

# Noise and Nonlinearity in Measles Epidemics: Combining Mechanistic and Statistical Approaches to Population Modeling

S. P. Ellner,<sup>1,\*</sup> B. A. Bailey,<sup>1,†</sup> G. V. Bobashev,<sup>1,‡</sup> A. R. Gallant,<sup>2</sup> B. T. Grenfell,<sup>3</sup> and D. W. Nychka<sup>1</sup>

1. Biomathematics Program, Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695-8203;

2. Department of Economics, University of North Carolina, Chapel Hill, North Carolina 27599-3305;

3. Department of Zoology, Cambridge University, Downing Street, Cambridge CB2 3EJ, United Kingdom

Submitted July 7, 1997; Accepted November 17, 1997

**abstract:** We present and evaluate an approach to analyzing population dynamics data using semimechanistic models. These models incorporate reliable information on population structure and underlying dynamic mechanisms but use nonparametric surface-fitting methods to avoid unsupported assumptions about the precise form of rate equations. Using historical data on measles epidemics as a case study, we show how this approach can lead to better forecasts, better characterizations of the dynamics, and a better understanding of the factors causing complex population dynamics relative to either mechanistic models or purely descriptive statistical time-series models. The semimechanistic models are found to have better forecasting accuracy than either of the model types used in previous analyses when tested on data not used to fit the models. The dynamics are characterized as being both nonlinear and noisy, and the global dynamics are clustered very tightly near the border of stability (dominant Lyapunov exponent  $\lambda \approx 0$ ). However, locally in state space the dynamics oscillate between strong short-term stability and strong short-term chaos (i.e., between negative and positive local Lyapunov exponents). There is statistically significant evidence for short-term chaos in all data sets examined. Thus the nonlinearity in these systems is characterized by the variance over state space in local measures of chaos versus stability rather than a single summary measure of the overall dynamics as either chaotic or nonchaotic.

**Keywords:** population dynamics, modeling, measles, time-series analysis, local Lyapunov exponents.

\* To whom correspondence should be addressed; E-mail: ellner@stat.ncsu.edu.

† Present address: Geophysical Statistics Project, National Center for Atmospheric Research, Boulder, Colorado 80307.

‡ Present address: Department of Mental Hygiene, School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205.

Am Nat. 1998. Vol. 151, pp. 425–440. © 1998 by The University of Chicago. 0003-0147/98/5105-0003\$03.00. All rights reserved.

Historical data on recurrent epidemics have generated sustained interest among epidemiologists and population biologists seeking to understand the causes and consequences of the complex mix of seasonal and irregular oscillations in disease prevalence (e.g., Rhodes and Anderson 1996; Keeling and Grenfell 1997; Levin et al. 1997). Much recent interest traces to influential papers by Schaffer and coworkers (reviewed by Kot et al. 1988; Schaffer et al. 1990), which contend that these data provide compelling evidence of chaotic dynamics occurring in a natural biological population. It remains uncertain whether any populations of macro-organisms exhibit chaos (Hastings et al. 1993). Recent analyses based on nonlinear time-series modeling (Turchin 1993; Ellner and Turchin 1995) and mechanistic population modeling (Hanski et al. 1993; Costantino et al. 1995; Dennis et al. 1995) have identified a few likely cases, but these methods need further testing and the data sets on macro-organisms are all short (usually 30–100 data points). Epidemic data series are longer, typically over 400 data points, and their accuracy is probably better due to the importance attached to human disease notifications. Consequently, as Tidd et al. (1993, p. 258), observe, “Chaos in childhood diseases is thus something of a test case on which hinges a good deal more than the dynamics of some half dozen pathogenic agents.”

Two main approaches have been used to analyze these data, leading to different conclusions about the dynamics. The first approach (Schaffer et al. 1990; Tidd et al. 1993) is built on comparisons between the data and simple mechanistic models consisting of a few differential equations (described below). For measles in particular, certain features of the data are consistent with model output only for parameter values at which the models are chaotic. Measles is consequently interpreted as a deterministic chaotic system with some noise superimposed (e.g., sampling effects due to finite population size).

The second approach is based on fitting phenomenological time-series models

$$x_{t+T} = f(x_t, x_{t-L}, \dots, x_{t-mL}, t) + e_t \quad (1)$$

(e.g., Nychka et al. 1992; Ellner and Turchin 1995; Ellner et al. 1995). Here,  $x_t$  is the log number of cases in time period  $t$  (quarterly, monthly, or weekly totals);  $T$  is the forecasting interval;  $e_t$  represents exogenous environmental noise;  $L$  is the time lag;  $m$  is the index for number of lags; and  $f$  is a flexible nonlinear model whose parameters are estimated from the data. Using (1), measles epidemics in large developed-world cities were estimated to be weakly stable rather than chaotic (Ellner et al. 1995) and clustered near the border between stable and chaotic dynamics. The dynamics were also estimated to have non-negligible noise (Ellner et al. 1995; mean  $r^2 = 0.82 \pm 0.09$  [SD],  $n = 12$  cities, for forecasts one-quarter ahead).

Both modeling approaches have significant limitations. A data analysis based on a fully specified mechanistic model assumes that we know everything there is to know about the processes causing the dynamics and that all relevant processes are included in the model. In fact, obtaining a low-dimensional mechanistic epidemic model requires drastic simplifying assumptions, and many potentially important aspects of the dynamics must be omitted (e.g., variability among individuals and the family and social structure of disease transmission). Data analyses that assume the models to be quantitatively accurate are consequently prone to bias. In particular, use of a deterministic model can strongly bias analyses toward “detecting” chaos in nonchaotic data (Ellner 1991). Conversely, an analysis based on (1) assumes that we know nothing whatsoever about any of the underlying processes and treats the data as simply a string of numbers. This reduces the precision of estimates and thus decreases the likelihood of detecting interesting features in the data (Tidd et al. 1993), including (but not limited to) chaotic dynamics. Tidd et al. compared the two approaches, based on their ability to forecast data that were not used in fitting the models, and concluded that the mechanistic approach was superior. However, the phenomenological models in their comparison were all linear, whereas the analyses based on (1) indicate that the dynamics are nonlinear as well as noisy (see table 1).

Our goals here are twofold. First, we complete the program initiated by Tidd et al. to provide a quantitative basis for deciding between these alternative characterizations of the dynamics, by adding to the comparison phenomenological models representing the hypothesis of noisy nonlinear dynamics. These complete the experimental design of Tidd et al. (1993), making it possible to reach conclusions about the contrast of main interest (mechanistic vs. phenomenological; table 1).

