

Estimating variability in models for recurrent epidemics: assessing the use of moment closure techniques

Alun L. Lloyd^{a,b,*}

^a Program in Theoretical Biology, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA

^b Biomathematics Graduate Program and Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA

Received 10 December 2002

Abstract

The major role played by demographic stochasticity in determining the dynamics and persistence of childhood diseases, such as measles, chickenpox and pertussis, has long been realized. Techniques which can be used to estimate the magnitude of this stochastic effect are of clear importance. In this study, we assess and compare the use of two moment closure approximations to estimate the variability seen about the average behavior of stochastic models for the recurrent epidemics seen in childhood diseases. The performance of the approximations are assessed using analytic techniques available for the simplest epidemiological model and using numerical simulations in more complex settings. We also present epidemiologically important extensions of previous work, considering variability in the SEIR model and in situations for which there is seasonal variation in disease transmission. Important implications of stochastic effects for the dynamics of childhood diseases are highlighted, including serious deficiencies of deterministic descriptions of dynamical behavior.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Demographic stochasticity; Moment closure; Epidemic models; Nonlinear dynamics

1. Introduction

Although deterministic models have contributed much to our understanding of the biological processes which underlie the spread of disease (see Dietz and Schenzle, 1985; Anderson and May, 1991; Diekmann and Heesterbeek, 2000, for reviews of the field), the importance of random effects in determining population-level patterns of disease incidence and persistence has long been realized (see, for instance, Bartlett, 1956, 1957, 1960a, b; Black, 1966; Anderson and May, 1991; Grenfell et al., 1995a; Bolker and Grenfell, 1995, 1996; Keeling and Grenfell, 1997; Nåsell, 1999; Andersson and Britton, 2000; Keeling et al., 2001; Rohani et al., 2002). The epidemiological impact of stochasticity has received most attention within the context of childhood diseases, most notably measles, on which this study is focused.

Stochasticity leads to two fundamental differences between the behavior of deterministic and stochastic models. Firstly, random effects give rise to variability in the course of epidemics (Isham, 1991): repeated simulation of a stochastic process starting with identical initial conditions leads to a collection of different realizations of the process, in contrast to similar repeated simulation of a deterministic model. Secondly, when disease incidence is low, the chain of disease transmission can be interrupted, leading to (at least local) extinction, or fadeout, of the disease (Bartlett 1956, 1957, 1960a, b). Fadeout is most likely to occur in the period immediately following a major epidemic (the so-called ‘inter-epidemic trough’). An important observation is that deterministic models cannot reproduce such effects: the ability of populations to recover from very low levels is a well-known weakness of many deterministic models (Engbert and Drepper, 1994; Bolker and Grenfell, 1995).

Knowledge of the variability between epidemic realizations is of direct interest when making predictions about the future course of an epidemic as it gives

*Corresponding address: Biomathematics Graduate Program and Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA.

E-mail addresses: alloyd@stat.ncsu.edu, alun@alunlloyd.com.

information about the likely accuracy of any forecast made. Variability is also informative regarding disease persistence: extinction is more likely to occur when high levels of variability are observed because individual realizations wander further away from their mean and so are more likely to come close to zero.

Estimation of variability by examination of a collection of model realizations is straightforward, although this can be quite expensive in terms of computer time, a problem exacerbated when large population sizes are considered. An alternative way to understand the variability between realizations involves the derivation of differential equations for the moments (the mean, standard deviation, skewness, and so on) of the distribution of the states of the system (Whittle 1957; Isham, 1991). The difficulty in applying this approach, in general, is that nonlinearities in the equations governing the behavior of the system lead to coupling between the equations for moments of different orders: for instance, the equations for the first-order moments may involve second-order moments, those describing the second-order moments may involve third-order moments, and so on.

Use of moment equations requires the deployment of an approximation technique—a moment closure approximation—which can truncate this set of equations at some order. Analogous moment closure techniques are also used in other contexts, such as spatial systems, when a finite set of dynamical equations is desired to describe an infinite (or very high) dimensional system (Bolker and Pacala, 1997; Keeling and Grenfell, 1997).

The simplest moment closure methods assume that the distribution of states follows some given distribution and then uses the known relationship between the moments of that distribution to truncate the set of moment equations. For instance, Whittle's multivariate normal (MVN) approximation (Whittle, 1957) assumes that the distribution is multivariate normal. As the third-order central moments of a MVN vanish, closure of the set of moment equations is achieved by setting third-order central moments to zero. The MVN has previously been used within an epidemiological context by Isham (1991), although the particular setting concerned an epidemic, rather than endemic, situation. A particularly fruitful application of related moment closure methods within the epidemiological literature has been within the context of macroparasite infections (see, for instance, Grenfell et al., 1995b).

Although the MVN approximation has been the one most widely used in the literature, and as such will be our main interest here, we shall see below that the resulting moment equations exhibit what appear to be important failings when population sizes are 'small'. In an attempt to provide an improved approximation in such situations, an alternative moment closure technique, which assumes that the distribution of states

follows a multivariate lognormal distribution, has been suggested (Keeling, 2000a, b).

We assess the use of these two moment closure approximations to estimate the impact of demographic stochasticity in models for recurrent epidemics in reasonably large populations. In the simplest case of a non-seasonal SIR model, the performance of the approximations can be compared against existing analytic expressions for the variability seen about the endemic equilibrium of the model. In more complex, but epidemiologically more realistic, model settings—for which there has been little previous systematic consideration of issues concerning variability—the behavior of the approximations is compared with results obtained from numerical simulation of the stochastic model. We explore in some detail the impact of including an exposed, but not yet infectious, class of individuals (examining the so-called SEIR model) and of the inclusion of seasonal variations in the transmission parameter.

This paper is organized as follows: Section 2 briefly outlines the biological background of the models under consideration and develops deterministic and stochastic formulations of the models. Section 3 discusses the derivation of moment equations to describe variability in stochastic formulations of the models, focusing on deployment of the multivariate normal approximation. Section 4 examines the performance of the moment equations in the well-known case of the non-seasonal SIR model. Section 5 considers the behavior of seasonally forced models, and highlights the increased importance of demographic stochasticity in such models. Appendix A briefly summarizes the development of moment equations based on the lognormal assumption. Appendix B extends the MVN moment equations to describe the SEIR model, and illustrates how the inclusion of an exposed class of individuals reduces variability in the model. Appendix C discusses dynamical differences between the behaviors of the deterministic equations and the MVN moment equations in the seasonally forced model.

2. Deterministic and stochastic model formulations

The SIR model and its variants have been central in the mathematical study of epidemics (Kermack and McKendrick, 1927; Bartlett, 1960b; Anderson and May, 1991). The population is divided into three classes: susceptible, infectious and recovered, and the numbers of individuals in these classes are written as S , I and R , respectively. Making the standard assumptions for an SIR model describing childhood diseases—in which it is assumed that the infection is non-fatal and confers permanent immunity upon recovery—in a well-mixed population of constant size, N (see, for instance,

Table 1

Transitions between classes in the SIR model. Here, the per-capita birth and death rates are written as μ , hence the average lifespan of individuals, L , equals $1/\mu$. The transmission parameter, β , equals α/N . (We absorb the N -dependence of the transmission parameter into β in this way as it simplifies notation somewhat.) The average duration of infectiousness, D , is $1/\gamma$

Event	Transition	Rate at which event occurs	Probability of transition in time interval $[t, t + dt]$
Birth	$S \rightarrow S + 1$	μN	$\mu N dt$
Susceptible death	$S \rightarrow S - 1$	μS	$\mu S dt$
Infection	$S \rightarrow S - 1, I \rightarrow I + 1$	βSI	$\beta SI dt$
Recovery	$I \rightarrow I - 1$	γI	$\gamma I dt$
Infectious death	$I \rightarrow I - 1$	μI	$\mu I dt$

Anderson and May, 1991; Diekmann and Heesterbeek, 2000), the movements of individuals between the classes are governed by the few simple rules, and accompanying parameters, listed in Table 1.

The deterministic SIR model, which treats the numbers of susceptibles, infecteds and recovered as continuously varying quantities, can then be written as

$$\frac{dS}{dt} = \mu N - \mu S - \beta SI, \tag{1}$$

$$\frac{dI}{dt} = \beta SI - \mu I - \gamma I. \tag{2}$$

Notice that an equation for the recovered class does not appear in the above set of differential equations as the assumption of a constant population size (or, equivalently, the assumed equality of birth and death rates) implies that $R = N - S - I$. It is also easy to see (by, for instance, rewriting Eqs. (1)–(2) in terms of the fractions of the population which are in each of the three classes, and noting the N -dependence of the parameter β discussed in Table 1) that the behavior of the deterministic model does not depend on the population size.

An alternative, stochastic, formulation implicitly recognizes that the population is made up of individuals, and that transitions between classes are random events (Bartlett, 1956, 1960b; Olsen et al., 1988). The rates at which various transitions occur can be reinterpreted to calculate the probabilities of each event occurring in an infinitesimal time interval dt . For instance, the probability of a birth occurring in the time interval $[t, t + dt]$ is $\mu N dt$. Realizations of this stochastic model can be generated by computer using standard Monte-Carlo techniques (Bartlett, 1957, 1960b; Renshaw, 1991).

As mentioned above, the stochastic formulation can exhibit disease fadeout: the number of infectives can fall to zero, as happens if there is a single infected individual who recovers or dies before passing on the disease. Further disease outbreaks cannot occur without the reintroduction of the infection by immigration of an infective individual from elsewhere. Examination of disease incidence records shows that disease fadeout followed by reintroduction is a frequent occurrence in

the real world (Bartlett, 1956, 1957, 1960a, b): models which are used to match actual disease patterns must take immigration into account. The introduction of immigration, however, has non-trivial implications for the dynamical patterns of disease incidence. If fadeouts are very common, then the timing and size of epidemics depends heavily on the immigration term, as witnessed by the sensitive dependence of dynamics to the level of immigration in the deterministic model (Engbert and Drepper, 1994; Bolker and Grenfell, 1995). As a consequence, we do not include immigration here, deferring discussion of its effects to a future study.

3. Derivation of moment equations

If the probability that the number of susceptibles and infecteds equals (s, i) at time t is written $p_t(s, i)$, then the Kolmogorov forward equations (Bartlett 1960b; Feller, 1968) (also known as the master equations) of the system can be obtained in the standard way: relate $p_{t+dt}(s, i)$ to p_t by considering the possible transitions which could occur in the time interval $[t, t + dt]$ and let dt tend to zero, giving

$$\begin{aligned} \frac{dp_t(s, i)}{dt} = & \mu N p_t(s - 1, i) + \mu(s + 1)p_t(s + 1, i) \\ & + \beta(s + 1)(i - 1)p_t(s + 1, i - 1) \\ & + (\gamma + \mu)(i + 1)p_t(s, i + 1) \\ & - \{\mu N + \mu s + \beta si + (\gamma + \mu)i\}p_t(s, i). \end{aligned} \tag{3}$$

(Notice that, since s and i must be non-negative, $p_t(s, i)$ is defined to be zero if either s or i is negative.)

