis of measles time series by Fine and Clarkson (1982). They used a deterministic discrete time epidemic model with variable transmission parameter to analyse measles data from England

and Wales for the period 1950–1965. Assuming homogeneous mixing between susceptible and infected individuals they obtained convincing estimates for the disease transmission rates. Later, Mollison and Din (1993) analysed the dynamic behaviour of the Fine and Clarkson (1982) model, concluding that it is not compatible with the data as it fails to explain the observed biennial cycle.

This raises a fundamental question: can the data be modelled by a parametric time series model that captures the basic SEIR mechanism? The aim of the present study is to address this question. We focus on fitting a dynamic recursive relationship that rigorously uses mechanistic rules from the SEIR model. Our modelling approach is based on a stochastic version of the Fine and Clarkson (1982) model, which is refined to allow for non-homogeneous mixing. Births are the fundamental 'fuel' that drives measles epidemics. To understand how changes in birth-rates affect measles dynamics we therefore base our modelling approach on a reconstruction of the susceptible population, which allows for birth-rate variations. We study the dynamic behaviour of the model and, in particular, the role of the birth-rate as a bifurcation parameter. Because of its unification of stochastic time series and epidemic modelling approaches, we call the new formulation the time series–susceptible–infected–recovered or TSIR model.

## 2. Data and background

Fig. 1 shows a flow diagram of the mechanism underlying the theoretical SEIR model (Anderson and May, 1991). The population is divided into four compartments comprising individuals who are susceptible (S) after birth and gradually pass through the exposed (E), infected (I) and recovered (R) states. For a directly transmitted viral disease, the transmission of infection between infectious and susceptible individuals is a non-linear function reflecting the contact process between individuals (Heesterbeek and Roberts, 1995). We assume that the transmission rate is time varying, as the contact rates of children vary over the school year (London and Yorke, 1973).

For measles, the exposed and infectious period, of around 2 weeks, is followed by lifelong immunity (Black, 1984). Essentially, infection occurs in epidemics, which extinguish themselves as waves of infection immunize susceptible individuals. Subsequent epidemics can occur only after susceptible densities are replenished by births. In England and Wales before vaccination this pattern gave rise to a regular biennial cycle, characterized by an alternation between major and minor epidemic years (Fig. 2(a)). However, an episode of annual cyclicity is apparent from 1945 to 1950 coinciding with the time of high birth-rates (Fig. 2(b)). Previous cross-sectional studies (Finkenstädt and Grenfell, 1998; Finkenstädt *et al.*, 1998), of data at the individual city level, have shown that a high birth-rate magnifies the minor

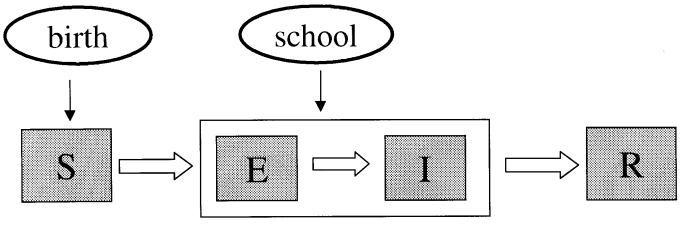


**Fig. 1.** Flow diagram of the SEIR compartmental model: S, susceptible; E, exposed; I, infected; R, recovered

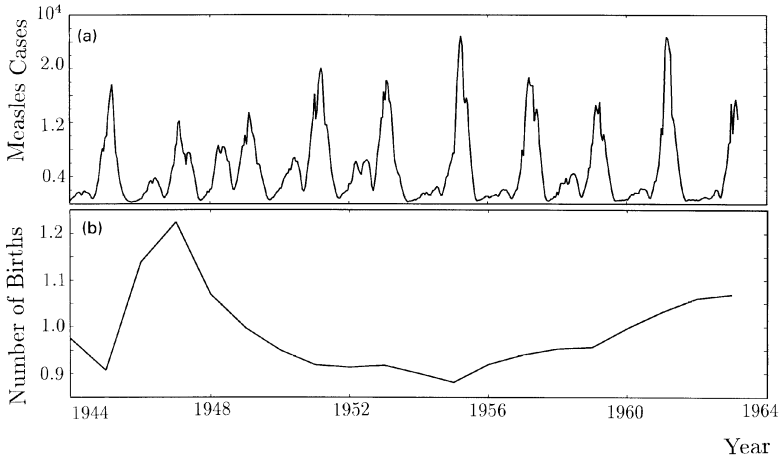


**Fig. 2.** Time series plot of (a) biweekly measles notifications and (b) yearly births for the aggregate of 60 cities in England and Wales, from 1944 to 1964

epidemics because of a quicker replacement of susceptible individuals after major epidemics (Grenfell *et al.*, 1994a, b; Grenfell, Bolker and Kleczkowski, 1995a). The rise of the minor epidemic de-emphasizes the biennial pattern, producing a predominantly annual cyclicity. These empirical findings make a compelling case that births should play a key role in a time series model for measles.

The analysis centres on the weekly notified measles cases for the aggregate of 60 cities in England and Wales from 1944 to 1964 (Fig. 2(a)), before the advent of vaccination. These are official notifications, taken from the Registrar General's *Weekly Reports*. Annual births-rates are taken from the *Annual Reports* of the Registrar General and are divided into appropriate subintervals of the year, assuming a constant birth-rate within the year.

## 3. The time series–susceptible–infected–recovered model

For measles, the characteristic timescale of the disease, i.e. the duration of the transition from infection to recovery and lifelong immunity, is 2 weeks (Black, 1984). The data are therefore aggregated into biweekly time steps. An infected person in biweek $t$ can hence be related to a contact between that individual as a susceptible and an infected individual in biweek $t-1$. Consider the following two-dimensional discrete time stochastic epidemic model for a childhood disease that entails lifelong immunity (see also Fine and Clarkson (1982)):

$$I_t = r_t I_{t-1}^{\alpha_1} S_{t-1}^{\alpha_2} \epsilon_t, \tag{1}$$

$$S_t = B_{t-d} + S_{t-1} - I_t + u_t. \tag{2}$$

Here $I_t$ and $S_t$ denote the number of infected and susceptible individuals respectively. Equation (1) describes the transmission of the infection through the multiplicative interaction between susceptible and infected individuals. The parameters $\alpha_1$ and $\alpha_2$ are mixing parameters of the contact process (Lui *et al.*, 1987). The case $\alpha_1 = \alpha_2 = 1$ corresponds to the standard assumption of homogeneous mixing (de Jong *et al.*, 1995). The transmission parameter $r_t$ is a factor of proportionality that is time varying with period 1 year. Equation (2) formulates the recursive relationship for the susceptibles class recruited by births $B_{t-d}$ and

depleted by the infected individuals leaving the class. The parameter $d$ denotes a small delay time allowing for the length of maternally derived immunity in infants before they become fully susceptible. It is known to be around 8 biweeks for measles (Anderson and May, 1991). We assume that $E(u_t) = 0$ and $V(u_t) = \sigma_u^2$, and $E\{\ln(\epsilon_t)\} = 0$, and $V\{\ln(\epsilon_t)\} = \sigma_\epsilon^2$ for the noise. We assume multiplicative rather than additive noise in equation (1) because preliminary analyses show that the variance increases with abundance.