Second, we present and evaluate an intermediate class of models that we call *semimechanistic*. The goal of this approach is to obtain more precise estimates by making use of reliable biological information, while using non-

**Table 1:** The experimental design of model comparisons in Tidd et al. (1993) and in this article

	<i>Mechanistic</i>	<i>Phenomenological</i>	<i>Semimechanistic</i>
Tidd et al. (1993):			
Linear	...	M1, M2, M3	...
Nonlinear	SEIR, RAS	...	...
This article:			
Nonlinear	SEIR	FNN, SNP	SC_FNN

Note: The prediction models compared by Tidd et al. (1993) were either nonlinear mechanistic models (SEIR, RAS), in which the only noise is demographic stochasticity due to finite population size, or linear phenomenological models driven by high levels of environmental noise (M1, M2, M3). The SEIR and RAS models are described in the text. M1 is a linear Gaussian autoregressive process, M2 is nonlinear transformation of a linear Gaussian autoregressive process, and M3 is M2 superimposed on an annual periodic trend. Thus, M1 is a special case of M2, which is a special case of M3. In this article, we compare SEIR with models that allow both nonlinearity and substantial levels of environmental noise (FNN, SNP, SC\_FNN models are described in the text).

parametric surface-fitting methods (described below) to avoid unsupported assumptions about the exact form of rate equations. In the case of measles, the semimechanistic model that we propose uses mechanistic information to define a priori a meaningful state space and to specify the model’s compartmental structure, but the contact rate equation is estimated from the data.

As with epidemics, our state of knowledge about ecological systems is typically somewhere between the extremes implicit in mechanistic versus phenomenological modeling: we do not have complete knowledge about all underlying mechanisms, but we are not completely ignorant. A model representing our state of knowledge therefore would be partially mechanistic but would remain flexible and phenomenological about poorly known aspects of the system (Wood 1994). We therefore believe that the semimechanistic approach will be advantageous for modeling the dynamics of many populations and ecosystems based on time-series data for a range of purposes including forecasting, control, and identifying underlying mechanisms. Here, taking measles as a case study, we illustrate how the semimechanistic approach leads to improved forecasting accuracy, improved characterization of measles dynamics, and additional information about the underlying causes.

## Models

### *Mechanistic*

The mechanistic model we consider is the seasonally forced basic SEIR model (defined below). We chose this

model because it was found to give the highest forecasting accuracy by Tidd et al. (1993), and it is the prototype for most of the more detailed measles models. A more complex model (e.g., the Realistic Age Structure [RAS] model; Schenzle 1984; Bolker and Grenfell 1993; Keeling and Grenfell 1997) need not give more accurate forecasts because of the difficulty of estimating accurately the larger number of parameters.

The SEIR model is a system of nonlinear differential equations that describes the changes over time in the fraction of Susceptible, Exposed (latent), Infective, and Recovered individuals in the population. In the case of measles the basic model takes the form

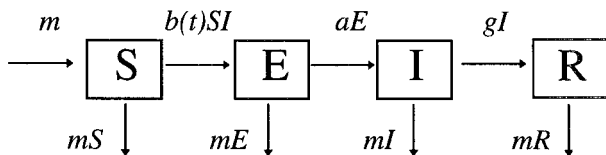
$$\begin{aligned} dS/dt &= m(1 - S) - b(t)SI; \\ dE/dt &= b(t)SI - (m + a)E; \\ dI/dt &= aE - (m + g)I; \end{aligned} \tag{2a}$$

and

$$dR/dt = gI - mR,$$

where  $m$  is the mortality rate (assumed to be the same for all classes),  $b(t)$  is the contact intensity, and  $1/a$  and  $1/g$  are the mean durations of the exposed and infectious periods. The differential equations in (2a) correspond to the compartment diagram shown in figure 1. This form of the model assumes that all newborns are susceptible, recovered individuals have permanent immunity (which is appropriate for measles), and population size is constant. Contact intensity,  $b(t)$ , is assumed to follow a deterministic annual cycle caused by seasonality and the school year. Conventionally,  $b(t)$  is modeled by a sine wave, but we used the more realistic form introduced by Kot et al. (1988):

$$\begin{aligned} b(t) &= \beta_0(1 + \beta_1\varphi(t)), \varphi(t) \\ &= 1.5(0.68 + \cos(2\pi t))/(1.5 + \cos(2\pi t)). \end{aligned} \tag{2b}$$



**Figure 1:** Compartment diagram of the SEIR model, for measles, corresponding to equation (2) in the text. The four compartments are the fractions of Susceptible, Exposed (latent), Infective, and Recovered individuals in the population, hence  $S + E + I + R = 1$ . All newborns are assumed to enter the susceptible class eventually (either at birth or after a time delay), and recovery is permanent. In the form of the model used here and by Tidd et al. (1993), the total population size is assumed to be constant, hence the recruitment rate ( $m$ ) equals the total death rate.

The contact rate equation in (2), namely  $b(t)SI$ , is the conventional mass-action equation, which is also used, for example, in the Lotka-Volterra predator-prey model and in host-parasitoid models in continuous time.

### Phenomenological

We considered two very different phenomenological nonlinear time-series models, feed-forward neural networks (FNN) and the seminonparametric (SNP) model of Gallant and Tauchen (1992). Given the amount of data available on measles, our past experience and comparative studies favor FNN for characterizing nonlinear population dynamics (McCaffrey et al. 1992; Ellner and Turchin 1995). The seminonparametric model is a recently proposed model, very different in form from FNN, which has been successful on economic data sets where linear models have proved inadequate.

The neural network model is equation (1) with

$$\begin{aligned} f(x_1, x_2, \dots, x_d, t) \\ = \beta_0 + \sum_{i=1}^k \beta_i G \left( \sum_{j=1}^d \gamma_{ij} x_j + \gamma_{i,d+1} \cos(2\pi t) \right. \\ \left. + \gamma_{i,d+2} \sin(2\pi t) + \mu_i \right), \end{aligned} \tag{3}$$

where  $G$  is a sigmoid function such as  $G(u) = \exp(u)/(1 + \exp(u))$ . In the original interpretation of (3) as a neural model, each term in parentheses on the right-hand side of (3) represents the stimuli to a single neuron, and  $G(\cdot)$  is the resulting firing rate of that neuron. However, equation (3) was used here strictly as a statistical model, with each  $x$  on the right-hand side being the log-transformed case report total from a past month. The cos and sin terms allow the model to include periodic variation in the contact rate intensity with period 1 yr; models omitting these seasonal covariates performed far less well. Software is available for fitting this model and for computing confidence intervals using methods developed elsewhere (Bailey et al. 1997), as part of the FUNFITS nonparametric regression package (Nychka et al. 1996).

The SNP model uses a series expansion to approximate the full one-step-ahead probability density of the time series conditional on past values, so it does not have the form of equation (1). The leading term is a linear autoregression with Gaussian noise. Higher-order terms result from polynomials multiplying the leading term's density and modifying its shape, in which the coefficients are themselves polynomial functions of current and past system states. A complete description of this model is given elsewhere (Gallant and Tauchen 1992). Seasonality was incorporated into the SNP model by including times

$t - 12$  mo and  $t - 24$  mo in the vector of past values used in making predictions at time  $t$  (see eq. [1]). Model fitting software for SNP is available on request from A. R. Gallant.

### *Semimechanistic*

The semimechanistic model was designed to incorporate the known compartmental structure of disease transmission embodied by figure 1, without specifying the form of the contact rate equation that gives the rate at which new cases arise. For convenience in fitting to monthly case report data, the model was formulated in discrete time:

$$S_{t+1} = S_t - C_{t+1} + r_{t+1},$$

and

$$x_{t+1} = g(S_t, x_t) + e_t. \quad (4)$$

Here  $C_t$  and  $x_t$  are case totals for month  $t$  (untransformed and log-transformed, respectively),  $S_t$  is an estimate of the number of susceptibles at the end of month  $t$ ,  $r_t$  is the total net recruitment into the susceptible population in month  $t$ , and  $e_t$  is random error.