For small population sizes, this set of coupled differential equations can be integrated numerically to study the evolution of the probability distribution of S and I (Jacquez and Simon, 1993), but this quickly becomes impractical as the population size increases.

The system of forward Eqs. (3) can be used to derive formulae for the rates of change of the expected numbers of susceptibles and infectives, $E(S)$ and $E(I)$, and higher moments such as the variances of the number of susceptibles and infectives and the covariance between the susceptible and infective numbers. The

most convenient approach for our purposes involves the use of the moment generating function (MGF), $M(\theta_1, \theta_2; t)$, defined as

$$M(\theta, t) = E(e^{\theta_1 S + \theta_2 I}). \quad (4)$$

The equation for the time evolution of the MGF is derived by multiplying Eq. (3) by $\exp(\theta_1 s + \theta_2 i)$ and summing over s and i , giving

$$\begin{aligned} \frac{\partial M}{\partial t} &= \beta(e^{\theta_2 - \theta_1} - 1) \frac{\partial^2 M}{\partial \theta_1 \partial \theta_2} + \mu(e^{-\theta_1} - 1) \frac{\partial M}{\partial \theta_1} \\ &+ (\gamma + \mu)(e^{-\theta_2} - 1) \frac{\partial M}{\partial \theta_2} \\ &+ \mu N(e^{\theta_1} - 1)M. \end{aligned} \quad (5)$$

Since the moment generating function can be written as

$$M(\theta_1, \theta_2) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} \theta_1^j \theta_2^{k-j} E(S^j I^{k-j}), \quad (6)$$

the expectation of $S^m I^n$ can be determined from the appropriate coefficient of $\theta_1^m \theta_2^n$ in the expansion of M . Hence, by expanding Eq. (5) in powers of θ_1 and θ_2 and equating coefficients, differential equations for the time evolution of the ordinary moments can be derived (Lloyd, 1996; Matis and Kiffe, 1999).

In order to derive equations for the variances and covariances, it is more convenient to consider the cumulant generating function, $K(\theta_1, \theta_2, t)$, defined as the logarithm of the MGF, $K(\theta_1, \theta_2, t) = \log(M(\theta_1, \theta_2, t))$, whose coefficients in the series expansion corresponding to (6) are known as the cumulants. The first two cumulants equal the mean and variance and the third cumulant is proportional to the skewness.

Eq. (5) can easily be transformed into an equation for the time derivative of K , giving

$$\begin{aligned} \frac{\partial K}{\partial t} &= \beta(e^{\theta_2 - \theta_1} - 1) \left(\frac{\partial^2 K}{\partial \theta_1 \partial \theta_2} + \frac{\partial K}{\partial \theta_1} \frac{\partial K}{\partial \theta_2} \right) \\ &+ \mu(e^{-\theta_1} - 1) \frac{\partial K}{\partial \theta_1} \\ &+ (\gamma + \mu)(e^{-\theta_2} - 1) \frac{\partial K}{\partial \theta_2} + \mu N(e^{\theta_1} - 1). \end{aligned} \quad (7)$$

Expansion of Eq. (7) gives the following equations for the time evolution of the moments of orders one and two:

$$\frac{dE(S)}{dt} = \mu N - \mu E(S) - \beta E(SI), \quad (8)$$

$$\frac{dE(I)}{dt} = \beta E(SI) - (\gamma + \mu)E(I), \quad (9)$$

$$\begin{aligned} \frac{d \text{Var}(S)}{dt} &= \mu N + \mu \{E(S) - 2\text{Var}(S)\} + \beta \{E(SI) \\ &- 2E(I)\text{Var}(S) \\ &- 2E(S)\text{Cov}(S, I) - 2T_{SSI}\}, \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{d \text{Var}(I)}{dt} &= (\gamma + \mu) \{E(I) - 2\text{Var}(I)\} + \beta \{E(SI) \\ &+ 2E(S)\text{Var}(I) \\ &+ 2E(I)\text{Cov}(S, I) + 2T_{SII}\}, \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{d \text{Cov}(S, I)}{dt} &= -(\gamma + 2\mu)\text{Cov}(S, I) - \beta \{E(SI) \\ &- E(I)[\text{Var}(S) - \text{Cov}(S, I)] \\ &+ E(S)[\text{Var}(I) - \text{Cov}(S, I)] \\ &- T_{SSI} + T_{SII}\}, \end{aligned} \quad (12)$$

where T_{SSI} is the third central moment $E(\{S - E(S)\}^2 \{I - E(I)\})$, and T_{SII} is defined similarly.

Notice that, as mentioned earlier, the equations for the first-order moments are not closed as they involve the second-order term $E(SI)$. It is interesting to note that the deterministic system consists of Eqs. (8) and (9) with $E(S)$ replaced by S , $E(I)$ replaced by I and $E(SI)$ replaced by SI . This last replacement is an approximation since $E(SI) = E(S)E(I) + \text{Cov}(S, I)$: the deterministic equations are not an exact representation of the mean behavior of the system for finite populations. Put another way, we could say that the deterministic model is obtained as the result of the simplest moment closure technique applied to (8) and (9), namely one in which the second-order central moments are set equal to zero.

The equations for the second-order moments are seen to contain terms which involve the third central moments. Although this expansion could be taken to higher and higher orders, yielding an infinite set of equations for all the moments of the system (cf. Matis and Kiffe, 1999), we only wish to consider terms of first and second order. Moment closure in our case corresponds to approximating the third-order terms, T_{SSI} and T_{SII} , by terms involving first- and second-order moments.

Application of the MVN to Eqs. (8)–(12) is trivial: since third-order central moments vanish for a multivariate normal distribution, one simply sets $T_{SSI} = T_{SII} = 0$. The use of this approximation can be justified by the results of Kurtz (1970, 1971), which show that the stochastic process can be approximated by a Gaussian diffusion process in the large N limit.

The lognormal approximation (Keeling, 2000a, b) also provides expressions for T_{SSI} and T_{SII} in terms of lower order moments, and so can be deployed by making the appropriate substitutions. These expressions, however, are somewhat more complex, and it turns out that the approximation is most succinctly described using so-called multiplicative moments (see Keeling, 2000a, b). A brief derivation of these moment equations (hereafter referred to as the MM equations) appears in Appendix A.

In the case of the MVN, non-negativity of population numbers imposes constraints on the mean and variance

of a normal distribution which is attempting to approximate the distribution of the states of the system: clearly, a normal distribution cannot be a good approximation if its standard deviation is sizeable compared to its mean. (As a concrete example, Näsell (1999) takes the ratio of the standard deviation to the mean being less than one third as an indication that the distribution of states is likely to be well-approximated by a MVN distribution.) In contrast, population numbers for a lognormal distribution are guaranteed to be positive and so it might be hoped that it could provide a better approximation when population sizes are small.

3.1. An aside on extinction and the stochastic process conditioned on non-extinction

In the absence of immigration, the ultimate fate of each realization of the stochastic model is that the infection will go extinct. The expected time until extinction may, however, be very long indeed when the population size is large: theoretical results indicate that this quantity, for large N , increases exponentially with N . It is possible to relate quantities such as the expected time until extinction to the level of variability seen in the model (Näsell, 1996, 1999, 2002; Andersson and Britton, 2000), but, since the main focus of this paper is variability itself, we do not pursue this further here.

By conditioning on non-extinction of the infection, one can define a new stochastic process that is related, via the extinction process, to the (unconditioned) process governing the system. The relationship between the conditioned and unconditioned processes is an important one and has been studied quite extensively for certain simple epidemic models (see, for instance, Jacquez and Simon, 1993; Näsell, 1996), but somewhat less so for multivariate models of the type considered here (see, however, Näsell, 1999).

When comparing the behavior of either the deterministic model or the moment equations to that of a collection of model realizations, it is natural to ask whether one should consider the original, unconditioned, process or the process conditioned on non-extinction. Although the Kolmogorov Eqs. (3) (and hence the resulting moment Eqs. (8)–(12)) describe the unconditioned stochastic process, the moment closure techniques that lead to the closed sets of moment equations—or, indeed, the deterministic model—take no account of extinction. Accordingly, we argue that it is more natural to compare the behavior of these equations with numerical estimates of moments obtained by conditioning on non-extinction (cf. Jacquez and Simon, 1993). (As a trivial example, the mean number of infectives in the unconditioned process will tend to zero over a sufficiently long time interval, but we would not, in general, expect to see such behavior in the

deterministic model or in the MVN or MM moment equations.) Of course, any differences between the unconditioned and conditioned processes will be small if the population size is sufficiently large that extinctions are infrequent over the timescale on which the simulations are carried out.

We remark that the Kolmogorov forward equations for the process conditioned on non-extinction of the infection can be derived quite easily: see Section 2.3, in particular equation 2.6, of Näsell (1999), and Sections II.A.5 and III.B.3 of Jacquez and Simon (1993). These equations differ from (3) in an important way: each one includes an additional term arising from the need to account for the rate at which the infection goes extinct. This extra term complicates the analysis of the resulting set of equations considerably: this is not surprising since the rate at which extinctions occur is not, in general, a quantity for which an exact expression is available. As a consequence, the usefulness of these equations for the problem at hand here is far from clear.

4. Variability in the non-seasonal model

For epidemiologically realistic parameter values, the deterministic model exhibits damped oscillations to a stable endemic equilibrium (Hethcote, 1974; Anderson and May, 1991). It can easily be shown (Anderson and May, 1991) that these oscillations have period approximately equal to $2\pi(AD)^{1/2}$ and damping time approximately equal to $2A$. Here D is the average duration of infectiousness and A is the average age at infection in the endemic state. A is approximately equal to the average lifespan, L , divided by the basic reproductive number, R_0 , with the latter being given by the expression $\beta N/(\mu + \gamma)$. (Notice that R_0 is independent of population size since β was assumed to be proportional to the reciprocal of N : see Table 1. Also notice that our restriction to childhood diseases separates the timescales between the average duration of infection (D is on the order of days), the average age at infection (A is on the order of a few years) and the average lifespan (L is on the order of tens of years).)