Time series measurements are available on births $B_t$ and reported cases $C_t$, where the latter generally underestimate (under-report) true cases. To fit system (1)–(2) to the data, we proceed in two stages. We first use equation (2) to estimate the level of reporting and to reconstruct the unobserved susceptible class. We then fit the transmission equation (1) using the reconstructed susceptible dynamics as a covariate. Throughout the analysis we check our methods with simulated data from the RAS model (Keeling and Grenfell, 1997), where all state variables can be simulated as a function of epidemiological parameters and an assumed birth-rate time series.

### 3.1. Reconstruction of the susceptible dynamics

The number of true cases $I_t$ is related to the reported cases $C_t$ by

$$I_t = \rho_t C_t, \tag{3}$$

where $\rho_t$ is the reporting rate at time $t$. The number of true cases is under-reported if $\rho_t > 1$ and is fully reported if $\rho_t = 1$. We assume that $\rho_t$ has a probability function with expected value $E(\rho_t) = \rho$. Incorporating equation (3) into equation (2) gives

$$S_t = B_{t-d} + S_{t-1} - \rho_t C_t + u_t. \tag{4}$$

Let $E(S_t) = \bar{S}$ so that $S_t = \bar{S} + Z_t$ with $E(Z_t) = 0$. The deviations $Z_t$ from their mean follow the same recursive relationship as $S_t$,

$$Z_t = B_{t-d} + Z_{t-1} - \rho_t C_t + u_t. \tag{5}$$

A successive iteration of equation (5) for an initial condition $Z_0$ yields

$$Z_t = Z_0 + \sum_{i=1}^{t} B_{i-d} - \sum_{i=1}^{t} \rho_i C_i + \sum_{i=1}^{t} u_i. \tag{6}$$

Equation (6) shows that the susceptible class essentially accounts for the temporal balance between all births entering and infected individuals leaving the compartment. Also, if cases were not corrected by their reporting level, $Z_t$ would not be stationary, as the difference between cumulative births and cases grows unboundedly in the presence of under-reporting.

For ease of notation, let

$$X_t = \sum_{i=1}^{t} C_i, \qquad Y_t = \sum_{i=1}^{t} B_{i-d}, \qquad U_t = \sum_{i=1}^{t} u_i, \qquad R_t = \sum_{i=1}^{t} (\rho_i - \rho) C_i$$

for $t = 1, \ldots, n$. Then

$$\begin{aligned} R_t &= R_{t-1} + (\rho_t - \rho) C_t, \\ U_t &= U_{t-1} + u_t \end{aligned} \tag{7}$$

are random walk processes. Since their conditional mean is determined by the previous value, i.e. $E(R_t|R_{t-1}) = R_{t-1}$ and $E(U_t|U_{t-1}) = U_{t-1}$, these processes may exhibit *prolonged* temporal

shifts away from mean 0. Using the notation introduced above, equation (6) can be re-arranged as

$$Y_t = -Z_0 + \rho X_t + R_t + Z_t - U_t. \tag{8}$$

For the almost noise-free case ($U_t \approx 0$) and constant reporting rate ($R_t \approx 0$), equation (8) is a simple linear regression relationship between cumulative births $Y_t$ and cumulative cases $X_t$ with constant slope $\rho$. Then the unobservable susceptible dynamics $Z_t$ are obtained exactly as the regression remainder. We first illustrate this with simulations from the RAS model (Keeling and Grenfell, 1997).

*3.1.1. Example: validation of the susceptible reconstruction against the realistic age-structured model data*

We used stochastic (Monte Carlo) simulations of the RAS model to generate time series of susceptible densities and biweekly cases. Full details of the model and simulations are given in Keeling and Grenfell (1997). The data from the stochastic RAS model represent an almost noise-free realization at constant reporting rate. Fig. 3 illustrates the regression relationship between cumulative cases and births for two different birth time series. The reconstructed variable $Z_t$ is the difference between $\hat{Y}_t$, the linear predictor obtained from the regression in equation (8), and the observed $Y_t$. Fig. 3(b) demonstrates that the linear regression relationship (8) is not affected by pronounced temporal changes in the birth impulse. Fig. 4 shows that the regression residuals of model (8) perfectly reconstruct the susceptible dynamics. The empirical correlation with the true $S_t$ was 0.9986. The slope was estimated as 1.76 which corresponds to a reporting probability of 56.8% whereas the true rate was 56%. For RAS data at the full reporting level, the slope was estimated as 1.004. The reconstruction for the susceptible individuals works extremely well, even though the TSIR model ignores the concentration of infection in school-aged children assumed in the RAS model.

Fig. 5(a) clearly shows that the remainder of the linear regression fitted to the observed data for England and Wales suffers from local shifts in the mean. We suppose that the effects from temporal fluctuations of the reporting rate $\rho_t$ are more severe than from the external noise $u_t$. Since the medical diagnosis of measles was relatively straightforward, we assume that the reporting rate mainly reflects the frequency at which infected children were sent to a doctor. The temporal fluctuations of the reporting rate may thus be influenced by various time-varying factors such as the state of the epidemic, reports in the media, family behaviour, school attitudes and the introduction of the National Health Service in the UK in 1948 (Fine and Clarkson, 1982).

Incorporating the recursive formulations (7) for $R_t$ and $U_t$, equation (8) can be rearranged as

$$\begin{aligned} Y_t &= R_{t-1} - U_{t-1} - Z_0 + \rho X_t + (\rho_t - \rho)C_t + Z_t - u_t, \\ &= R_{t-1} - U_{t-1} - Z_0 - (\rho_t - \rho)X_{t-1} + \rho_t X_t + Z_t - u_t. \end{aligned} \tag{9}$$

The last formulation suggests the use of locally linear estimation methods in estimating the reporting rates $\rho_t$ and the susceptible dynamics $Z_t$. Note that the conditional mean of equation (9) given the previous states $R_{t-1}$ and $U_{t-1}$ is

$$E(Y_t | R_{t-1}, U_{t-1}) = R_{t-1} - U_{t-1} - Z_0 + \rho_t X_t. \tag{10}$$

Allowing for temporal shifts in the conditional mean, we treat the term $R_{t-1} - U_{t-1} - Z_0$