The first line in (4) has exactly the same interpretation as the first line in (2). Both are mass-balance equations for the number of susceptibles, accounting for the net effect of individuals recruiting into the susceptible class and of individuals leaving the susceptible class when they catch the disease. To be exact, the first line of (4) really should read  $S_{t+1} = S_t - E_{t+1} + r_{t+1}$ , where  $E_t$  is the number of individuals catching the disease in month  $t$ . However, because the latent period for measles (roughly 1 wk) is short compared with the sampling interval (1 mo), we can assume that  $C_t \doteq E_t$  and thereby avoid using  $E_t$  as an additional state variable (Bobashev 1997; G. Bobashev, S. Ellner, D. W. Nychka, and B. T. Grenfell, unpublished manuscript). We initially allowed additional lags, replacing the second line with  $x_{t+1} = g(S_t, x_t, x_{t-L}, x_{t-2L}, \dots) + e_t$ , but these models performed less well than (4).

The main substantive difference between (2) and (4) is that the number of new cases is not assumed to follow a particular parametric rate equation ( $b(t)SI$  in the SEIR model) but is fitted by a phenomenological regression model  $g$ . Otherwise the models are functionally equivalent and operate at the same level of detail (e.g., both ignore age and spatial structure within the population). However, the fitted rate equation in (4), because it can take any functional form mandated by the data, can in principle reflect the net effect of processes and heterogeneities that are not explicitly included in the model.

We used two different forms for  $g$  in fitting equation

(4) to data: the neural network model (3), which we call the SC\_FNN model (SC for susceptibles and cases), and a kernel regression model (Härdle 1990; Cheng and Tong 1992), SC\_kernel. A brief description of kernel regression is given in appendix A. Kernel regression models can be fitted orders of magnitude faster than neural network models. Kernel regression was therefore used when estimating forecasting accuracy by cross validation (see app. A), which requires repeatedly refitting the model to different data sets. Cross validation on data sets of the size analyzed here is computationally infeasible at present for the neural network model.

### Methods for Evaluating the Models

Our comparisons of alternative models are based on their forecasting accuracies when applied to data not used in fitting the model, exactly as in Tidd et al. (1993). The data were monthly case report totals for measles in five large cities (fig. 2). Forecasts forward from each time  $t$  are based on a vector of log-transformed past values,  $\mathbf{X}_t = (x_t, x_{t-L}, \dots, x_{t-mL})$ , and on time of year ( $t$ ) to account for the strong seasonal variation in contact rate. Log-transformation was necessary to stabilize variances. Details of the methods for fitting the models, selecting an optimal state vector  $\mathbf{X}_t$ , and generating the forecasts are given in appendixes B and C.

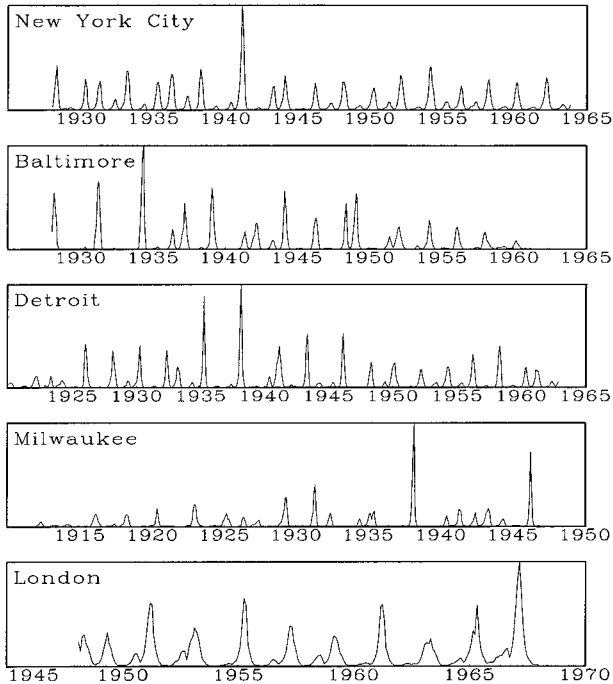
For each time  $t$  and for prediction intervals  $T_p = 1-24$  mo ahead, each model  $M$  was used to produce a forecast  $\hat{x}_{t+T_p}^M$  of  $x_{t+T_p}$ , which was the conditional mean of  $x_{t+T_p}$  given  $\mathbf{X}_t$  in that model. Only the first half of each data series was used in fitting the models, and predictions were made only for the second half of each data series. Prediction accuracy was measured by prediction  $r^2$ , which is computed by the usual formula for  $r^2$  in a regression analysis:  $1 - (\text{mean square error})/(\text{variance of data})$ . In this nonregression case,  $r^2$  may be negative if the model's forecasts are highly inaccurate. A negative prediction  $r^2$  means that model-based predictions are less effective than simply taking the unconditional mean value as the prediction would be.

As a baseline for evaluating the performance of these models, we used a prediction method that uses only the average trend over the year: the forecast for time  $k$  is simply the mean of all values from the same month in the first half of the data series. We refer to this as the *annual trend forecast*.

## Results

### *Comparisons of Forecasting Accuracy*

Forecasts based on SEIR were slightly less accurate overall than forecasts based on the phenomenological time-



**Figure 2:** The data used to evaluate forecasting accuracy: measles monthly case report totals for New York City, Baltimore county, Detroit, Milwaukee, and London. New York City data are from London and Yorke (1973), London data are from the Registrar General's Weekly Report; Baltimore, Detroit, and Milwaukee compiled and provided by William M. Schaffer. The time span for each series was chosen to terminate before vaccination (which began in the 1960s) had any evident effect and to avoid any obvious causes of nonstationarity. The data for London therefore do not include the years of the Second World War, during which many children were evacuated. For U.S. cities those years are included; virtually all cases occur before the age of 15, so the absence of those old enough for military service would not greatly affect disease transmission.

series models representing the “noisy nonlinearity” hypothesis, namely, the SNP model, and the FNN model (fig. 3). Semiparametric (SNP) forecasts were markedly better than SEIR for London and Milwaukee and nearly identical to SEIR in prediction accuracy for the other cities. Feed-forward neural networks forecasts (FNN) were markedly better than SEIR for Milwaukee but were only slightly better on average than SEIR for the other four cities.