In the stochastic model, random effects prevent the system from settling down into equilibrium (Bartlett, 1956, 1960b; Näsell, 1999; Aparicio and Solari, 2001): individual realizations continue to oscillate around the endemic level predicted by the deterministic model. The moment equations capture this behavior and settle into an endemic state, with equilibrium values of S and I close to those predicted by the deterministic model. The equilibrium values of the second-order moments show that the numbers of infectives seen in realizations of the model can vary considerably about their mean level (Fig. 1).

In the particular case of the non-seasonal SIR model, the level of variability estimated by the moment equations can be compared to results obtained using one of several analytic approximation techniques. If the fluctuations of S and I about their mean level are assumed to be small, an analytic expression for variability can be obtained using a linear approximation of the governing equations (Bartlett, 1960b; Nåsell, 1999, 2002). For the diseases under consideration, it can be shown (Schenzle and Dietz, 1987; Lloyd, 1996; Nåsell, 1999, 2002) that the coefficient of variation of the number of infectives (i.e. the standard deviation expressed as a fraction of the mean) can, to a very good approximation, be written as

$$cv = \frac{L}{D\sqrt{R_0 N}}. \quad (13)$$

Since this expression is obtained using a linearization of the governing equations, it should only be expected to provide a good description of variability when N is large. Notice that linearization guarantees that (13) scales as $N^{-1/2}$ for all values of N .

The form of (13) is familiar from the study of demographic stochasticity in many simple population models (May, 1973), and states that as the population size increases, variability decreases and so the deterministic equations become a better and better description of the system (cf. the results of Kurtz, 1970, 1971). The factor L/D appearing in the analytic approximation (13) highlights the reason why demographic stochasticity can

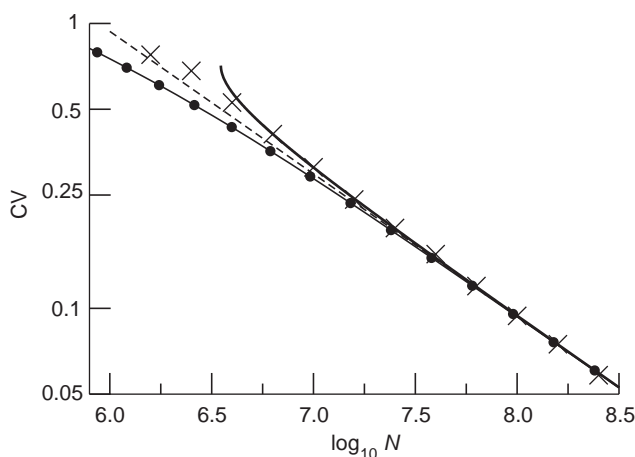


Fig. 1. Coefficient of variation of the fluctuations in the numbers of infectives for various population sizes (note the logarithmic scales). The crosses represent numerically obtained estimates, based on averages taken over one thousand realizations of the model. The heavy unbroken line represents the estimate based on the MVN moment equations. The light unbroken line with circles denotes the estimate based on the MM equations. The broken line shows the estimate based on Bartlett's stochastic linearization approximation. Parameter values are as follows: $\mu = 1/70$ year $^{-1}$, $\gamma = 50$ year $^{-1}$, $\beta = 750/N$ year $^{-1}$.

play such an important role in the dynamics of many childhood diseases: variability is large because of the separation of timescales between the infection process (occurring over days) and the demographic process (replacement of susceptibles, which occur over years). Also notice that variability about the endemic equilibrium decreases as R_0 increases: this effect can be explained by noting that oscillations about the endemic equilibrium are more strongly damped for larger values of R_0 (Aparicio and Solari, 2001, see also Lloyd, 2001b).

Fig. 1 compares the variability estimated by the MVN equations, the MM equations, Bartlett's stochastic linearization technique (13), and the variability observed in numerical simulations of the model. (Recall that the numerically obtained estimates employ moments conditioned on non-extinction of the infection.) In order to accommodate a large range of population sizes and to illustrate the $N^{-1/2}$ scaling, both axes in Fig. 1 have logarithmic scales: note that this tends to de-emphasize any differences between the various curves shown.

As previously discussed, using parameter values which are not unreasonable for a childhood disease such as measles (which has $R_0 \approx 15$ and $D \approx 5-7$ days), variability is large for many realistic population sizes (Bartlett, 1960b; Schenzle and Dietz, 1987; Lloyd, 1996; Nåsell, 1999). For populations of sizes of the order of a million, the coefficient of variation is not much smaller than one, showing that the number of infectives fluctuates significantly about its mean. This size of fluctuation implies that the number of infectives is quite likely to hit zero, and so fadeouts are common.

The moment equations are seen to perform well over a wide range of population sizes. In contrast to the deterministic model, the equilibrium values of the moment equations do depend on the population size. The second-order terms scale in such a way that the coefficient of variation of the number of infectives exhibits the expected $N^{-1/2}$ scaling for large values of N , in agreement with the prediction based on stochastic linearization.

Over an intermediate range of population sizes, there are small, but noticeable, differences between the curves in Fig. 1. The MVN predicts that variability is greater than the estimate provided by Bartlett's approximation, in agreement with the behavior observed in the numerical simulations. On the other hand, the MM equations underestimate variability, in some instances by a substantial amount.

Numerical integrations of the MVN moment equations diverge for the smallest population sizes shown in Fig. 1. A closer examination, using numerical bifurcation analysis, shows that there is a critical population size ($N \approx 10^{6.5}$) at which the stable equilibrium of the MVN moment equations disappears in a saddle-node bifurcation. We shall see similar phenomena below, and we shall return to consider its implications for the

interpretation of the deterministic equations for small population sizes in the discussion.

While the MM equations underestimate variability, their equilibrium remains stable over a wider range of population sizes. This reduced tendency to exhibit divergence was previously noted by Keeling (2000a, b). The MM equations do exhibit unexpected behavior as the population size becomes smaller still, however, with the equilibrium point becoming unstable as N decreases below about 10^5 . This appears to happen via a Hopf bifurcation, and a stable limit cycle can be observed for smaller population sizes.

For the smallest population sizes shown in Fig. 1, extinction of the infection is a frequent event (for instance, with a population size of one million, over 80% of realizations undergo extinction before 20 years). Given that the moment equations take no account of extinction, it is unsurprising that they fail to provide adequate estimates of variability in this region. Indeed, for populations of small size, even numerical estimation of variability becomes problematic as the frequency of extinction becomes high. (Note that we do not place much significance in the observation that Bartlett's linearization provides the best estimate of variability for the smallest population size shown in Fig. 1. Since Bartlett's approximation underestimates variability for moderately sized populations, but increases without bound as N becomes small, the line is bound to intersect the curve drawn through the numerically estimated variability at some point.)

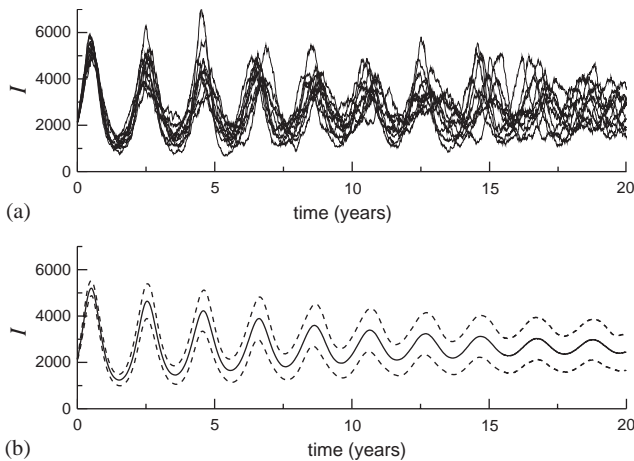


Fig. 2. (a) Numbers of infectives seen in ten realizations of the stochastic SIR model for $N = 10^7$. Initially, the numbers of susceptibles and infectives were taken to be close to the equilibrium, (S^*, I^*) , of the corresponding deterministic model, with $S = 1.05S^*$ and $I = 0.8I^*$. (b) Estimates of the average (solid line) and standard deviation (illustrated as mean \pm standard deviation) of the number of infectives seen in the stochastic model, as provided by the moment equations. Parameter values and initial conditions are as in (a), and initially the second-order central moments are taken to equal zero (as all realizations in (a) were started at identical initial conditions). Results obtained by averaging over realizations of numerical simulations are almost identical (results not shown).

It is instructive to examine the transient behavior of the system of moment equations. Integrating the moment equations from an initial condition for which second-order central moments are zero shows how variability develops between realizations started at the same initial condition. Fig. 2 illustrates the behavior of a set of such realizations, together with the variability seen between them as estimated by the MVN moment equations. The realizations are seen to oscillate about the endemic equilibrium, but slowly drift out of phase with each other. In the terminology of Nisbet and Gurney, these are phase-forgetting quasi-cycles (Nisbet and Gurney, 1982). Interestingly, this behavior, which is captured very well by either set of moment equations, explains the damped oscillations exhibited by the deterministic model: the oscillations in the mean behavior damp away as the different realizations gradually become out of phase.

5. The seasonally forced model

Seasonal variation in disease transmission plays an important role in the transmission dynamics of many childhood diseases: since schools are centers for the spread of such diseases, the probability of disease transmission is considerably higher during school terms than it is during vacations (London and Yorke, 1973; Fine and Clarkson, 1982). These seasonal variations in transmission can be shown to be responsible for the multi-annual recurrent epidemic patterns seen in real-world records of disease incidence (see, for instance, London and Yorke, 1973; Olsen et al., 1988; Olsen and Schaffer, 1990; Bolker and Grenfell, 1995). Whilst non-seasonal models for childhood diseases generally exhibit damped oscillations to a steady endemic level, the inclusion of seasonality leads to the maintenance of oscillatory incidence patterns, and deepens the inter-epidemic troughs. Seasonality, therefore, can heighten the role played by demographic stochasticity in determining both the dynamics and the persistence of childhood diseases, even in sizeable populations (see, for instance, Yorke et al., 1979; Engbert and Drepper, 1994; Bolker and Grenfell, 1995; Grenfell et al., 1995a).

Many studies of the effects of seasonality have used a simple form of forcing in which the transmission parameter, β , is taken to vary sinusoidally (Dietz, 1976)

$$\beta(t) = \beta_0(1 + \beta_1 \cos 2\pi t). \quad (14)$$

More realistic forcing functions have also been considered, either using more complex phenomenological functional forms (Kot et al., 1988), or forms which accurately depict details of the opening and closing of schools (Schenzle, 1984). For simplicity, the forcing described by Eq. (14) is employed here; the behavior of

models with more realistic forcing functions will be discussed elsewhere.