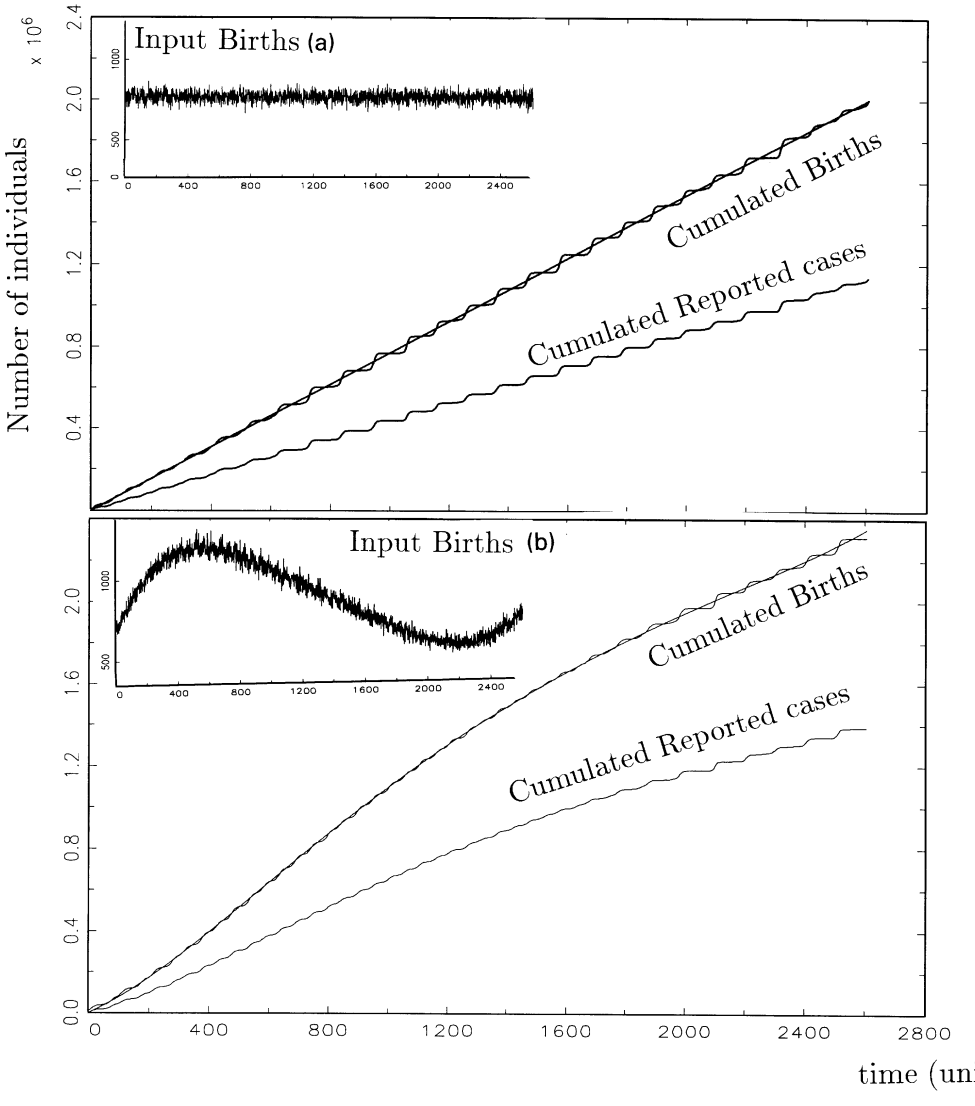


**Fig. 3.** Reconstruction of the susceptible population from simulated data from the RAS model for two different birth time series as shown in the upper left-hand corner of (a) and (b) (in both graphs, the lower oscillatory line plots the cumulative reported cases at a 56.8% rate of under-reporting; the upper non-oscillatory line is the time series of cumulative births $Y_t$; the predictor $\hat{Y}_t$ obtained from the linear regression is the oscillatory line projected onto the cumulative births; the susceptibility variable $Z_t$ is the difference between $Y_t$ and the linear predictor)

as a temporally varying intercept and approximate relationship (9) by *localized* linear regressions of $Y_t$ on $X_t$ in neighbourhoods of $X_t$ with slope $\rho_t$,

$$Y_t = \text{int}_t + \rho_t X_t + Z_t - u_t. \tag{11}$$

To fit equation (11), we apply locally linear regressions as described in Fan and Gijbels (1996). The method allows for a straightforward estimation of the local slopes $\rho_t$. However,

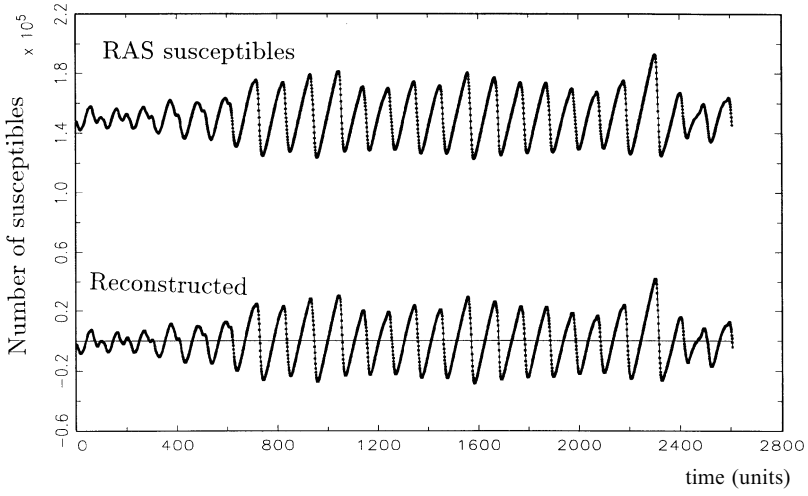


**Fig. 4.** Time plot of susceptible individuals from the RAS model and reconstructed $Z_t$ from the linear regression

further work not documented here shows that other local regression concepts such as splines also work.

The parameter that tunes the size of the neighbourhood (bandwidth) must be chosen carefully. Automatic selection procedures such as cross-validation are not useful here, as they seek to explain the cyclic pattern in the remainder as part of the regression curve and therefore select a parameter that reduces the remainder to white noise. It is desirable to choose a bandwidth that balances the explanatory power of the local linear model in equation (11) and its deviation from the global linear model to preserve the cyclic pattern in the remainder. Let $\hat{m}_{h,t}(x)$ denote the local estimator at point $x$ with smoothing parameter $h$. Here, the Gaussian kernel with bandwidth $h$ was used. The sums of squares of errors are

$$\mathrm{SSE}_1(h) = \sum_{t=1}^{n} \{ Y_t - \hat{m}_{h,t}(X_t) \}^2.$$

The sums of squares of deviations of the local estimator from the linear predictor $\hat{Y}_t$ of the global linear model are

$$\mathrm{SSE}_2(h) = \sum_{t=1}^{n} \{ \hat{Y}_t - \hat{m}_{h,t}(X_t) \}^2.$$

$\mathrm{SSE}_1$ is increasing in $h$, i.e. the smaller the bandwidth the better is the explanatory power of the local model — but the more the cyclicity in the remainder is reduced to white noise. In contrast, $\mathrm{SSE}_2$ is decreasing in $h$ since — for a larger neighbourhood size — the local model approaches the global linear model, preserving the cyclicity in the remainder. We chose the bandwidth at the intersection point between $\mathrm{SSE}_1(h)$ and $\mathrm{SSE}_2(h)$. One can perform a further fine tuning of the smoothing parameter, on the basis of a residual analysis at the stage of fitting the transmission equation (1). Fig. 6 illustrates the choice of bandwidth for the measles data for England and Wales at 0.3 times the standard deviation of $X_t$. Fig. 5(b) shows that the smooth reconstruction of the susceptible population produces a remainder that is locally stationary around zero.

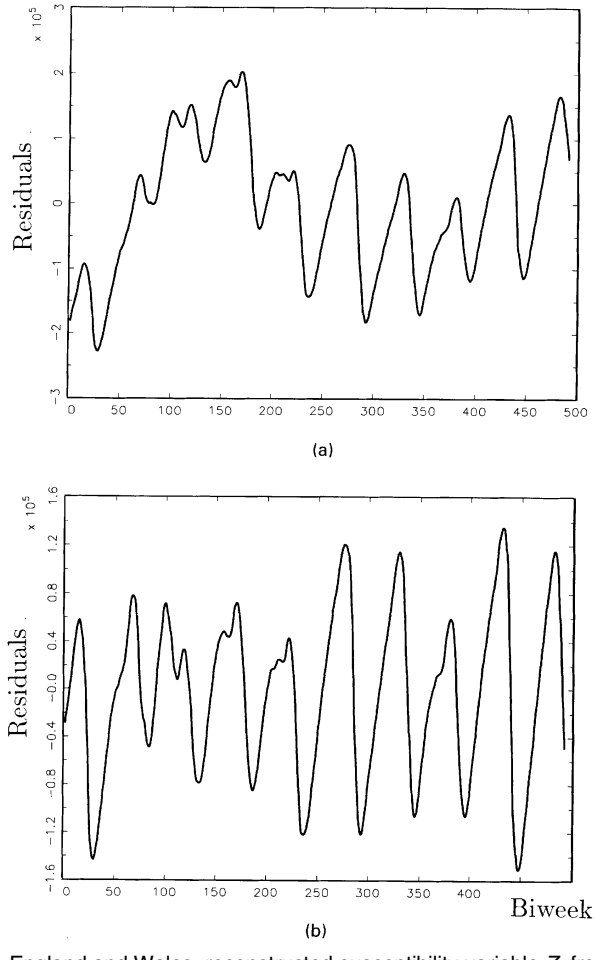


**Fig. 5.** Measles in England and Wales: reconstructed susceptibility variable $Z_t$ from (a) global linear regression and (b) local linear regression