The semimechanistic SC\_FNN model was better overall than SEIR in each of the cities (fig. 4) and had a higher prediction  $r^2$  than SEIR in 89% of the (city  $\times$  prediction interval) combinations examined. Comparing all four models, the highest average prediction  $r^2$  was achieved by SC\_FNN in three cities (New York, Detroit, Milwaukee), FNN in one city (Baltimore), and SNP in one city (London). The SC\_FNN model outperformed

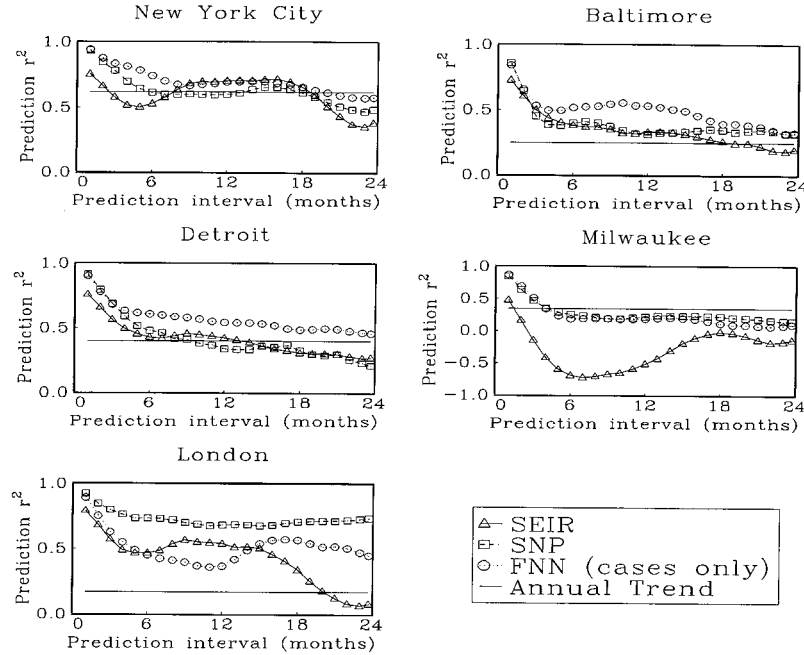
FNN on average (mean prediction  $r^2 = 0.61$  for SC\_FNN, 0.53 for FNN) and in 71% of the (city  $\times$  prediction interval) combinations: the semimechanistic model “beats” the purely phenomenological model. Both SNP and SEIR had lower average prediction  $r^2$  than either of the neural-net based models (mean prediction  $r^2 = 0.52$  for SNP, 0.32 for SEIR), supporting our earlier conclusions that neural network regression is well suited for population time series with larger sample sizes.

Forecasting based on SEIR has been reported to be more successful at “peak-to-peak” prediction, in which the next outbreak maximum is predicted based on the most recent outbreak maximum and intermediate months are ignored (Tidd et al. 1993; Kendall et al. 1994). We therefore compared the peak-to-peak prediction accuracies of the SEIR and SC\_FNN models. Forecasts were generated by tabulating the time series of outbreak maxima in long simulations of the model and using nonlinear regression (function *smooth.spline* in S-plus, Chambers and Hastie 1992; StatSci 1995) to estimate the relationship between successive maxima. The semimechanistic SC\_FNN model gave consistently more accurate forecasts than SEIR did (table 2). In contrast with SEIR, the peak-to-peak forecasts from SC\_FNN were less accurate than the 1-yr-ahead forecasts discussed above that were obtained by iterating forward 1 mo at a time.

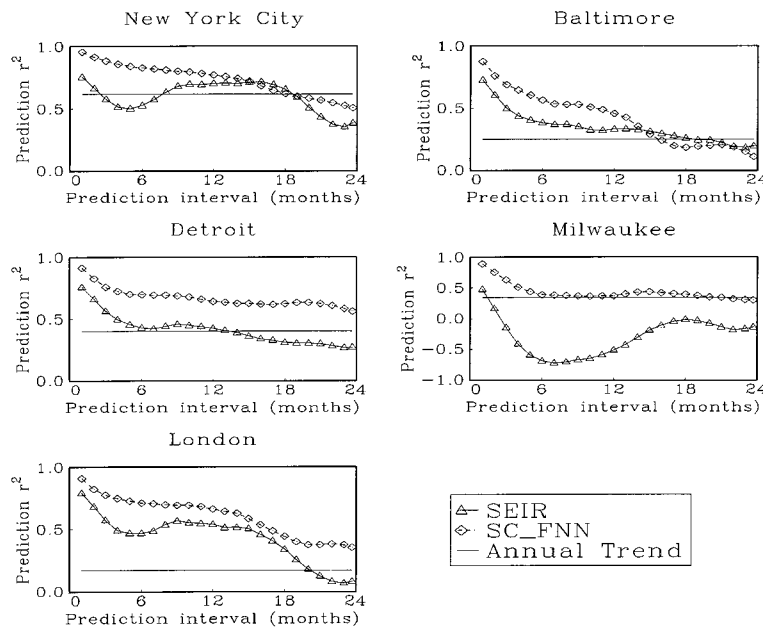
#### Using the Fitted Model

The comparison of forecasting accuracy indicates that the semimechanistic model SC\_FNN provides the best description of the dynamics. The effort of constructing and fitting this model is repaid by information that it provides about the dynamics. In this section we present three such examples. First, the form of the fitted model may suggest underlying mechanisms, or narrow the range of possible mechanisms, so that the semimechanistic model is a step toward obtaining an improved mechanistic model. Second, the fitted model can be used to characterize the significant features of the dynamics. Some such features of interest for population and epidemic dynamics are the overall extent of chaos versus stability, the level of predictability versus unpredictable “noise” in the dynamics, and how both of these vary as a function of the system's current state. Finally, the fitted model can be used to evaluate a proposed mechanistic model, or compare alternative proposed models, by fitting the same model to output from the mechanistic model and comparing with the model fitted to the data.

*Improving the Mechanistic Model.* There are only two mechanistic differences between the SC\_FNN and SEIR models: the former has a general contact rate equation



**Figure 3:** Comparison of prediction accuracy between the mechanistic SEIR model and the phenomenological SNP and FNN models, for model-based forecasts 1–24 mo ahead. Forecasts were made for the second half of the data series for each city, with only the first half of the data series used for fitting the models. Forecasting accuracy was measured by prediction  $r^2$ , the proportion of variance accounted for by the fitted model. Methods for fitting the models, and generating the forecasts, are described in appendixes B and C, respectively. The annual trend forecast (*solid line*) is the mean of all values from the same month in the first half of the data series.



**Figure 4:** Comparison of prediction accuracy between the mechanistic SEIR model and the semimechanistic SC\_FNN model for model-based forecasts 1–24 mo ahead. The curve for the SEIR-based forecasts is the same as that in figure 3. Methods are as described in the text and figure 3 legend.

**Table 2:** Prediction  $r^2$  for peak-to-peak prediction by the SEIR model and the semimechanistic neural net model SC\_FNN

	<i>New York</i>	<i>Baltimore</i>	<i>Detroit</i>	<i>Milwaukee</i>	<i>London</i>
SC_FNN	.51	.43	.47	-.11	-.05
SEIR	.43	.17	.19	-.76	-.09

Note: See text for definition and description of SC\_FNN and SEIR.

instead of assuming a priori the mass-action equation  $b(t)SI$ , and it also takes into account nonconstant recruitment rates. The better performance by SC\_FNN suggests that an improved mechanistic model could be obtained by adding one or both of these features to the SEIR model. It is easy enough to modify the SEIR model (2) so that the recruitment rate is nonconstant. Modifying the contact rate equation is less straightforward, because the fitted rate equation in SC\_FNN has no mechanistic meaning. We need to replace it with a simple functional form whose parameters can be interpreted in terms of underlying processes, without sacrificing accuracy.