As the existing analytic methods do not cover the seasonally forced case, we assess the performance of the moment equations by comparison with variability estimates obtained by numerical simulation of the model. We focus on two particular cases: one for which the corresponding deterministic model exhibits annual dynamics ($\beta_1 = 0.01$) and the second for which biennial dynamics is observed ($\beta_1 = 0.08$). In both cases, before we can compare predicted and observed variability, we must briefly discuss the behavior observed in the simulation studies.

5.1. *Weak seasonality, $\beta_1 = 0.01$*

With this weak level of seasonality, we see that the behavior of model realizations depends quite strongly on the population size. For small populations (e.g. of the order of a few million), the realizations undergo somewhat irregular oscillations with period close to the natural frequency of the system (about 1.9 years in this case) (Fig. 3a). At first sight, apart from their increased amplitude, these oscillations appear quite similar to those seen in the unforced model.

As the population size is increased, increasingly regular oscillations which more resemble the annual oscillations of the deterministic system are observed (Fig. 3b), as is expected, since for large enough N , variability should become sufficiently small that the system approaches the deterministic limit. As N increases, it becomes clear that seasonality, as in the deterministic model, imposes a definite phase on solutions; in the unforced models epidemic peaks could occur at any time of the year, but seasonality leads to epidemic peaks occurring at definite times of the year (Lloyd and May, 1996).

The average behavior of the realizations undergoes annual oscillations in both of these cases (Figs. 3c and d). Notice that, even for the smaller population, whose realizations exhibited almost biennial dynamics, the average pattern is annual, although with an amplitude much lower than that seen in individual realizations. This is exactly analogous to what was observed in the non-forced case: the biennial component—due to the stochastic fluctuations—has no definite phase and so tends to disappear when between-realization averaging is carried out, although in this case these fluctuations are superimposed on an underlying annual (rather than a constant endemic) pattern.

As the population size increases it becomes clear that the variability between realizations also changes over the course of a cycle, with, in this case, the maximum variability observed during the increasing phase of the epidemic, and roughly coincident with the mid-point between epidemic trough and peak.

The behavior seen in these cases can be understood by noticing that variability within and between realizations arises both from seasonality and from demographic stochasticity. The importance of stochasticity, therefore, will depend on the relative sizes of the oscillations in the average behavior due to seasonality and the fluctuations due to demographic stochasticity. For small population sizes, the latter effect will tend to dominate, and the dynamics of the average will poorly reflect the dynamics of individual realizations. As the population size increases, the variability due to stochastic effects decreases. Most of the variability within individual realizations then arises from the effects of seasonality, and so the behavior of individual realizations more closely follows the average. As the amplitude of the oscillations caused by seasonality depends in a complex way on the model parameters (in particular, R_0 and β_1) (Kuznetsov and Piccardi, 1994), an analytic exploration of this effect would appear to be far from straightforward.

For large population sizes, both sets of moment equations provide excellent estimates of both the mean behavior and the variability seen about this mean (Fig. 4), exhibiting annual oscillations that closely follow the numerically obtained estimates. However, the same cannot be said for smaller populations. Numerical integration of the MVN equations leads to divergent behavior when N is below approximately $10^{6.9}$ (notice that this includes the population size shown in Figs. 3a and c). As in the unforced case, the variability estimated by the MVN equations is large just above this critical point, warning that the MVN approximation is close to breaking down.

The MM equations again underestimate variability for intermediate-sized populations. Numerical integration of the MM equations is less likely to lead to divergent behavior as N is reduced, although the stable annual cycle seen for large N becomes unstable when N falls below about 10^5 .

Notice that variability—both as seen in the numerical simulations and as estimated by the moment equations—is somewhat larger than in the unforced case, as witnessed by the larger critical value of N needed for stable annual behavior of the MVN equations compared to that needed for stable equilibrium behavior in the unforced MVN equations.

5.2. *Moderate seasonality, $\beta_1 = 0.08$*

With the stronger level of seasonality, the number of infectives in the deterministic model undergoes larger biennial oscillations, with deeper inter-epidemic troughs than seen before. The stochastic model exhibits correspondingly increased levels of variability.

For moderately sized populations, individual realizations exhibit biennial cycles similar to those seen in the

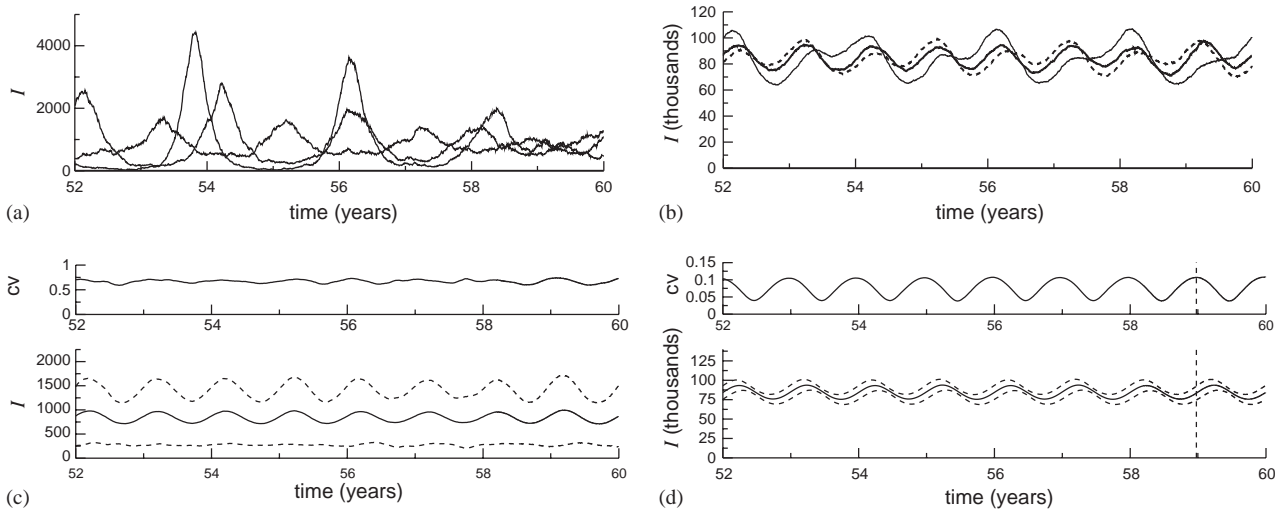


Fig. 3. Numbers of infectives seen in realizations of the weakly forced ($\beta_1 = 0.01$) stochastic SIR model, for two different population sizes: (a) $N = 10^{6.5}$ and (b) $N = 10^{8.5}$. The average and standard deviation (plotted as average \pm standard deviation) and variability of the numbers of infectives, calculated from numerical simulations of the model, for these two population sizes are shown in (c) and (d), respectively. All other parameters as in Fig. 1. Initial conditions were chosen to lie on the annual cycle of the corresponding deterministic model. Notice that results are shown after 52 simulated years have elapsed as there is a short initial transient period before realizations settle into their long-term behavior. The dotted lines in (d) illustrate a point at which the variability reaches its maximum.

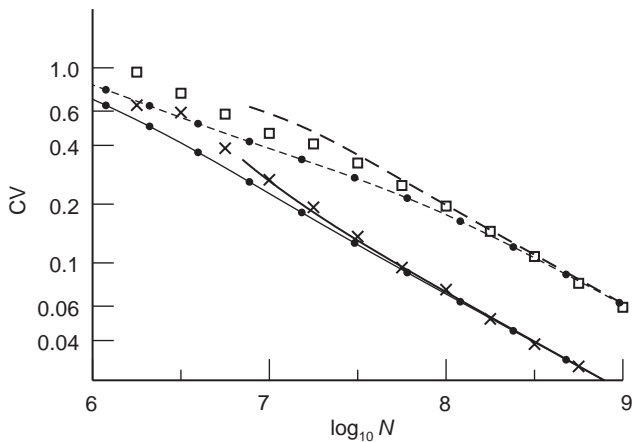


Fig. 4. Variability predicted by the moment equations and observed in numerical simulation of the weakly forced stochastic seasonal model for various population sizes. The heavy curves depict values obtained from the MVN equations, with the solid line representing the minimum value of the coefficient of variation of the number of infectives, taken over an epidemic cycle, and the dashed line its maximum value. The light curves with circles depict values obtained from the MM equations, with the minimum and maximum values of cv denoted by the unbroken and broken curves, respectively. Symbols represent minimum and maximum values of the coefficient of variation (crosses and squares, respectively), obtained by averaging over 1000 realizations of the model. All other parameter values are as in Fig. 3.

deterministic model, but that are clearly less regular (Fig. 5a). Whilst the epidemic peaks occur at very similar times in different realizations, the sizes of the peaks vary somewhat. Occasionally, the phase of the biennial epidemic pattern shifts by a year, after which the large outbreaks occur in odd, rather than even, years

(or vice versa). This effect is illustrated by the realization plotted as the broken curve in Fig. 5a.

When the population size is large, model realizations closely resemble the biennial oscillations of the corresponding deterministic model (Fig. 5b). (The chance of a realization undergoing a phase shift, as just described, appears to be insignificant for this population size: it never occurred in any of our simulations.)

For large population sizes, the average behavior exhibits biennial oscillations (Fig. 5d). The peaks in variability coincide with the peak of the small outbreak that occurs in inter-epidemic years, and troughs in variability coincide with the peaks of the large epidemics.

An interesting consequence of the phase-shifting phenomenon which occurs for moderately sized populations is that the numerically estimated average behavior has a strong annual component in its dynamics (Fig. 5c). Correspondingly, the distribution of infectives at a given point in time exhibits bimodality, arising from the occurrence of major outbreaks in either odd or even years. (Indeed, since there is no preference for large peaks to occur in odd or even years, it will eventually be the case that those realizations in which the infection has not died out will be evenly split between the two possibilities and the average behavior will exhibit purely annual oscillations.)