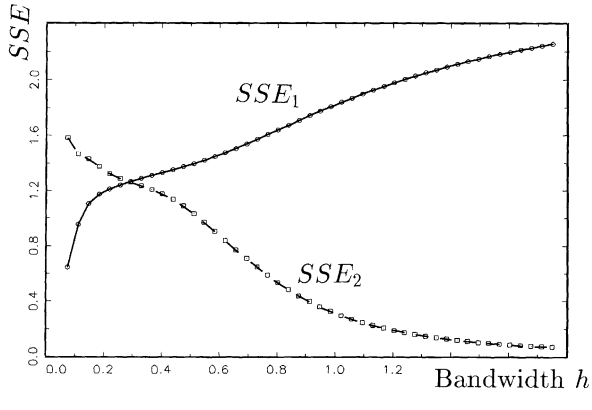


**Fig. 6.** Measles in England and Wales: bandwidth choice of local linear regression (the bandwidths are reported in units of standardized $X_t$)

### 3.2. Fitting the transmission equation

Consider the transmission equation (1). We first apply a logarithmic transformation to turn equation (1) into a linear additive model:

$$\ln(I_t) = \ln(r_t) + \alpha_1 \ln(I_{t-1}) + \alpha_2 \ln(S_{t-1}) + \ln(\epsilon_t). \tag{12}$$

The obvious benefit is the additive separability of the seasonal component, the interaction terms and the error. To measure the effect of the annual assembly and dispersal of children to and from schools (Fine and Clarkson, 1982), we allow for an annual seasonal variation in the transmission parameter $r_t$. We assume that the annual pattern is constant through time, reflecting a constant school–holiday pattern. This is achieved with a simple deterministic seasonality,

$$r_t = r_{t \bmod 26}. \tag{13}$$

This representation is much more flexible than the seasonal forcing in the standard seasonally forced SEIR model where, in the simplest approach, seasonal forcing is generally modelled by a periodic function

$$r_t = \beta_0\{1 + \beta_1 \cos(2\pi t)\}.$$

Here $\beta_0$ is the average transmission rate and $\beta_1$ denotes the amplitude of variation around $\beta_0$ (London and Yorke, 1973; Dietz, 1976). Furthermore, in Schenzle's (1984) RAS model, each epidemiological compartment exhibits an age-specific contact rate so $r_t$ is replaced by a contact matrix. Earn *et al.* (1999) showed that the behaviour of the aggregate RAS model over the different age classes can be reformulated by an appropriately forced SEIR model. However, the periodicity of the seasonal forcing in the aggregate RAS model is not yet fully understood. Our empirical approach postulates that it can be approximated by the most general representation of an annual periodicity that is compatible with biweekly data.

Next, to replace $S_t$ in the transmission equation by its reconstructed variation $Z_t$ around the mean, we expand $\ln(S_t)$ around $\bar{S}$:

$$\ln(S_t) = \ln(\bar{S} + Z_t) \approx \ln(\bar{S}) + \frac{1}{\bar{S}} Z_t + O\left(\frac{Z_t^2}{\bar{S}^2}\right). \tag{14}$$

Neglecting the expansion remainder and incorporating equations (13) and (14) into equation (12) gives

$$\ln(I_t) = \ln(r^*_{t \bmod 26}) + \alpha_1 \ln(I_{t-1}) + \zeta Z_{t-1} + \ln(\epsilon_t), \tag{15}$$

where $r_i^* = r_i \bar{S}^{\alpha_2}$ and $\zeta = \alpha_2/\bar{S}$. Equation (15) is a linear autoregressive process of order 1 in $\ln(I_t)$ with deterministic seasonality and additive covariate $Z_{t-1}$. The parameters ($r^*_{t \bmod 26}$, $\alpha_1$, $\zeta$) may hence be estimated by least squares. If $\ln(\epsilon_t)$ is normal then the least squares estimators are, in addition, maximum likelihood estimates. For estimation, we replace $I_t$ by the number of reported cases multiplied by the estimated local reporting rates, $\hat{\rho}_t C_t$, and $Z_t$ is the remainder of the local linear fit of equation (11) as shown in Fig. 5(b).

We checked that the first-order approximation works well for simulated data from the RAS model. The residuals of the above regression, using either $\ln(S_t)$ or the expansion with $Z_t$, were almost perfectly correlated. There was virtually no difference in the goodness of fit with both $R^2$-values being different from 1 in the third digit only. We also considered the second-order expansion of equation (14) which implies that $Z_t^2$ be included as a regressand (Table 1). However, we found that, for any real data applications, its coefficient is not — or is

**Table 1.** Estimation results for England amd Wales pre-vaccination data (1944–1964)

| | *Coefficient* | *t-ratio* | *Standard error* |
|---|---|---|---|
| $\hat{\alpha}_1$ | 0.9689 | 152.891 | 0.00634 |
| $\hat{\zeta}$ | $1.573 \times 10^{-6}$ | 17.680 | $8.896 \times 10^{-8}$ |
| (Order 2 expansion) | $-2.38 \times 10^{-12}$ | −2.040 | $1.17 \times 10^{-12}$ |
| $ln(\hat{r}^*_{t\,\mathrm{mod}26})$, *biweek* | | | |
| 1 | 0.155 | 2.514 | 0.062 |
| 2 | 0.571 | 9.320 | 0.061 |
| 3 | 0.460 | 7.260 | 0.063 |
| 4 | 0.341 | 5.261 | 0.065 |
| 5 | 0.302 | 4.612 | 0.065 |
| 6 | 0.343 | 5.231 | 0.065 |
| 7 | 0.245 | 3.735 | 0.066 |
| 8 | 0.154 | 2.370 | 0.065 |
| 9 | 0.309 | 4.886 | 0.063 |
| 10 | 0.404 | 6.397 | 0.063 |
| 11 | 0.323 | 5.094 | 0.063 |
| 12 | 0.238 | 3.765 | 0.063 |
| 13 | 0.202 | 3.228 | 0.063 |
| 14 | 0.203 | 3.284 | 0.062 |
| 15 | 0.074 | 1.223 | 0.061 |
| 16 | −0.095 | −1.606 | 0.059 |
| 17 | −0.218 | −3.819 | 0.057 |
| 18 | −0.031 | −0.571 | 0.054 |
| 19 | 0.433 | 8.224 | 0.053 |
| 20 | 0.531 | 9.928 | 0.054 |
| 21 | 0.479 | 8.705 | 0.055 |
| 22 | 0.397 | 7.046 | 0.056 |
| 23 | 0.444 | 7.731 | 0.057 |
| 24 | 0.411 | 7.027 | 0.058 |
| 25 | 0.291 | 4.886 | 0.060 |
| 26 | 0.509 | 8.498 | 0.060 |
| Standard deviation (residuals) | 0.1174 | | |
| $R^2$ | 0.9887 | | |
| $\frac{1}{n}\sum_{t=1}^{n}\hat{\rho}_t$ | 1.947 | | |

only borderline — significant and hence the corresponding term is dropped in the subsequent analysis.