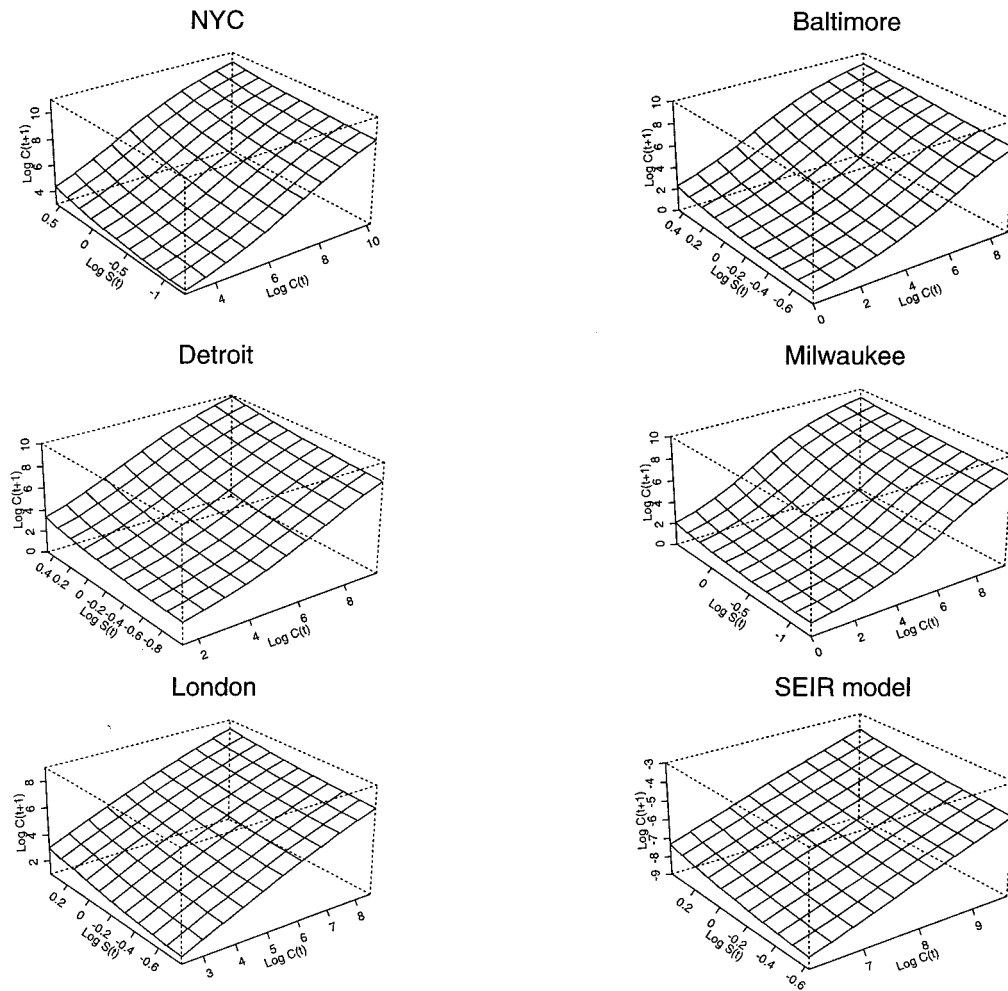
Simply plotting the fitted rate equations provides some information about the contact rate (fig. 5). For any month, the fitted equation is a response surface giving the expected number of new cases as a function of the current number of cases and susceptibles. If the contact rate followed the mass-action equation  $b(t)SI$ , then on log-scale the fitted surfaces should approximate the plane  $\log(C(t+1)) = A_0 + A_1 \log(C(t)) + A_2 \log(S(t))$ , with  $A_1 = A_2 = 1$ . The surface's nonlinearity should be weak, resulting only from the data being discrete-time running averages from the underlying continuous-time process. Thus when our fitting procedure is applied to output from the SEIR model, the estimated surface is very close to linear and the least-squares approximating plane has slopes  $A_0 = 0.98$ ,  $A_1 = 0.96$ . (These values are for the January surface shown in fig. 5; results for all other months are nearly identical.) The estimated surfaces in figure 5 appear to be linear or nearly so in  $\log S(t)$  but are markedly nonlinear in  $\log C(t)$ .

In contrast, the fitted seasonal periodicity is quite close to that in the SEIR model (eq. [2b]). Our model does not have an explicit seasonality term, but the seasonal dependence can be extracted by plotting  $\log(\text{predicted cases at time } t+1)$  as a function of time  $t$ , with the current  $C$  and  $S$  held constant. The resulting plots are very similar to equation (2b), the only consistent difference being a slightly broader trough of lowered transmission over the summer.

These results suggest a contact rate of the form  $b(t)Sf(I)$ , where  $f$  is nonlinear. Another study (Bobashev

1997) examined the adequacy of this form for the contact rate, using weekly case report data on six large cities in Britain. Nonparametric (spline) estimates of  $b(t)$  and  $f(I)$  were obtained for each city (see Bobashev 1997 for details of the methods), and this nonparametric model gave a statistically better fit to the data than the mass-action equation (based on  $F$ -tests for significance of the added degrees of freedom). The parametric form  $f(I) = I + I_0$ ,  $I_0$  a small positive constant, was found to describe the data at low values of  $I$  typical of the troughs between epidemics. Presumably  $I_0$  represents cases arising due to contact with infectives who live in other cities and are counted as cases in their home city. At high values of  $I$  there is some saturation in the number of additional cases per additional infective (Bobashev 1997; this is also visible in fig. 5). Possible mechanistic explanations for the saturation include spatial clustering of infectives, that is, when there are many infectives, most susceptibles are in areas where infectives happen to be rare (see Rhodes and Anderson 1996 for how this can affect the dynamics) and heterogeneity of susceptibility (more-susceptible individuals are infected early, so as the epidemic grows the transmission rate decreases). Each of these could be evaluated by constructing a model incorporating the mechanism and seeing if the contact rate surface estimated from model output matches those derived from the data. However, this level of mechanistic understanding is not necessary for creating a more accurate model at the same level of detail as the SEIR equations; all we need is the more accurate contact rate equation estimated from the data.

*Stability versus Chaos, and Local Chaos.* Stability versus chaos can be quantified by the dominant Lyapunov exponent  $\lambda$ . The Lyapunov exponent gives the long-term sensitivity to initial conditions, that is, the rate of growth (or decay) over time of the effect of a small perturbation in the system's state. Thus  $\lambda > 0$  indicates sensitive dependence on initial conditions, which is the classical defining feature of chaotic dynamics. The mathematical definition of  $\lambda$  for systems with random perturbations, and methods for estimating  $\lambda$  in such cases, are reviewed elsewhere (Ellner and Turchin 1995).