Both sets of moment equations give good predictions of the variability seen when the population size is large (Fig. 6). As the population size becomes smaller, the moment equations yet again provide poorer estimates of the observed variability. The moment equations fail to

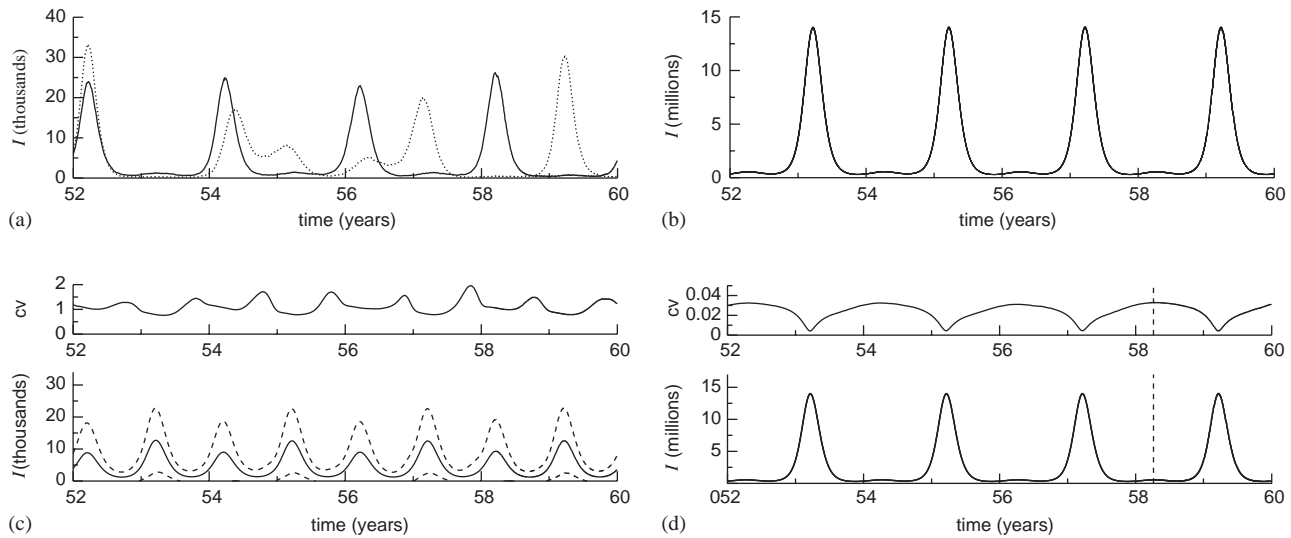


Fig. 5. Numbers of infectives seen in realizations of the moderately forced ($\beta_1 = 0.08$) stochastic SIR model, for two different population sizes: (a) $N = 10^{7.25}$ and (b) $N = 10^{10}$. As for Fig. 3, average, standard deviation and variability of the numbers of infectives for these two cases are shown in (c) and (d). All other parameters as in Fig. 1. Initial conditions were chosen to lie on the biennial cycle of the corresponding deterministic model. Notice that results are shown after 52 simulated years have elapsed as there is a short initial transient period before realizations settle into their long-term behavior.

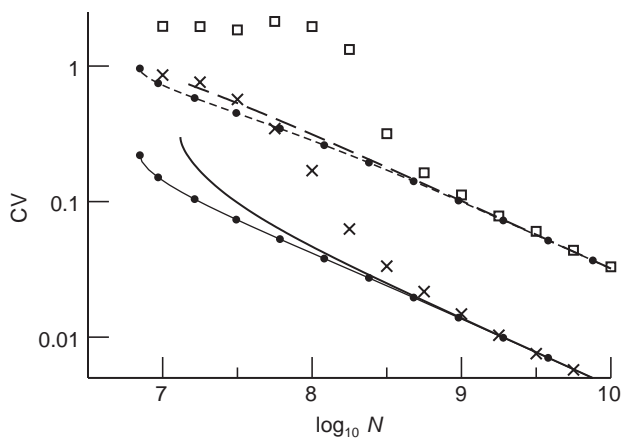


Fig. 6. Variability predicted by the moment equations and observed in numerical simulation of the moderately forced stochastic seasonal model for various population sizes. The heavy curves depict values obtained from the MVN equations, with the solid line representing the minimum value of the coefficient of variation of the number of infectives, taken over an epidemic cycle, and the dashed line its maximum value. The light curves with circles depict values obtained from the MM equations, with the minimum and maximum values of cv denoted by the unbroken and broken curves, respectively. Symbols represent minimum and maximum values of the coefficient of variation (crosses and squares, respectively), obtained by averaging over 1000 realizations of the model. All other parameter values are as in Fig. 5.

capture either the rapid increase in variability or the annual dynamics seen in the numerical estimates of the average behavior that accompany the phase-shifting phenomenon. In this situation, the failure of the moment equations not only reflects disease extinction,

but also the dynamical twist introduced by the shifting of phases: while the dynamics of the numerically estimated average exhibits a strong annual component, the moment equations exhibit purely biennial oscillations. It is, of course, not at all surprising that the moment equations cannot adequately describe such situations as the moment closure approximations assume a unimodal distribution of states. Interestingly, the failure of the moment equations occurs for somewhat smaller levels of variability than seen in previous situations in which it has failed, with an average variability of about 0.13 for $N = 10^{8.5}$, the population size below which differences become noticeable in Fig. 6.

The situation described for the smaller population size provides an interesting example of a case in which the behavior of the average of the system (as estimated by numerical simulation) is poorly represented by either the moment equations or the deterministic system.

We again notice that the MVN equations diverge for population sizes below a certain critical value ($N \approx 10^{7.1}$). Yet again, the MM equations consistently underestimate variability. Compared to the previous cases, the MM equations are much more prone to divergence as N is decreased: their stable biennial attractor is destroyed in a saddle-node bifurcation as N falls below $N \approx 10^{6.9}$. Below this point, the MM equations exhibit divergent behavior.

As we have already remarked, the behavior of the moment equations, unlike that of the corresponding deterministic system, depends on population size. Their divergence for small population sizes is a pointed example of the fact that the moment equations can

exhibit qualitatively different dynamical patterns as the population size is varied. This behavior—and its important implications for the use of deterministic models in predicting the dynamics of childhood diseases—is investigated in more detail in Appendix C.

6. Discussion

The moment equations obtained using either moment closure approximation can, in many situations, successfully predict the amount of between-realization variability generated by demographic stochasticity in finite populations. When this variability is small, we can be confident that the deterministic model will provide a good description of both the average behavior of the system and that of individual realizations. When this variability is large, or in the extreme cases where the moment equations diverge, use of the deterministic model would appear to be problematic. The moment equations, therefore, can allow the estimation of the population size required such that stochastic effects may be safely ignored in a given situation. Given that the SIR model, particularly when seasonally forced, often exhibits large variability for population sizes corresponding to large cities—and in some situations, variability remains large even when population sizes are chosen to be larger than the largest modern-day cities—this observation has clear implications for the use of deterministic descriptions of epidemic systems such as those considered here.

The large variability seen in this study is partly a consequence of the use of the simplest SIR model. As is shown in Appendix B, the inclusion of an exposed class reduces variability somewhat, leading to increased persistence in smaller populations, but seasonally forced SEIR models still exhibit substantial variability in realistically sized populations. Indeed, one of the major efforts in the modeling of childhood diseases has been an attempt to address the disparity between population sizes needed for disease persistence within the model framework and those seen in real world data by the inclusion of additional biological detail (such as age or spatial structure) within the model (Bolker and Grenfell, 1995; Keeling and Grenfell, 1997). Inclusion of immigration of infective individuals, which must be done with some care given its potential impact on disease dynamics, and more realistic descriptions of the seasonal transmission parameter may also be important, and we shall discuss these effects in a future study.

The numerical analysis has highlighted situations in which the moment equations fail to give good estimates of variability, and even cases of complete failure of the moment equations as they diverge. Generally, failures of the moment equations are associated with high levels of variability in the numerical simulations. The potential

divergence of the MVN equations is strongly indicated by considering the behavior of the moment equations as the population size is decreased from large values towards smaller values and observing the increasing variability (cf. Figs. 1, 4, 6 and 10). The failure of the moment equations in such situations should be seen as a clear indication that variability is sufficiently large that fadeout is a very common occurrence and hence that the assumptions underlying the approximation methods (and also the deterministic model) are not valid.

We do not yet have a simple criterion for determining the validity of the solutions of the moment equations without recourse to generating model realizations. As mentioned earlier, the methods used to justify the MVN suggest a simple consistency check, namely that the variability (as measured, for instance, by the coefficient of variation) remains small, to ensure its validity. Several authors have, however, illustrated situations in which the MVN appears to give reasonable results, even in situations in which it might not at first be expected to. Isham (1995), for instance, presents simulation results in which it is known that the underlying population distribution is far from normal, and yet the MVN still performs well. This observation is pursued further by Stark et al. (2001), where the MVN is viewed as one of a general family of possible moment closure techniques, and in which it is shown that other closure techniques can often outperform the MVN (see also Näsell, 2003).

Possibilities for improvement of the moment equations include the extension of the moment system to third order (see, for instance Matis and Kiffe, 1999), or the use of a moment closure method more likely to be suited to situations in which population sizes come close to zero. Indeed, the observed failings of the MVN appear, in particular its frequent divergence, to have motivated Keeling's development of the MM equations (Keeling, 2000a, b). Whilst the MM equations are less prone to divergence, we have seen that they are not immune to such behavior. Furthermore, they were seen to underestimate variability and hence would tend to predict enhanced persistence of a disease, compared to predictions made on the basis of the MVN.

However, whilst improved moment closure approximations may give a more accurate estimate of variability when the system comes close to fadeout, they do not address the central point at issue, namely that extinction leads to bimodality in the distribution of states (of the unconditioned process). Moment closure techniques of the type discussed here do not take the possible extinction of the infection into account. For many purposes, however, (e.g. in determining the applicability of a deterministic description of the system) one may not be so interested in obtaining accurate estimates of variability, but rather to understand in broad terms whether variability is likely to be small or large.

In this study, attention was focussed on particularly simple dynamical situations, ignoring the more complex dynamical issues that arise, for instance, either when the system exhibits multiple attractors or when noise can excite ordinarily unstable dynamical objects (such as ‘chaotic repellors’ or unstable saddles) (Rand and Wilson, 1991; Engbert and Drepper, 1994; Keeling et al., 2001; Billings and Schwartz, 2002; Rohani et al., 2002). Variability is likely to be significantly higher in such situations and it is unclear to what degree moment equations of the type developed here would be informative. It is possible, however, to examine transient behavior (see Fig. 2), which might be useful in many situations, such as determining whether stochasticity is likely to bring trajectories close to basin boundaries. Notice, however, that the non-local nature of such questions requires additional information concerning the structure of phase space and the positioning of pertinent dynamical objects within it (Keeling et al., 2001; Rohani et al., 2002).

An increasing number of studies in the last decade have highlighted examples of epidemiological and ecological systems (both model and real-world) whose dynamical properties result from an interplay between stochastic effects and nonlinear dynamics (Higgins et al., 1997; Grenfell et al., 1998; Keeling et al., 2001; Rohani et al., 2002). Studies such as these strongly caution against the over-reliance on dynamical predictions made on the basis of asymptotic behavior of deterministic models, unless one can be sure that population sizes are large enough to ensure that stochastic effects will not play a large role. Against this background, techniques which indicate the likely importance of stochastic effects, preferably without recourse to generation of model realizations via numerical simulation, are of clear importance and interest. The wider deployment of techniques such as the MVN presented here can only come if we gain a better understanding of their validity and limitations. This study has provided a first step in this direction, but has already highlighted some potential problems, many of which arose in situations where fadeouts occurred most frequently. Unfortunately, given the interest in the extinction or persistence of diseases or populations, these are the sorts of situations in which epidemiologists and population biologists are often most interested.