## 4. Results for pre-vaccination England and Wales measles data

### *4.1. The model fit*

Table 1 summarizes the fit of the TSIR model to the biweekly data for England and Wales (Fig. 2). The average local slope for the reporting rate is 1.95 which means that about 51% of the cases were reported. This result is in line with Fine and Clarkson (1982), who stated that approximately two-thirds of measles infections were notified in England and Wales.

The coefficient $\hat{\alpha}_1 = 0.97$ is only slightly less than 1, indicating that the infective contact process is nearly homogeneous. However, the small difference from 1 is significant in terms of its $t$-value. Furthermore, simulations of the dynamical model with $\alpha_1$ set equal to 1 fail to reproduce the observed biennial cyclicity. The reconstructed susceptibility variable $Z_{t-1}$

clearly has explanatory power in the transmission equation. Not only is the coefficient $\zeta$ significantly different from 0 in terms of its $t$-value, but also separate regressions show that $Z_{t-1}$ explains the predominant biennial cycle *in conjunction* with $\ln(I_{t-1})$. If either of these variables is omitted from regression (15), the autocorrelogram of the residuals still reveals a strong biennial pattern. Furthermore, if we ignore the seasonality and assume that the transmission parameter in equation (15) is constant, then the regression remainder is subject to strong annual cycles whereas the biennial component vanishes.

This balance between annual and biennial epidemics is exactly consistent with previous work on the SEIR and RAS models (Grenfell, Bolker and Kleczkowski, 1995a). In particular, the predominant biennial cycle is a result of the non-linear contact process between susceptible and infected individuals. For constant transmission rate and births, the systematic part of the transmission process converges towards the equilibrium given by the number of births. The transient behaviour is then characterized by dampened biennial cycles and cycles of lower lengths if birth input is increased.

Fig. 7 plots the estimated annual pattern of the transmission parameters (mean centred), along with a school-term indicator function for the UK. The overall pattern of the seasonality is qualitatively similar to that estimated by Fine and Clarkson (1982). The transmission parameter attains its lowest values during the summer break in biweeks 15–18. The start of school around biweek 19 after the long summer break coincides with a pronounced rise in transmission. Other prominent seasonal dips in the transmission parameter coincide closely with the Easter (around biweek 8) and Christmas (around biweek 25) holidays; both are succeeded by steep rises with the start of the term. The effect of the Christmas holidays is captured with a lag, probably due to delays in notification over Christmas (Fine and Clarkson, 1982). However, the most drastic decline and subsequent rise of the transmission parameters coincides with the long summer school holiday. Since the observed cases are in

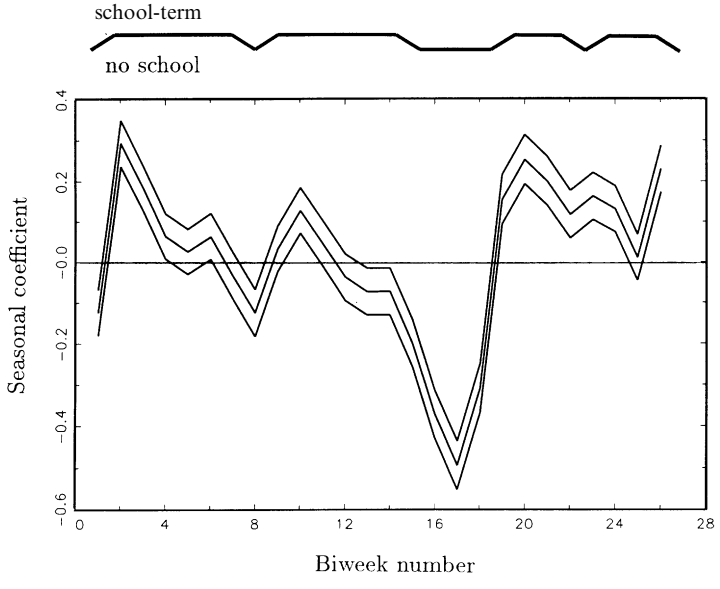


**Fig. 7.** Estimates of seasonal coefficients (mean centred) of measles in England and Wales within a confidence band of width two standard deviations: the uppermost graph shows the school-term indicator function (high level, school-term; low level, no school) for England and Wales

fact aggregates over different age classes, we are probably seeing here the seasonal pattern of the effective transmission parameter that encompasses age structure (Schenzle, 1984; Earn *et al.*, 1999). The estimated seasonal shape does not suggest a representation by means of a simpler function with few parameters, and in particular not a reduction towards a simple on–off school-term indicator function. The estimated transmission parameters are not merely measuring the rate of contact between school-children; otherwise they would be constant within school-terms. The fact that they are declining through the school-term (for instance from biweeks 10 to 14) after an initial rise at the beginning of the term underlines that schools are the main, but not the only, carriers of the mixing between infected and susceptible individuals (Schenzle, 1984; Anderson and May, 1991). Transmission rates are high when the epidemic initially spreads through schools where it quickly depletes a large number of susceptible children. After that, the effective transmission rate declines as children from isolated schools or in the preschooling age become infected. Keeling *et al.* (1997) arrived at similar conclusions from the study of correlation models for childhood epidemics. Schenzle (1984) also found a biennial variation in the aggregate transmission parameter of the RAS model. Our analysis does not identify this as a significant influence on the measles data for England and Wales.

Overall, the results of fitting the transmission equation with annual periodic forcing are encouraging, as this simple modelling approach manages to reduce substantially the remainder towards uncorrelated and symmetric noise. The $R^2$-value of 0.9887 is convincingly high. This is mainly due to the large explanatory power of the previous number of infected individuals. Fig. 8(a) shows that the one-step-ahead predictions from the model almost overlap the data.

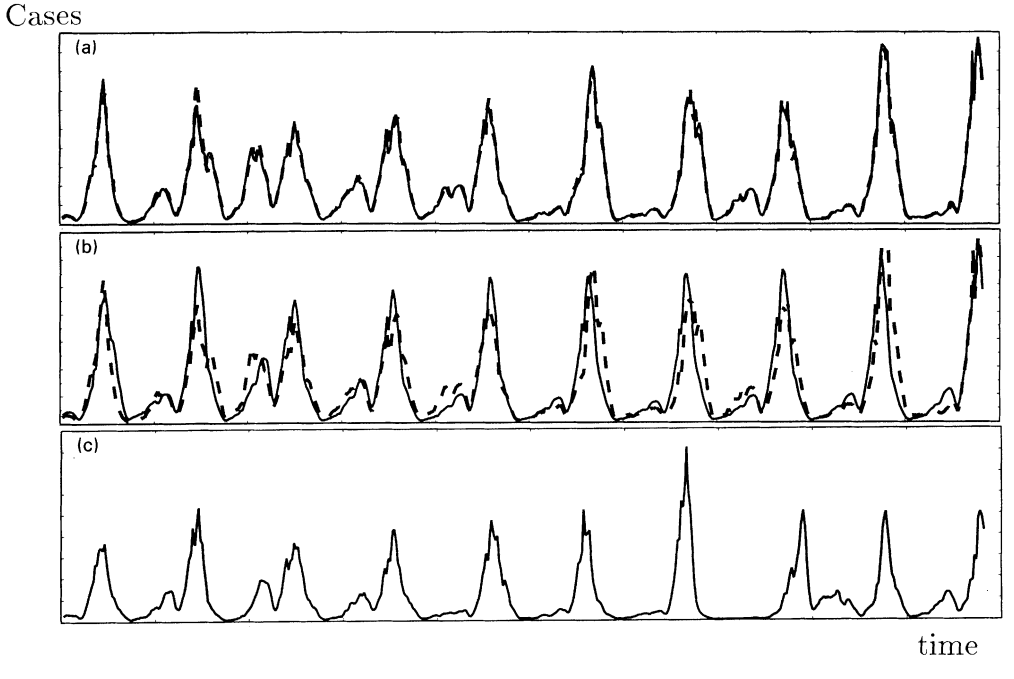


**Fig. 8.** Pre-vaccination for measles in England and Wales: (a) model fit (– – –, observed; ———, model predictions); (b) observed series (– – –) and iterated deterministic skeleton with observed births input for the same initial condition (———); (c) realization of the stochastic model with observed births input