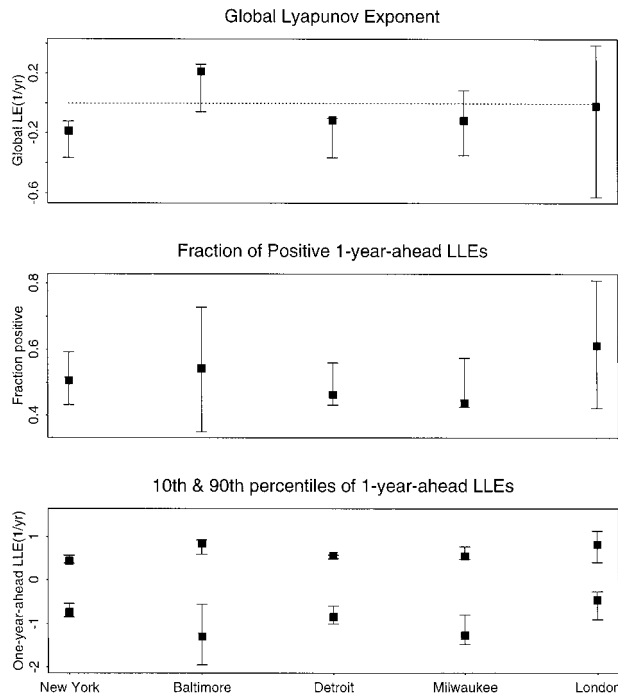
**Figure 5:** Plots of the fitted rate equations in the SC\_FNN model, showing on logarithmic scale the number of new cases,  $C(t + 1)$ , as a function of the current number of cases,  $C(t)$ , and number of susceptibles  $S(t)$ . The model generates different surfaces for each month. Those shown are for January, but those for other months are qualitatively similar. For comparison, the bottom right panel shows the plot obtained by applying exactly the same fitting and plotting procedures to a times series of cases generated by the SEIR model with mass-action contact rate  $\beta SI$ ; the relatively weak nonlinearity in this surface results from the fitted model being in discrete rather than continuous time. The difference in scales for  $\log C(t)$  versus  $\log S(t)$  produces the visual impression that the contact rate is insensitive to changes in  $S(t)$ ; in fact, for the least-squares approximating plane to the SEIR model surface, the slopes in  $\log C(t)$  and  $\log S(t)$  are nearly identical.

To estimate  $\lambda$ , we refitted the SC\_FNN model using the entire data series but retaining the model specification that gave the best prediction  $r^2$ . The estimates of  $\lambda$  are shown in figure 6 (top panel), with 95% confidence intervals obtained by methods developed in another paper (Bailey et al. 1997). These estimates support our previous conclusion (Ellner et al. 1995) that measles epidemic dynamics are clustered near the transition between stable and chaotic dynamics ( $\lambda \approx 0$ ). They are neither strongly stable nor strongly chaotic, and for three of the five cities, the 95% confidence interval for  $\lambda$  contains

zero. Estimates from the semimechanistic model were far less variable (over cities) than those obtained in an earlier study (Ellner et al. 1995) with the phenomenological FNN model (fig. 7). This supports Tidd et al.'s (1993) criticism of phenomenological models: by ignoring biological information they needlessly sacrifice precision.

Because the dynamics are neither strongly stable nor strongly chaotic, global measures such as  $\lambda$  are less informative than short-term measures that characterize the "local" dynamics in different regions of state space (Yao and Tong 1995). Local chaos versus stability can be char-

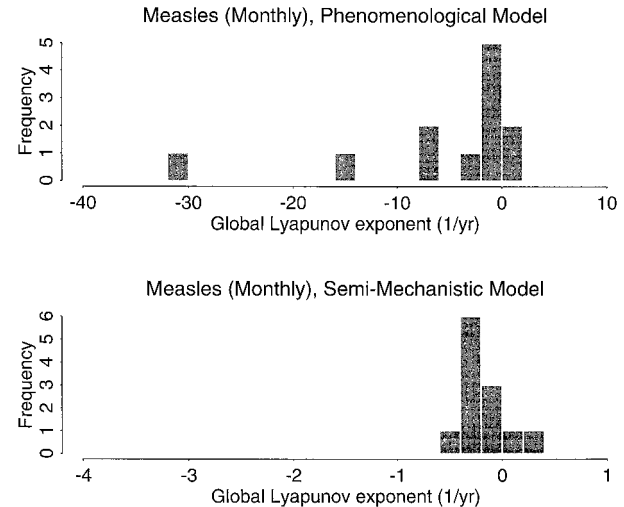




**Figure 6:** Estimates of global and local Lyapunov exponents for each of the cities. Estimated values are indicated by the solid box, and the error bars show 95% confidence intervals computed by those described elsewhere (Bailey et al. 1997). The top panel shows the estimated global Lyapunov exponent  $\lambda$ , with the dashed line at 0 indicating the boundary between stable ( $\lambda < 0$ ) and chaotic ( $\lambda > 0$ ) dynamics. The two lower panels refer to the local Lyapunov exponents  $\lambda_m(t)$  for  $m = 12$  mo ahead, which gives the short-term rate of growth ( $\lambda_m(t) > 0$ ) or decrease ( $\lambda_m(t) < 0$ ) in the effect of a perturbation at time  $t$ . The middle panel shows 95% confidence intervals for the fraction of  $\lambda_m(t)$  values that are positive; the bottom shows 95% confidence intervals for the tenth and ninetieth percentiles of the distribution of  $\lambda_m(t)$  over the course of the data series.

acterized by the local Lyapunov exponents, denoted  $\lambda_m(t)$  (e.g., Abarbanel et al. 1991, 1992; Wolff 1992; Bailey 1996). The exponent  $\lambda_m(t)$  is the short-term rate of growth ( $\lambda_m(t) > 0$ ) or decrease ( $\lambda_m(t) < 0$ ) in the effect of a perturbation at time  $t$ , over the time interval from  $t$  to  $(t + m)$ . If the system is deterministic,  $\lambda_m(t)$  depends only on the system state at time  $t$ ; in noisy systems it is a random quantity that depends on the system trajectory between times  $t$  and  $(t + m)$ .

In each of the cities, there was substantial variation in the local Lyapunov exponents over the course of epidemics, including a roughly 50:50 split between positive and negative values (fig. 6). The variation in local exponents within each city (as indicated by the gap between the



**Figure 7:** Histograms of estimated global Lyapunov exponents for measles epidemics in 12 cities, those shown in figure 2 plus Copenhagen, Bristol, Liverpool, Manchester, Newcastle, Birmingham, and Sheffield. Upper panel shows estimates obtained from the phenomenological time-series model (eq. [1]), taken from an earlier study (Ellner et al. 1995). Lower panel shows estimates obtained from the semimechanistic SC\_FNN model (eq. [4]). Note the tenfold difference in horizontal scale between the panels.

tenth and ninetieth percentiles of the distribution) is far larger than the variation between cities in the global exponent. Thus in all cases, the near-zero global Lyapunov exponent  $\lambda$  results from an alternation between short-term sensitive and insensitive dependence on initial conditions.

So is measles chaotic? Yes, and no. Alternating between local stability and local chaos causes the dynamics to look and act chaotic much of the time, even if strictly speaking they are not. This kind of behavior is not the same as noise-induced chaos (Rand and Wilson 1991), in which random perturbations cause the system to be genuinely chaotic in the strict sense. Nonetheless, during the periods of local chaos, the system has strong short-term sensitivity to any perturbations it receives, exactly as in a truly chaotic system. Small causes “now” have larger effects “a bit later,” even though much later the effects are damped out during the periods of local stability. This kind of quasi-chaotic behavior has received little attention in the literature on deterministic chaos because without continual random perturbations it is only the long-term insensitivity to perturbations that matters. With random perturbations, every time the system enters a region of local chaos, the potentially erratic behavior gets kicked into action immediately. Local chaos is therefore the more relevant property for systems subject to re-

current random perturbations, and it appears to be the rule in measles epidemics.