Acknowledgments

This work was supported by the Wellcome Trust, the Medical Research Council, the Leon Levy and Shelby White Initiatives Fund and the Florence Gould Foundation. The author wishes to thank Jonathan Dushoff and an anonymous referee for their helpful comments.

Appendix A. The lognormal approximation; the method of multiplicative moments

As an alternative to the multivariate normal approximation, Keeling (2000a, b) assumed that the realizations could be described by a multivariate lognormal distribution, i.e. that the logarithms of the state variables could be assumed to follow a multivariate normal distribution. The resulting moment equations can be most compactly expressed in terms of the variables \hat{V}_S , \hat{V}_I and ξ , where

$$E(S^2) = \hat{V}_S E(S)^2, \quad (\text{A.1})$$

$$E(I^2) = \hat{V}_I E(I)^2, \quad (\text{A.2})$$

$$E(SI) = \xi E(S)E(I). \quad (\text{A.3})$$

For reasons clear from their definition, these three variables are termed the multiplicative moments. Eqs. (A.1)–(A.3) can be rearranged to give expressions for the central moments in terms of the multiplicative moments: notice that equating variances and covariances to zero corresponds to setting the multiplicative moments equal to one.

The moment closure approximation made by Keeling corresponds to making the following set of substitutions:

$$E(S^3) = \hat{V}_S^3 E(S)^3, \quad (\text{A.4})$$

$$E(S^2I) = \hat{V}_S \xi^2 E(S)^2 E(I), \quad (\text{A.5})$$

$$E(SI^2) = \hat{V}_I \xi^2 E(S) E(I)^2, \quad \text{and} \quad (\text{A.6})$$

$$E(I^3) = \hat{V}_I^3 E(I)^3. \quad (\text{A.7})$$

It is fairly straightforward to show algebraically that this indeed corresponds to the assumption that the distribution of (S, I) is bivariate lognormal. Furthermore, Eqs. (A.4)–(A.7) can be directly rearranged to provide expressions for the third-order central moments T_{SSI} and T_{SII} in terms of first- and second-order moments, providing a closed set of moment equations from (8)–(12), although the resulting equations are somewhat messy.

The multiplicative moment equations are more naturally expressed in terms of the variables \hat{V}_S , \hat{V}_I and ξ . Using the method described in detail in Keeling (2000a, b), it is straightforward to obtain the following set of equations (for compactness, and using the same notation as Keeling, the expectations of S and I are written here as S and I):

$$\frac{dS}{dt} = \mu N - \beta SI \xi - \mu S, \quad (\text{A.8})$$

$$\frac{dI}{dt} = \beta SI\xi - (\mu + \gamma)I, \quad (\text{A.9})$$

$$SI \frac{d\xi}{dt} = \beta SI\xi(S\xi\{\hat{V}_S - 1\} - I\xi\{\hat{V}_I - 1\} - 1) - \mu NI(\xi - 1), \quad (\text{A.10})$$

$$I^2 \frac{d\hat{V}_I}{dt} = \beta SI\xi(1 + 2\hat{V}_I I\{\xi - 1\}) + (\mu + \gamma)I, \quad (\text{A.11})$$

$$S^2 \frac{d\hat{V}_S}{dt} = \mu N(1 + 2S\{1 - \hat{V}_S\}) + \beta SI\xi(1 + 2\hat{V}_S S\{1 - \xi\}) + \mu S. \quad (\text{A.12})$$

It should be pointed out that these equations differ slightly from those presented by Keeling. Apart from his consideration of a metapopulation model with n distinct patches (from which it is straightforward to obtain the equations describing a single, well-mixed, patch), we have included terms describing deaths of susceptibles and infectives. (For childhood diseases, since both the average duration of infection and the mean age at infection are short compared to the average lifespan, these terms are of small size.) Inspection of Eqs. (A.8)–(A.12) shows that inclusion of these terms in the model simply leads to four extra terms: $-\mu S$ in (A.8), $-\mu I$ in (A.9), μI in (A.11) and μS in (A.12). Notice that we also retain the notation μN to describe births (as opposed to Keeling's birth term B).

Appendix B. Moment equations and estimates of variability for the SEIR model

The SIR model makes the unrealistic assumption that individuals immediately become infectious upon contracting the infection. In reality, there is usually a latent period between acquisition of infection and the start of infectiousness. This can be accounted for within the model framework by allowing newly infected individuals to enter an exposed class, where they remain for an average of $1/\sigma$ time units before moving into the infectious class. If, as is taken to be the case in the SEIR model, it is assumed that the duration of latency is exponentially distributed, and the number of exposed individuals is written as E , then the movement of individuals between the exposed and infectious classes occurs at rate σE . Notice that the SIR model is recovered as the $\sigma \rightarrow \infty$ limiting case of the SEIR model.

An equation for the time evolution of the cumulant generating function can be derived in an analogous way

to that described in Section 2 for the SIR model

$$\begin{aligned} \frac{\partial K}{\partial t} = & \beta(e^{\theta_2 - \theta_1} - 1) \left(\frac{\partial^2 K}{\partial \theta_1 \partial \theta_3} + \frac{\partial K}{\partial \theta_1} \frac{\partial K}{\partial \theta_3} \right) \\ & + \mu(e^{-\theta_1} - 1) \frac{\partial K}{\partial \theta_1} + \sigma(e^{\theta_3 - \theta_2} - 1) \frac{\partial K}{\partial \theta_2} \\ & + \mu(e^{-\theta_2} - 1) \frac{\partial K}{\partial \theta_2} + (\gamma + \mu)(e^{-\theta_3} - 1) \frac{\partial K}{\partial \theta_3} \\ & + \mu N(e^{\theta_1} - 1) + \nu(e^{\theta_3} - 1). \end{aligned} \quad (\text{B.1})$$

The vector θ now has three components, $(\theta_1, \theta_2, \theta_3)$, corresponding to S , E and I , respectively. Notice that, for completeness, Eq. (B.1) includes a term representing a constant immigration of infectives at rate ν .

Expansion of (B.1) and application of the MVN yields the following set of equations for the moments of orders one and two of the SEIR model

$$\frac{dE(S)}{dt} = \mu N - \mu E(S) - \beta E(SI), \quad (\text{B.2})$$

$$\frac{dE(E)}{dt} = \beta E(SI) - (\sigma + \mu)E(E), \quad (\text{B.3})$$

$$\frac{dE(I)}{dt} = \nu + \sigma E(E) - (\gamma + \mu)E(I), \quad (\text{B.4})$$

$$\begin{aligned} \frac{d \text{Var}(S)}{dt} = & \mu\{N + E(S) - 2\text{Var}(S)\} + \beta E(SI) \\ & - 2\beta\{E(S)\text{Cov}(S, I) + \text{Var}(S)E(I)\}, \end{aligned} \quad (\text{B.5})$$

$$\begin{aligned} \frac{d \text{Var}(E)}{dt} = & (\mu + \sigma)\{E(E) - 2\text{Var}(E)\} + \beta E(SI) \\ & + 2\beta\{E(S)\text{Cov}(E, I) \\ & + \text{Cov}(S, E)E(I)\}, \end{aligned} \quad (\text{B.6})$$

$$\begin{aligned} \frac{d \text{Var}(I)}{dt} = & \nu + (\mu + \gamma)\{E(I) - 2\text{Var}(I)\} \\ & + \sigma\{E(E) + 2\text{Cov}(E, I)\}, \end{aligned} \quad (\text{B.7})$$

$$\begin{aligned} \frac{d \text{Cov}(S, E)}{dt} = & -(2\mu + \sigma)\text{Cov}(S, E) \\ & - \beta\{E(SI) - E(S)\text{Cov}(S, I)\} \\ & + \beta\{\text{Var}(S)E(I) - \{E(S)\text{Cov}(E, I) \\ & + \text{Cov}(S, E)E(I)\}\}, \end{aligned} \quad (\text{B.8})$$

$$\begin{aligned} \frac{d \text{Cov}(S, I)}{dt} = & -(2\mu + \gamma)\text{Cov}(S, I) + \sigma\text{Cov}(S, E) \\ & - \beta E(S)\text{Var}(I) \\ & - \beta \text{Cov}(S, I)E(I), \end{aligned} \quad (\text{B.9})$$

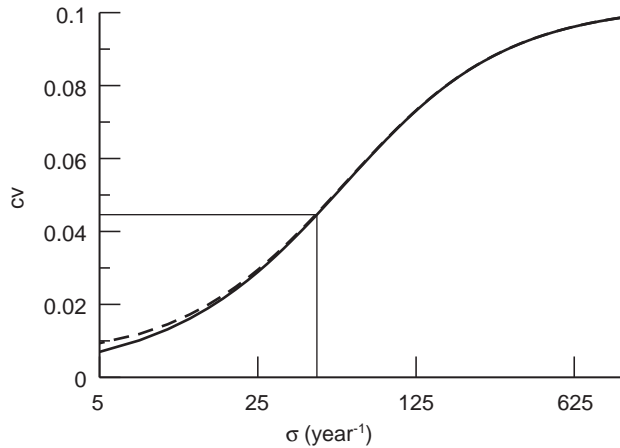


Fig. 7. Variability seen in the SEIR model with different average durations of the latent period. The broken curve denotes the coefficient of variation of the number of infectives and the solid curve denotes the coefficient of variation of the number of individuals who are infected (i.e. the sum of the exposed and infective individuals). Parameter values are $\beta = 1010.7/N \text{ year}^{-1}$, $\gamma = 73.0 \text{ year}^{-1}$ and $\mu = 0.02 \text{ year}^{-1}$. This set of parameter values has been commonly used in the literature, and represents a disease with a 5 day infectious period and a R_0 value of about 14, in a population with average lifespan 50 years (Olsen et al., 1988). A population size of 10^8 was used, although this value (provided N is sufficiently large that fadeouts are not so frequent) is not important here as changing it would merely rescale the y -axis. The light line highlights a parameter value commonly used in the literature ($\sigma = 45.6 \text{ year}^{-1}$, corresponding to an 8-day latent period).

$$\begin{aligned} \frac{d\text{Cov}(E, I)}{dt} = & -(2\mu + \sigma + \gamma)\text{Cov}(E, I) \\ & - \sigma\{E(E) - \text{Var}(E)\} \\ & + \beta\{E(S)\text{Var}(I) \\ & + \text{Cov}(S, I)E(I)\}. \end{aligned} \quad (\text{B.10})$$

Although it has long been realized in numerical simulations that inclusion of the exposed class tends to enhance persistence by reducing variability about the mean, an equation of the form (13) has only recently been derived for the SEIR model (Andersson and Britton, 2000), and so there has been less systematic study of variability in this model. For parameter values appropriate for childhood diseases, and writing $r = \gamma/\sigma$, the variability about the endemic equilibrium of the SEIR model is approximately

$$\text{cv} = \frac{L}{D\sqrt{R_0 N}} \left(\frac{1}{1 + r + r^2 + (1 + r)^2/R_0} \right). \quad (\text{B.11})$$

Fig. 7 illustrates how variability depends on the average duration of the latent period, confirming that variability is indeed lower in SEIR models with latent periods of realistic durations. The light line indicates the value of the latency parameter assumed in several modeling studies of measles dynamics. For this set of parameter values, variability is roughly 40% of that seen in the

corresponding SIR model, in agreement with the value of the factor that appears in brackets in Eq. (B.11).