### 4.2. Simulating long-term dynamics

A much more substantial test of the TSIR approach is to assess whether the dynamics of the estimated model are compatible with the data. In particular, we assess how the dynamic behaviour of the model reacts to changes in the birth-rate. The following two-dimensional system simulates the behaviour of the estimated model: for infected individuals,

$$\ln(I_t) = \hat{r}^*_{t\,\mathrm{mod}26} + \hat{\alpha}_1 \ln(I_{t-1}) + \hat{\zeta} Z_{t-1} + \epsilon_t; \tag{16}$$

for susceptible individuals,

$$Z_t = B_{t-d} + Z_{t-1} - I_t. \tag{17}$$

We use the term *stochastic model* if the dynamical noise in equation (16) is uncorrelated Gaussian noise and refer to the model with $\epsilon_t = 0$ as the *deterministic skeleton* (Tong, 1990). The standard deviation of the regression residuals does not only reflect dynamical noise. It also incorporates errors arising from other possible sources of uncertainty, such as measurement noise and errors due to the fact that the model may be too simplistic biologically. For illustration, we set the standard deviation of the Gaussian dynamic noise at 80% of the estimated standard deviation. This probably still overestimates the importance of dynamical noise.

We use graphical tools (bifurcation diagrams and phase plots) to analyse the non-linear behaviour of simulated time series from the model (May, 1976). As starting values, we used the observed pair $(I_1, Z_1)$.

Figs 8–10 display the simulation results. Fig. 8(b) depicts the iterated deterministic skeleton. Considering that the 20-year-ahead prediction is only conditioned on the initial values $I_1$, $Z_1$ and the observed births, it traces the observed data remarkably well. In particular, the mechanism captures the biennial cycle as well as the higher minor epidemics during the time of high birth-rates around 1947. Fig. 8(c) illustrates one realization of the stochastic model. The fact that the deterministic skeleton fits the data better than the stochastic realizations indicates that we may be overestimating the level of the dynamical noise.

Fig. 9 is a more systematic analysis of the role of the birth-rate as a bifurcation parameter. Starting from initial conditions $I_1 = B$ and $Z_1 = 0$, the model is iterated assuming a time constant birth-rate, as plotted on the horizontal axis. The observed range of births for England and Wales is marked along the axis. We simulate the model for 100 years, after a transient time of 100 years. The time path of $\log(I_t)$ is measured at biweek 10 of each year and is plotted against the input births. We chose biweek 10 as roughly the peak of the major epidemic years. For a given birth-rate, the deterministic time path has an annual (or biennial) cycle if the sequence of 100 measurement points is plotted onto one (or two) points.

The resulting bifurcation diagram (Fig. 9) illustrates that the deterministic skeleton

(a) converges towards annual cycles for very low birth-rates,
(b) bifurcates towards 2-year cycles with an increasing difference in amplitude between the major and minor epidemic years for increasing birth-rates and
(c) finally collapses back to annual cycles at high birth-rates.

The observed birth range spans the behaviour around the latter bifurcation. The stochastic realizations scatter around the deterministic bifurcation and little information is revealed about their cyclicity.

The effect of even small amounts of dynamical noise may be substantial in a non-linear

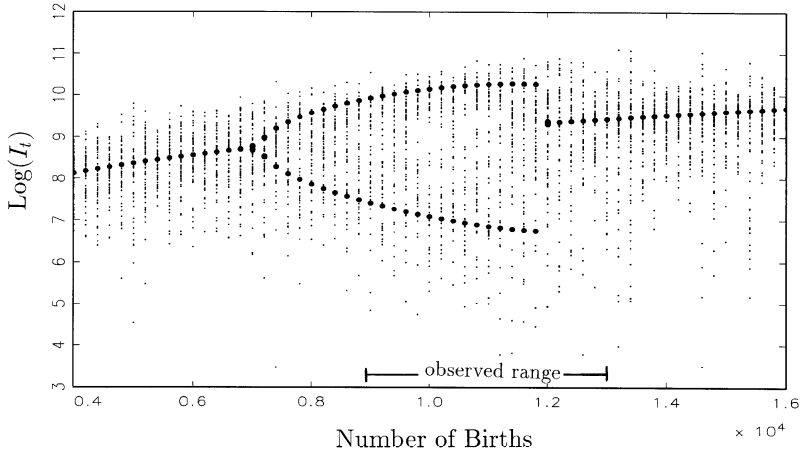


**Fig. 9.** Bifurcation diagram: number of births against log($I_t$) at biweek 10 of each year from simulations (100 years) of the deterministic skeleton (•) and the stochastic model with dynamical noise (·) (the transient time was 100 years)

system (Rand and Wilson, 1991). For instance, the stochastic realizations may exhibit intermittent biennial cyclicity although the corresponding deterministic attractor is annual (Grenfell, Bolker and Kleczkowski, 1995b). We explore the question of *predominant* cyclicity in the presence of dynamical noise by means of methods based on joint density estimation (see Finkenstädt *et al.* (1998)). Fig. 10(a) shows the median biennial waves arising from 384 years of stochastic simulations for two levels of the birth-rate denoted by *high* and *low*. Fig. 10(b) shows a delay plot of the overall number of cases in epidemic year $t$ against year $t-1$. The graph illustrates the negative relationship between successive epidemics along a line with slope $-1$. For a higher birth-rate this is shifted outwards as more cases occur. The predominant cyclicity of the stochastic realizations can be inferred from the joint density in Fig. 10(c):

(a) for the lower birth-rate, high epidemic outbreaks are generally followed by low epidemics and vice versa, resulting in a bimodal joint density; hence, the most frequently encountered cyclicity is biennial;
(b) the higher birth-rate gives rise to successive epidemics of similar magnitude. The corresponding joint density is unimodal and the predominant cyclicity is hence annual.

Here, the bifurcation of the joint density from unimodal towards bimodal is the stochastic equivalent to the deterministic bifurcation scenario, with birth-rate as the bifurcation parameter. It seems that the qualitative dynamic behaviour is essentially preserved, even if the dynamical noise input is as high as 80% of the estimated overall noise. The results in Fig. 10 also show that the TSIR model can capture the negative density dependence between successive epidemics that is also seen in data at the individual city level (Finkenstädt *et al.*, 1998).