*Evaluating Mechanistic Models: Unpredictability as a Probe.* The prediction errors of the fitted semimechanistic model are due in part to the intrinsic unpredictability of the dynamics, but errors are also produced by other factors, such as limited sample size and reporting error. Nonetheless, we can use prediction accuracy as a “probe” into the adequacy of a proposed mechanistic model. We do this by comparing the apparent unpredictability of the real data with that of equally long simulated data sets from the proposed model that are also contaminated by (simulated) reporting errors.

As a simple example, we consider the SEIR model (eq. [2]) implemented as a finite-population simulation. The finite-population model treats the right-hand sides of the differential equation as specifying stochastic transition rates between the compartments shown in figure 1. The model was implemented by using a 2-h time step and Poisson-distributed transitions with mean given by transition rate  $\times$  time step. A small rate of immigration by infectives was added to keep the disease from fading out completely by chance. The immigration rate was adjusted up from zero until the lower tail of the distribution of monthly case totals roughly matched that for New York City, corrected for underreporting. A natural model of the reporting process is independent random sampling, with each case having a probability  $P < 1$  of being reported. However, a better fit to the power spectrum of the data is obtained by assuming that cases are nonindependent and, in particular, that they are reported in clusters of size 2 or 3 (Ellner et al. 1995). Here we used a cluster size of 2 and reporting probability  $P$  chosen so that the mean number of cases in the simulation matched that in the data. Prediction  $r^2$  was computed for the real time series and for model output, in both cases by fitting the semimechanistic model to the time series of monthly case totals. The contact rate ( $g$  in eq. [4]) was estimated by kernel regression because this allows us to use the entire time series (rather than just the second half) to estimate unpredictability by cross validation (see the section “Semimechanistic”).

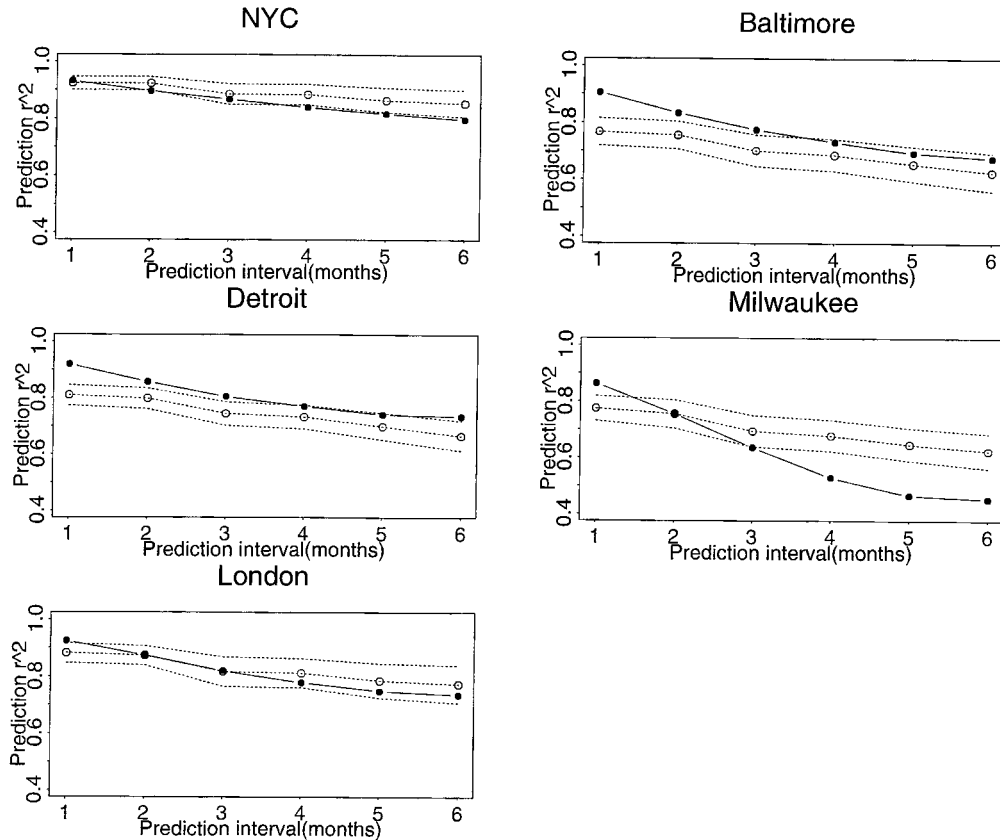
The results (fig. 8) show a good but imperfect match between model and data—indeed surprisingly good, given that model parameters were not adjusted to achieve a close match. However, the 1-mo-ahead predictability of the data was consistently higher than that of the simulation model (five of six cities). Given that the model omits numerous complexities of the real disease, this discrepancy is more likely due to the reporting process than to the real world being less stochastic than the model. For example, if a fraction of cases are reported in the month

after they actually occur, this would smooth out the data and increase the apparent predictability. Milwaukee is clearly identified as an “outlier,” with predictability falling off far more rapidly than the model predicts. The major outbreaks in 1938 and 1945 (see fig. 1) are the likely culprits. In those instances, 2 yr of small outbreaks were followed by an especially large outbreak, whereas previously that had not been true. On the whole, the results in figure 8 indicate that the unpredictability in the data can be accounted for by demographic stochasticity interacting with nonlinear dynamics (in particular, the “noise amplification” [Deissler and Farmer 1992] that occurs during periods of short-term chaos). This accords with the conclusions stated above based on the semi-mechanistic model.

## Discussion

The main biological implication of our analyses is that measles epidemics are best described as a mixture of nonlinearity and noise, with neither component being small enough to disregard. The dynamics are not essentially deterministic, nor are they just random measurement errors masking a simple seasonal trend. Epidemics are likely to be perturbed by both environmental variability and demographic stochasticity (Bartlett 1957; Grenfell et al. 1995c). The biological details omitted from our models, such as age structure and spatial heterogeneity, may also contribute to forecasting errors. However, the observed levels of predictability are quite high for at least 6 mo into the future. Moreover, the levels of unpredictability are reasonably well matched by a model in which the only stochastic forces are demographic stochasticity and random reporting errors, both of which are undoubtedly present in reality.

The estimates of the local and global Lyapunov exponents (figs. 6 and 7) indicate that the epidemic dynamics are clustered very near the border between stable ( $\lambda < 0$ ) and unstable ( $\lambda > 0$ ) dynamics, but there are large fluctuations in the degree of short-term chaos versus stability (i.e., short-term sensitivity vs. insensitivity to initial conditions). These findings imply that interactions between noise and nonlinearity are a significant aspect of the dynamics. During periods of short-term sensitivity to initial conditions ( $\lambda_m(t) > 0$ ) the system is acting as a “noise amplifier,” in which the nonlinear disease transmission dynamics amplify the effect of demographic stochasticity and/or exogenous random shocks. During these periods, much of the unpredictability is generated by endogenous nonlinearity, as in a chaotic system. When  $\lambda_m(t) < 0$ , there is insensitivity to initial conditions, as in a stable system with random perturbations. Characterizing this kind of dynamics as chaotic (if  $\lambda > 0$ ) or nonchaotic (if



**Figure 8:** Comparison of apparent unpredictability of the data with that of output from the finite-population SEIR model with simulated reporting errors. Unpredictability is  $1 - r^2$  for the semimechanistic model with kernel regression estimate of the contact rate (see text for details); this can be interpreted as the proportion of variance unexplained by the model. Solid lines are the unpredictability's for the data; dashed lines are the mean  $\pm 2$  standard errors, from 25 simulated data sets equal in length to the real data. Total population size for the simulations was 5 million for New York (NYC), 2 million for London, and 1 million for the others.