Appendix C. Dynamics of the seasonally forced moment equations

In this appendix, we examine the dynamics of the seasonally forced MVN moment equations as either the strength of seasonality or the population size is varied, using the software package CONTENT (Kuznetsov and Levitin, 1995–1997) to perform numerical bifurcation analyses. The second of these analyses has no counterpart in the deterministic literature as the behavior of the deterministic system as seasonality is increased has been comprehensively studied in the literature (Olsen et al., 1988; Olsen and Schaffer, 1990; Schwartz, 1985; Aron and Schwartz, 1984; Engbert and Drepper, 1994; Earn et al., 2000; Lloyd, 2001a), however, and it is useful to briefly review some relevant findings before discussing the behavior seen in similar analyses of the moment equations.

The changing behavior of the deterministic model as seasonality is varied is illustrated in Fig. 8, which shows the number of infectives seen at the start of each year. Regular annual oscillations are observed with weak levels of seasonality. When the level of seasonality is moderate, i.e. over a wide range of values of the baseline transmission parameter, regular biennial cycles are seen

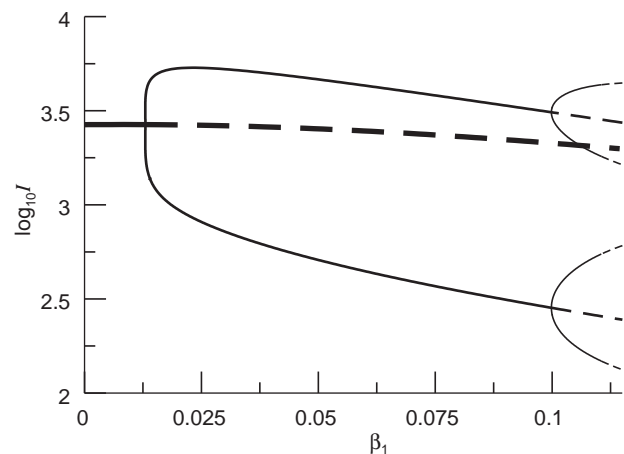


Fig. 8. Numbers of infectives seen at the start of each year (i.e. at times $t = n$, where n is an integer) in the deterministic model, for varying strengths of seasonality. For small values of β_1 , a single point is observed, corresponding to annual behavior. For larger values of β_1 , two points are observed, corresponding to biennial behavior. Solid lines correspond to stable solutions, dashed lines to unstable solutions. The thickest curve denotes annual behavior, the medium curve denotes biennial behavior and the lightest curve shows period four solutions. Other behaviors are also seen for certain parameter ranges; only those which involve the main annual and biennial attractors are shown. Parameter values are as in Fig. 1, except that N was taken to equal 10^7 .

(Aron and Schwartz, 1984; Kuznetsov and Piccardi, 1994; Earn et al., 2000), although there are different mechanisms which lead to the appearance of these cycles (Kuznetsov and Piccardi, 1994; Lloyd, 2001a).

For the parameter values illustrated in Fig. 8, biennial (and higher period) cycles appear via the period doubling process (Aron and Schwartz, 1984). Further period doublings can be seen as seasonality is strengthened; with sufficiently strong forcing, chaotic behavior would be generated (Olsen et al., 1988; Olsen and Schaffer, 1990). The important point to notice is that as seasonality is strengthened, the amplitude of the oscillations increases and the numbers of infectives in the troughs between epidemics decreases (in the chaotic regime, the numbers of infectives fall to levels well below a single individual).

Additional periodic solutions exist for some parameter values (Schwartz, 1985; Engbert and Drepper, 1994), for instance triennial cycles are often observed. Importantly for the behavior of the stochastic model, these solutions are again often associated with low numbers of infectives in inter-epidemic periods.

Numerical bifurcation analyses of the moment equations as seasonality is varied are presented in Fig. 9, which employs two different population sizes, $N = 10^8$ and 10^{10} . Even though the former case represents a population which is much larger than any present-day city, significant variability is observed. As in the discussion of variability in the main text, we remark that the large population sizes employed here reflect the large variability seen in the model: populations of the size of tens or hundreds of millions are required for persistence of infection within this model.

There is a general correspondence between the behavior of the moment equations and the deterministic

system. Several differences are apparent, however, and as might be expected, these differences are more pronounced for the smaller population size. Most notably, the transition between annual and biennial behavior is more complex. Unlike the deterministic model, the moment equations do not exhibit a period doubling bifurcation. The stable annual solution continues to exist beyond the point at which biennial cycles appear, although it is associated with high levels of variability (and so would be unlikely to be observed as a stable behavior in realizations of the stochastic model). To what extent these differences, particularly in the vicinity of bifurcation points where the model comes close to neutral stability (in the linear sense), reflect failings of the MVN approximation remains to be seen.

An important observation from Fig. 9 is the tendency of variability to increase with increasing seasonality, in agreement with intuition based on the increasing amplitude of epidemics as the system becomes more strongly forced.

Finally, we examine in more detail the behavior of the moment equations at a given level of seasonality (in this case we examine the moderately forced, $\beta_1 = 0.08$, model) as N is varied (Fig. 10). The average variability of the predominant stable behavior—the biennial cycle—exhibits the expected $N^{-1/2}$ scaling, but as we saw earlier, there is a population size below which it becomes unstable and is no longer observed. (Notice that there is a range of population sizes over which the biennial cycle is not stable: further bifurcations occur within this window.)

In addition to the expected period two solution there is also an annual attractor, however, which exhibits substantially greater variability than the biennial attractor.

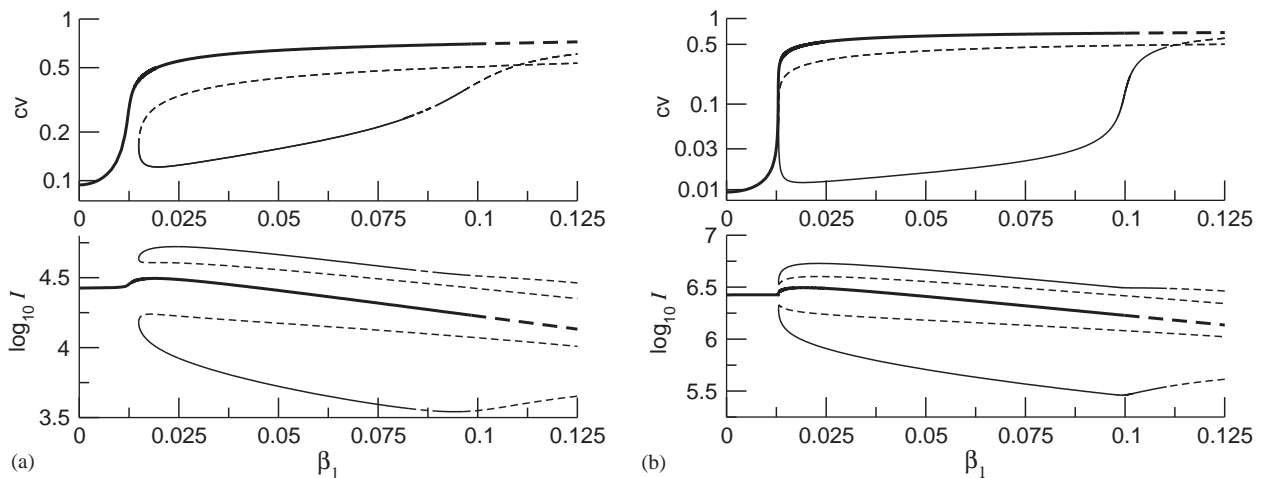


Fig. 9. Numbers of infectives at the start of each year (lower panel) and variability (averaged over an epidemic cycle) predicted by the MVN equations (upper panel) over a range of strengths of seasonal forcing, for populations of size (a) $N = 10^8$ and (b) 10^{10} . Heavy lines denote annual behavior and lighter lines biennial behavior, solid curves stable dynamics and broken curves unstable dynamics. All other parameter values as in Fig. 1.

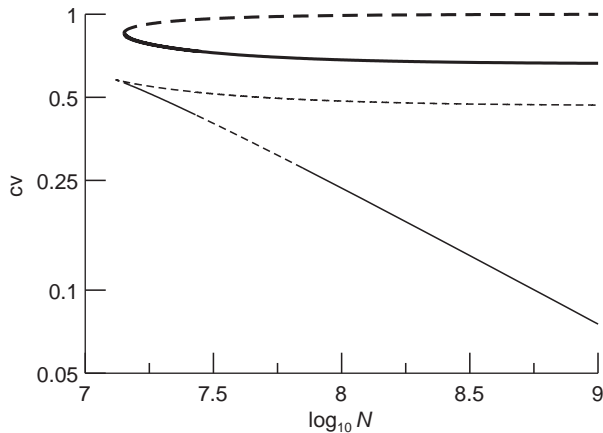


Fig. 10. Average variability predicted by the moment equations for the moderately forced model ($\beta_1 = 0.08$) over a range of population sizes. Annual behavior is represented by the heavy curve, and biennial behavior by the lighter curve. Stable behavior is indicated by solid curves and unstable behavior by broken curves. Notice that both stable annual and biennial behaviors are seen for the same set of parameter values, in contrast to the corresponding deterministic model. All other parameter values as in Fig. 1.

This corresponds to the large-variability annual solution seen after the appearance of biennial cycles in Fig. 9. We remark that this annual solution does not give a good description of the annual dynamics observed in the simulation-based average behavior of the realizations (Fig. 5c); as discussed in the main text, we would not expect the MVN to be able to capture this behavior. Interestingly, the average variability of this solution does not appear to scale in the expected way, decreasing only very slowly as N increases. We suggest that this annual solution does not, therefore, correspond to an average of some collection of realizations, and is instead an artefactual solution arising from the multivariate normal approximation. Interestingly, in his original development of the MVN, Whittle noticed equilibrium points of the moment equations which did not correspond to equilibria of the corresponding deterministic system (Whittle, 1957). In the examples he studied, however, such spurious equilibria of the moment equations were, in contrast to the annual solution described here, all linearly unstable.