## 5. Discussion

This study has attempted to bridge the gap between non-linear dynamical models and stochastic time series in epidemics. The dynamics of childhood diseases, and in particular

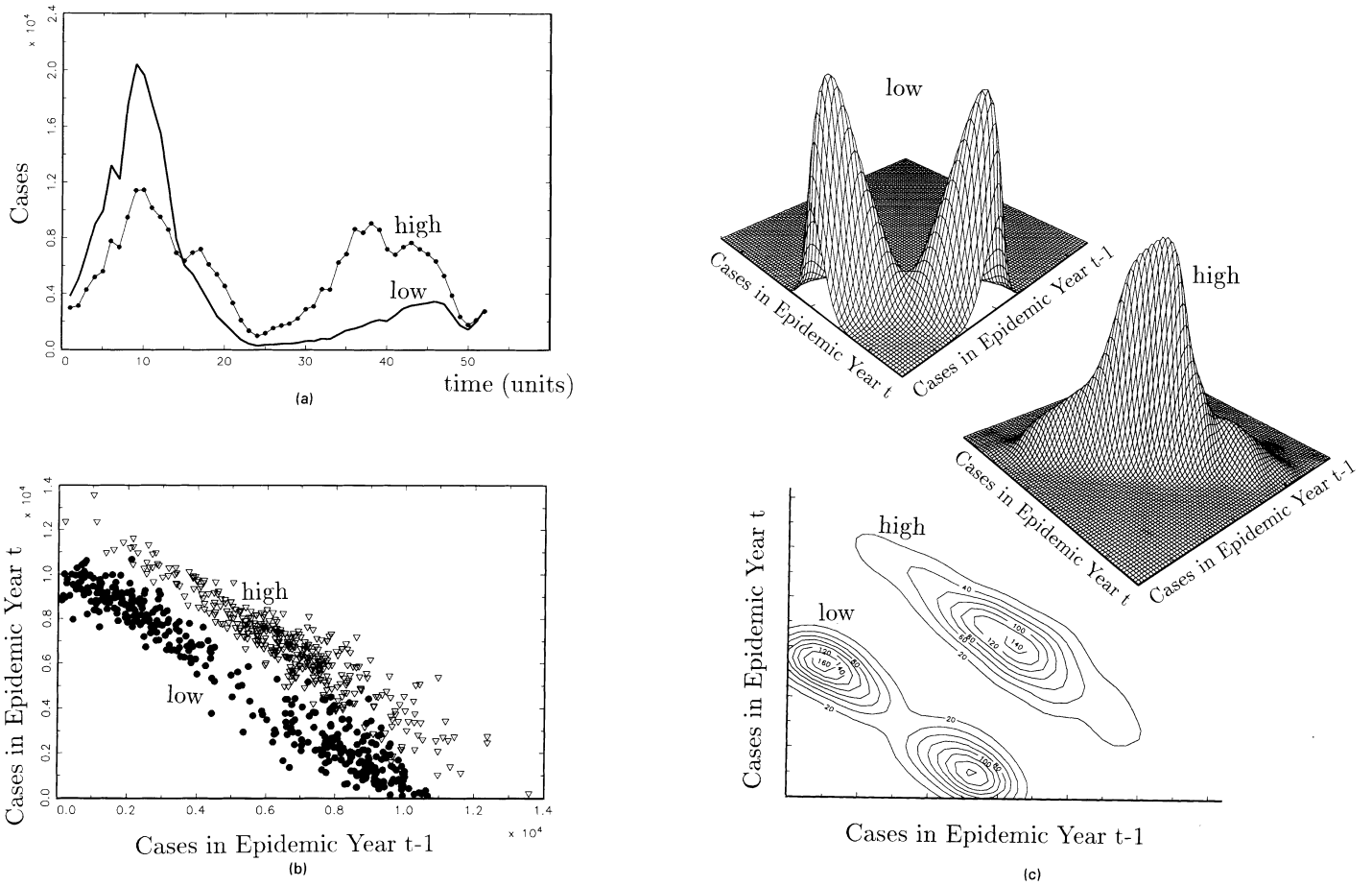


**Fig. 10.** Stochastic realization (384 years) of the model with dynamical noise for two different levels of births input, 11000 (low) and 14000 (high): (a) median biennial waves (———, low births; •, high births); (b) delay plot of cases in successive epidemic years (•, low births; ▽, high births) formed by summing over the cases of epidemic years of biweeks 1–26 and 27–52 as shown in (a); (c) kernel joint densities (two-dimensional product Gaussian kernel, bandwidth 0.7 × normal referencing rule) of the scatterpoints in delay plot (b) (the bimodal joint density corresponds to low births and the unimodal joint density corresponds to high births)

measles, served as our case-study. We developed a discrete time stochastic embellishment of the SEIR model related to early work by Fine and Clarkson (1982) and demonstrated methods of fitting the model to data on birth-rates and case reports. The TSIR model fits the data remarkably well. Moreover, it complies with the substantial requirement, that its dynamical behaviour explains the observed cyclicity. We estimated the seasonal pattern of transmission rates and studied the role of the birth-rate as a bifurcation parameter. Our results suggest the following:

(a) at the aggregate population level, an annual seasonal forcing encompasses heterogeneities due to age structure; the estimated seasonal forcing shows a more complicated pattern than the simple sinusoidal or term-time seasonal forcing;
(b) the birth-rate drives bifurcations in the dynamic behaviour of the system.

Since the TSIR model incorporates a mechanism dictated by epidemic theory, it has potential for studying a variety of further epidemiological questions. We address three of them here.

Firstly, it is important to understand and predict the effect of vaccination programmes. Vaccination modulates the recruitment rate of the susceptible individuals (equation (2)) in effectively the same way as births. The TSIR model's ability to estimate the effect of birth-rate variations also gives us the potential to examine the much more important question of childhood vaccination policies for developing countries.

Secondly, the natural history of measles infection fits the SIR paradigm very closely. It is therefore not surprising that the rigorously parameterized TSIR model performs well. It will be interesting to extend this modelling approach to explore the dynamics of other major childhood microparasites with a more complex natural history, such as whooping-cough and rubella. Fitting the TSIR model to incidence data for these infections might give us a more focused way of exploring whether their natural history fits the SIR paradigm. An interesting area for future statistical work will be comparing the performance of the TSIR model and semimechanistic approaches (Ellner *et al.*, 1998), especially for more biologically complex infections than measles. We are also exploring the use of serological data to estimate the mean susceptible density $\bar{S}$ directly. This will allow us to estimate potential saturation in the contact process, via the parameter $\alpha_2$.

Thirdly, this study has focused on the *endemic dynamics* of the disease observed in a large human host population. Measles dynamics at the individual city level and, in particular, in small communities are much more irregular. For populations below a critical community size of about 250000 inhabitants (Bartlett, 1957, 1960) the numbers of susceptible individuals become depleted to the extent that the chain of transmission is interrupted. Infection generally disappears after major epidemics so we need to allow in equation (1) for the influx of infections from large centres. This necessitates discrete stochastic modelling of *epidemic* dynamics (Finkenstädt *et al.*, 2000) where the disease goes locally extinct and only recurs as a result of spatial migration of infected individuals. The pattern of spatial coupling is thus of great importance here.

## Acknowledgements

This research was supported by the Wellcome Trust (BF and BG). Thanks are due to Matt Keeling, who generated the RAS model data, and to Howell Tong, Qiwei Yao, Peter Kuhbier, Ottar Bjørnstad and David Rand for their constructive comments. Thanks are also due to the referees for suggestions that have improved the presentation.

## References


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

Bailey, N. T. J. (1975) *The Mathematical Theory of Infectious Diseases and Its Applications*. London: Griffin.

Bartlett, M. S. (1957) Measles periodicity and community size. *J. R. Statist. Soc.* A, **120**, 48–60.