To summarize the results of this section: the qualitative behavior of the moment equations, unlike the deterministic model, depends on the population size. The moment equations also exhibit a more complex dependence on the strength of seasonality; in particular, the transition between annual and biennial behavior is more subtle than the simple period doubling seen in the deterministic model with this set of parameters. The moment equations can also exhibit additional dynamical changes not seen in the corresponding deterministic model.

References

- Anderson, R.M., May, R.M., 1991. *Infectious Diseases of Humans*. Oxford University Press, Oxford.
- Andersson, H., Britton, T., 2000. Stochastic epidemics in dynamic populations: quasi-stationarity and extinction. *J. Math. Biol.* 41, 559–580.
- Aparicio, J.P., Solari, H.G., 2001. Sustained oscillations in stochastic systems. *Math. Biosci.* 169, 15–25.
- Aron, J.L., Schwartz, I.B., 1984. Seasonality and period-doubling bifurcations in an epidemic model. *J. Theor. Biol.* 110, 665–679.
- Bartlett, M.S., 1956. Deterministic and stochastic models for recurrent epidemics. In: Neyman, J. (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4. University of California Press, Berkeley, pp. 81–109.
- Bartlett, M.S., 1957. Measles periodicity and community size. *J.R. Stat. Soc. A* 120, 48–70.
- Bartlett, M.S., 1960a. The critical community size for measles in the United States. *J.R. Stat. Soc. A* 123, 37–44.
- Bartlett, M.S., 1960b. *Stochastic Population Models*. Methuen, London.
- Billings, L., Schwartz, I.B., 2002. Exciting chaos with noise: unexpected dynamics in epidemic outbreaks. *J. Math. Biol.* 44, 31–48.
- Black, F.L., 1966. Measles endemicity in insular populations: critical community size and its evolutionary implications. *J. Theor. Biol.* 11, 207–211.
- Bolker, B., Grenfell, B., 1995. Space, persistence and dynamics of measles epidemics. *Philos. Trans. R. Soc. London B* 348, 309–320.
- Bolker, B., Pacala, S.W., 1997. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theor. Popul. Biol.* 52, 179–197.
- Bolker, B.M., Grenfell, B.T., 1996. Impact of vaccination on the spatial correlation and persistence of measles dynamics. *Proc. Natl. Acad. Sci. USA* 93, 12648–12653.
- Diekmann, O., Heesterbeek, J.A.P., 2000. *Mathematical Epidemiology of Infectious Diseases*. Wiley, Chichester.
- Dietz, K., 1976. The incidence of infectious diseases under the influence of seasonal fluctuations. *Lect. Notes Biomath.* 11, 1–15.
- Dietz, K., Schenzle, D., 1985. Mathematical models for infectious disease statistics. In: Atkinson, A.C., Fienberg, S.E. (Eds.), *A Celebration of Statistics*. Springer, New York, pp. 167–204.
- Earn, D.J.D., Rohani, P., Bolker, B.M., Grenfell, B.T., 2000. A simple model for complex dynamical transitions in epidemics. *Science* 287, 667–670.
- Engbert, R., Drepper, F.R., 1994. Chance and chaos in population biology-models of recurrent epidemics and food chain dynamics. *Chaos Solitons Fractals* 4, 1147–1169.
- Feller, W., 1968. *An Introduction to Probability Theory and its Applications*. Wiley, New York.
- Fine, P.E.M., Clarkson, J.A., 1982. Measles in England and Wales—I: an analysis of factors underlying seasonal patterns. *Int. J. Epidemiol.* 11, 5–14.
- Grenfell, B.T., Bolker, B.M., Kleczkowski, A., 1995a. Seasonality and extinction in chaotic metapopulations. *Proc. R. Soc. London B* 259, 97–103.
- Grenfell, B.T., Wilson, K., Isham, V.S., Boyd, H.E.G., Dietz, K., 1995b. Modelling patterns of parasite aggregation in natural populations: trichostrongylid nematode–ruminant interactions as a case study. *Parasitology* 111, S135–S151.
- Grenfell, B.T., Wilson, K., Finkenstadt, B.F., Coulson, T.N., Murray, S., Albon, S.D., Pemberton, J.M., Clutton-Brock, T.H., Crawley, M.J., 1998. Noise and determinism in synchronized sheep dynamics. *Nature* 394, 674–677.
- Hethcote, H.W., 1974. Asymptotic behaviour and stability in epidemic models. In: van den Driessche, P. (Ed.), *Mathematical Problems in*

- Biology, Lecture Notes in Biomathematics, Vol. 2. Springer, Berlin, pp. 83–92.
- Higgins, K., Hastings, A., Sarvela, J.N., Botsford, L.W., 1997. Stochastic dynamics and deterministic skeletons: population behavior of Dungeness crab. *Science* 276, 1431–1435.
- Isham, V., 1991. Assessing the variability of stochastic epidemics. *Math. Biosci.* 107, 209–224.
- Isham, V., 1995. Stochastic models of host-macroparasite interaction. *Ann. Appl. Probab.* 5, 720–740.
- Jacquez, J.A., Simon, C.P., 1993. The stochastic SI model with recruitment and deaths. I. Comparison with the closed SIS model. *Math. Biosci.* 117, 77–125.
- Keeling, M.J., 2000a. Metapopulation moments: coupling, stochasticity and persistence. *J. Anim. Ecol.* 69, 725–736.
- Keeling, M.J., 2000b. Multiplicative moments and measures of persistence in ecology. *J. Theor. Biol.* 205, 269–281.
- Keeling, M.J., Grenfell, B.T., 1997. Disease extinction and community size: modeling the persistence of measles. *Science* 275, 65–67.
- Keeling, M.J., Rohani, P., Grenfell, B.T., 2001. Seasonally forced disease dynamics explored as switching between attractors. *Physica D* 148, 317–335.
- Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London A* 115, 700–721.
- Kot, M., Schaffer, W.M., Truty, G.L., Graser, D.J., Olsen, L.F., 1988. Changing criteria for imposing order. *Ecol. Modelling* 43, 75–110.
- Kurtz, T.G., 1970. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probab.* 7, 49–58.
- Kurtz, T.G., 1971. Limit theorems for sequences of jump Markov processes approximating ordinary differential equations. *J. Appl. Probab.* 8, 344–356.
- Kuznetsov, Y.A., Levitin, V.V., 1995–1997. CONTENT: A multiplatform environment for analyzing dynamical systems. Dynamical Systems Laboratory, Centrum voor Wiskunde en Informatica, Amsterdam, available at <http://ftp.cwi.nl/CONTENT>
- Kuznetsov, Y.A., Piccardi, C., 1994. Bifurcation analysis of periodic SEIR and SIR epidemic models. *J. Math. Biol.* 32, 109–121.
- Lloyd, A.L., 1996. Mathematical models for spatial heterogeneity in population dynamics and epidemiology. Ph.D. Thesis, University of Oxford.
- Lloyd, A.L., 2001a. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. *Proc. R. Soc. London B* 268, 985–993.
- Lloyd, A.L., 2001b. Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theor. Popul. Biol.* 60, 59–71.
- Lloyd, A.L., May, R.M., 1996. Spatial heterogeneity in epidemic models. *J. Theor. Biol.* 179, 1–11.
- London, W.P., Yorke, J.A., 1973. Recurrent outbreaks of measles, chickenpox and mumps. I. seasonal variation in contact rates. *Am. J. Epidemiol.* 98, 453–468.
- Matis, J.H., Kiffe, T.R., 1999. Effects of immigration on some stochastic logistic models: a cumulant truncation analysis. *Theor. Popul. Biol.* 56, 139–161.
- May, R.M., 1973. *Stability and Complexity in Model Ecosystems*. Princeton University Press, Princeton.
- Nåsell, I., 1996. The quasi-stationary distribution of the closed endemic SIS model. *Adv. Appl. Probab.* 28, 895–932.
- Nåsell, I., 1999. On the time to extinction in recurrent epidemics. *J.R. Stat. Soc. B* 61, 309–330.
- Nåsell, I., 2002. Stochastic models of some endemic infections. *Math. Biosci.* 179, 1–19.
- Nåsell, I., 2003. An extension of the moment closure method. *Theor. Popul. Biol.* 264, 233–239.
- Nisbet, R.M., Gurney, W.S.C., 1982. *Modelling Fluctuating Populations*. Wiley, Chichester.
- Olsen, L.F., Schaffer, W.M., 1990. Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science* 249, 499–504.
- Olsen, L.F., Truty, G.L., Schaffer, W.M., 1988. Oscillations and chaos in epidemics: a nonlinear dynamic study of six childhood diseases in Copenhagen, Denmark. *Theor. Popul. Biol.* 33, 344–370.
- Rand, D.A., Wilson, H.B., 1991. Chaotic stochasticity: a ubiquitous source of unpredictability in dynamics. *Proc. R. Soc. London B* 246, 179–184.
- Renshaw, E., 1991. *Modelling Biological Populations in Space and Time*. Cambridge University Press, Cambridge.
- Rohani, P., Keeling, M.J., Grenfell, B.T., 2002. The interplay between determinism and stochasticity in childhood diseases. *Am. Nat.* 159, 469–481.
- Schenzle, D., 1984. An age-structured model of pre- and post-vaccination measles transmission. *IMA J. Math. Appl. Med. Biol.* 1, 169–191.
- Schenzle, D., Dietz, K., 1987. Critical population sizes for endemic virus transmission. In: Fricke, W., Hinz, E. (Eds.), *Räumliche Persistenz und Diffusion von Krankheiten*. Heidelberger Geographische Arbeiten, Heidelberg, pp. 31–42.
- Schwartz, I.B., 1985. Multiple stable recurrent outbreaks and predictability in seasonally forced nonlinear epidemic models. *J. Math. Biol.* 21, 347–361.
- Stark, J., Iannelli, P., Baigent, S., 2001. A nonlinear dynamics perspective of moment closure for stochastic processes. *Nonlinear Anal.* 47, 753–764.
- Whittle, P., 1957. On the use of normal approximation in the treatment of stochastic processes. *J. R. Stat. Soc. B* 19, 268–281.
- Yorke, J.A., Nathanson, N., Pianigiani, G., Martin, J., 1979. Seasonality and the requirements for perpetuation and eradication of viruses in populations. *Am. J. Epidemiol.* 109, 103–123.