———(1960) The critical community size for measles in the United States. *J. R. Statist. Soc.* A, **123**, 37–44.

Black, F. L. (1984) Measles. In *Viral Infections of Humans: Epidemiology and Control* (ed. A. S. Evans), pp. 397–418. New York: Plenum.

Bolker, B. M. and Grenfell, B. T. (1996) Impact of vaccination on the spatial correlation and dynamics of measles epidemics. *Proc. Natn. Acad. Sci. USA*, **93**, 12648–12653.

Casdagli, M. (1989) Nonlinear prediction of chaotic time series. *Physica* D, **35**, 335–356.

Dietz, K. (1976) The incidence of infectious diseases under the influence of seasonal fluctuations. *Lect. Notes Biomath.*, **11**, 1–15.

Earn, D. J., Rohani, P., Bolker, B. M. and Grenfell, B. T. (1999) A simple model for complex dynamical transitions in epidemics. *Science*, to be published.

Ellner, S. P. (1991) Detecting low-dimensional chaos in population dynamics data: a critical review. In *Chaos and Insect Ecology* (eds J. Logan and F. Hain), pp. 63–90. Charlottesville: University Press of Virginia.

Ellner, S. P., Bailey, B. A., Bobashev, G. V., Gallant, A. R., Grenfell, B. T. and Nychka, D. W. (1998) Noise and nonlinearity in measles epidemics: combining mechanistic and statistical approaches to population modeling. *Am. Nat.*, **151**, 425–440.

Ellner, S. P. and Turchin, P. (1995) Chaos in a noisy world: new methods and evidence from time series analysis. *Am. Nat.*, **145**, 343–375.

Fan, J. and Gijbels, I. (1996) *Local Polynomial Modelling and Its Applications*. London: Chapman and Hall.

Fine, P. E. M. and Clarkson, J. A. (1982) Measles in England and Wales: I, an analysis of factors underlying seasonal patterns. *Int. J. Epidem.*, **11**, 5–14.

Finkenstädt, B. F., Bjørnstad, O. N. and Grenfell, B. T. (2000) A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks. *Preprint*. Department of Zoology, University of Cambridge, Cambridge.

Finkenstädt, B. F. and Grenfell, B. T. (1998) Empirical determinants of measles metapopulation dynamics in England and Wales. *Proc. R. Soc. Lond.* B, **265**, 211–220.

Finkenstädt, B. F., Keeling, M. J. and Grenfell, B. T. (1998) Patterns of density dependence in measles dynamics. *Proc. R. Soc. Lond.* B, **265**, 753–762.

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

Grenfell, B. T., Bolker, B. M. and Kleczkowski, A. (1995a) Demography, seasonality and the dynamics of measles in developed countries. In *Epidemic Models: Their Structure and Relation to Data* (ed. D. Mollison), pp. 248–268. Cambridge: Cambridge University Press.

———(1995b) Seasonality and extinction in chaotic metapopulations. *Proc. R. Soc. Lond.* B, **259**, 97–103.

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

Grenfell, B. T., Kleczkowski, A., Ellner, S. P. and Bolker, B. M. (1994a) Non-linear forecasting and chaos in ecology and epidemiology: measles as a case study. In *Chaos and Forecasting* (ed. H. Tong), pp. 321–345. Singapore: World Scientific.

———(1994b) Measles as a case-study in nonlinear forecasting and chaos. *Phil. Trans. R. Soc. Lond.* A, **348**, 515–530.

Grenfell, B. T., Kleczkowski, A., Gilligan, C. A. and Bolker, B. M. (1995) Spatial heterogeneity, nonlinear dynamics and chaos in infectious diseases. *Statist. Meth. Med. Res.*, **4**, 160–183.

Heesterbeek, J. A. P. and Roberts, M. G. (1995) Mathematical models for microparasites of wildlife. In *Ecology of Infectious Diseases in Natural Populations* (eds B. T. Grenfell and A. P. Dobson), pp. 90–122. Cambridge: Cambridge University Press.

de Jong, M., Diekmann, O. and Heesterbeek, H. (1995) How does transmission of infection depend on population size? In *Epidemic Models: Their Structure and Relation to Data* (ed. D. Mollison), pp. 84–94. Cambridge: Cambridge University Press.

Keeling, M. J. and Grenfell, B. T. (1997) Disease extinction and community size: modeling the persistence of measles. *Science*, **275**, 65–67.

Keeling, M. J., Rand, D. A. and Morris, A. J. (1997) Correlation models for childhood epidemics. *Proc. R. Soc. Lond.* B, **264**, 1149–1156.

Kermack, W. O. and McKendrick, A. G. (1933) A contribution to the mathematical theory of epidemics: Part III, Further studies of the problem of endemicity. *Proc. R. Soc. Lond.* A, **141**, 92–122.

London, W. P. and Yorke, J. A. (1973) Recurrent outbreaks of measles, chickenpox and mumps: I, Seasonal variation in contact rates. *Am. J. Epidem.*, **98**, 453–468.

Lui, W. M., Hethcote, H. W. and Levin, S. A. (1987) Dynamical behaviour of epidemiological models with nonlinear incidence rates. *J. Math. Biol.*, **25**, 359–380.

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

———(1986) Population biology of microparasitic infections. *Biomath*., **17**, 405–442.
McLean, A. R. and Anderson, R. M. (1988a) Measles in developing countries: Part I, Epidemiological parameters and patterns. *Epidem. Infectn*, **100**, 111–133.
———(1988b) Measles in developing countries: Part II, The predicted impact of mass vaccination. *Epidem. Infectn*, **100**, 419–442.
Miller, E. (1996) Immunisation policies — successes, failure and the future. *J. Clin. Pathol*., **49**, 620–622.
Mollison, D. and Din, S. U. (1993) Deterministic and stochastic models for the seasonal variability of measles transmission. *Math. Biosci*., **117**, 144–177.
Morse, D., Oshea, M., Hamilton, G., Soltanpoor, N., Leece, G., Miller, E. and Brown, D. (1994) Outbreak of measles in a teenage school population — the need to immunize susceptible adolescents. *Epidem. Infectn*, **113**, 355–365.
Nychka, D., Ellner, S., Gallant, A. R. and McCaffrey, D. (1992) Finding chaos in noisy systems. *J. R. Statist. Soc*. B, **54**, 399–426.
Olsen, L. F. and Schaffer, W. M. (1990) Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science*, **249**, 499–504.
Rand, D. A. and Wilson, H. B. (1991) Chaotic stochasticity: a ubiquitous source of unpredictability in epidemics. *Proc. R. Soc. Lond*. B, **246**, 179–184.
Schaffer, W. M. and Kot, M. (1985) Nearly one dimensional dynamics in an epidemic. *J. Theoret. Biol*., **112**, 403–427.
Schenzle, D. (1984) An age-structured model of pre- and post-vaccination measles transmission. *IMA J. Math. Appl. Med. Biol*., **1**, 169–191.
Sugihara, G., Grenfell, B. T. and May, R. M. (1990) Distinguishing error from chaos in ecological time series. *Phil. Trans. R. Soc. Lond. Biol*., **330**, 235–251.
Takens, F. (1981) Detecting strange attractors in turbulence. *Lect. Notes Math*., **898**, 366–381.
Tidd, C. W., Olsen, L. F. and Schaffer, W. M. (1993) The case for chaos in childhood epidemics: II, Predicting historical epidemics from mathematical models. *Proc. R. Soc. Lond*. B, **254**, 257–273.
Tong, H. (1990) *Non-linear Time Series*. Oxford: Oxford University Press.