$\lambda < 0$ ) might be correct, but it would be incomplete and misleading. The complexity of the dynamics is not simply “noise,” nor is it chaos in the classical sense but, rather, is a result of the interaction between noise and local nonlinearities. Similar qualitative behavior has been discovered recently in several other cyclic populations: pacific Dungeness crabs (Higgins et al. 1997) and Fennoscandian voles (P. Turchin and S. P. Ellner, unpublished manuscript).

Our results also have more general implications for the study of dynamics in other ecological and epidemiological systems. Our results support Tidd et al.'s (1993) argument that analyses based on phenomenological time-series models omit useful information and therefore sacrifice precision needlessly. However, contrary to Tidd et al., our results do not support analyses based on a simple mechanistic model such as SEIR. The main flaw in the Tidd et al. study was the use of linear models as the benchmark for comparison with the mechanistic model.

A linear model is not a fair “straw man” because it is far from the best phenomenological model available. A nonlinear autoregressive model such as a neural net model or a completely nonparametric model such as SNP would be more appropriate benchmarks.

The most accurate description of the dynamics, as measured by out-of-sample forecasting accuracy, was obtained with a semimechanistic model that retains the compartmental structure of the SEIR model but uses the data to estimate the form of some rate equations. The main difference between the SEIR and semimechanistic models is that the former uses the conventional “strong homogeneous mixing” contact rate equation (Anderson and May 1991), in which the infection rate is proportional to the product of the number of susceptibles and the number of infectives. The SEIR and semimechanistic models operate at exactly the same level of detail—city-wide monthly case totals. Thus the difference in performance is not simply a matter of one model incorporating

more detail but indicates that the conventional contact rate equation may be incorrect, which might explain some of the persistent discrepancies between the data and SEIR-type models (Bolker and Grenfell 1993). Alternatively, the greater flexibility of the semimechanistic model may allow it to capture the main effects of some biological "details" that are omitted from both models but can have significant effects on the dynamics: age structure (Schenzle 1984; Anderson and May 1991; Bolker and Grenfell 1993), changes in the clustering of infectives (Rhodes and Anderson 1996), and variation in susceptibility among individuals. Thus we cannot infer that the semimechanistic model is "correct," simply that it does the best job (so far) of capturing the dynamic structure present in the available time-series data. It is also not yet clear to what extent the dynamics are influenced by external demographic factors such as secular trends in birth rates (Grenfell et al. 1995a, 1995b, 1995c). Of the models considered here, only the semimechanistic model allows for variation in birth rates.

This case study has exploited the length and accuracy of epidemic data and our knowledge of underlying epidemiological processes and parameter values. Can a semi-mechanistic approach be useful in practice, even with poorer data or less knowledge? Several case studies suggest that the answer is sometimes yes. One such study (Ellner et al. 1997) considered the problem of estimating differential delay equations of the sort that arise from age- or stage-structured population models in continuous time (see, e.g., Gurney et al. 1983; McCauley et al. 1996 and references therein). The simplest example of such models is

$$\frac{dx}{dt} = f(x(t - \tau)) + g(x(t)).$$

This can be interpreted as a delayed-recruitment model, with  $x(t)$  the number of adults at time  $t$ ,  $\tau$  the maturation time, and the  $f()$  and  $g()$  terms representing recruitment into the adult class and deaths of adults, respectively. We assumed that the form of the equation was known (this is the mechanistic information analogous to eq. [4] above), but that  $\tau$ ,  $f$ , and  $g$  had to be estimated from a single time series of  $x(t)$  values. As few as 100 data points were sufficient to determine the general shape of both  $f$  and  $g$  and to estimate the maturation time  $\tau$  very accurately, even if the sampling frequency was so low that  $x(t)$  and  $x(t - \tau)$  are never observed simultaneously (e.g., if  $\tau = 17$  d and samples are taken every 3 d).

One of us recently compared (S. P. Ellner, unpublished data) phenomenological, mechanistic, and semi-mechanistic models for laboratory populations of the

blowfly *Lucilia sericata* (R. Smith, unpublished data); details of this analysis will appear elsewhere. The data were 200 every-other-day samples of larvae, pupae, and total adults (mature plus immature), with three replicates at each of two initial age distributions. Experimental data (Daniels et al. 1991; Simkiss et al. 1993) were available to parameterize all rate equations for a mechanistic stage-structured model (similar to Gurney et al. 1983). This mechanistic model exhibited the strong periodicity observed in the data but with the wrong period unless all noise in the model was turned off completely, including demographic stochasticity. The phenomenological model was equation (1) fitted to the total number of adults; the output from this model was noncyclic. The semimechanistic model adopted the stage structure from the mechanistic model (eggs, larvae, pupae, immature, and mature adults) but estimated all transition rates from the time series using an estimate of the (unobserved) number of mature adults. This model cycled with the correct period and roughly correct amplitudes. Moreover, it successfully predicted the qualitative changes in dynamics observed in a second experiment where the pupal mortality was higher due to a low dose of cadmium in the larval diet. The difference between the semimechanistic and mechanistic models was informative about underlying mechanisms. The former omitted as nonsignificant one of the regulating mechanisms observed in the independent experiments, a strong effect of larval crowding on the subsequent size and fecundity of those individuals as adults. Thus the density-dependent variation in adult size was apparently not a significant regulating factor in the observed population cycles.

The polar alternatives of fitting simple mechanistic models, and purely descriptive statistical models, have been the dominant approaches for modeling and quantifying population fluctuations. While it is recognized that both approaches have limitations, both are widely used. Our results suggest that a combination of mechanistic and statistical population modeling, reflecting the actual state of knowledge about the system, can be useful in practice for improved forecasting and characterization of population dynamics. Moreover, the fitted model can provide information on the mechanisms driving the observed dynamics. By applying this approach to measles, we have obtained a resolution to the longstanding debate on whether the dynamics are noise or chaos; we have presented quantitative evidence that our analysis is based on models with superior average performance to those used previously, across a range of representative data sets; and we have uncovered evidence for consistent departures from the conventional mass-action equation of disease transmission that can be incorporated into improved mechanistic models.

### Acknowledgments

We thank our volunteer manuscript reviewers (B. Kendall, N. Pack, P. Schliekelman, and J. Trexler), the referees (A. Hastings, G. Ruxton, and two anonymous), and the editor and associate editor for the time and thought that they put into improving the manuscript. We also thank B. Bolker, M. Easterling, B. Kendall, A. Kleczkowski, S. Peck, and many seminar attendees for helpful discussions. This research was supported in part by National Science Foundation grants DMS-9217866 to D.W.N. and S.P.E. and SBR-9514189 to A.R.G., an Industrial Traineeship from Becton Dickinson to B.A.B., and the Isaac Newton Institute for Mathematical Sciences; B.T.G. thanks the Wellcome Trust and the Royal Society for support.

## APPENDIX A

### Kernel Regression

Consider a data set  $\{(X_i, y_i), i = 1 \dots n\}$ , where the  $X$ 's are observations of the independent variables and the  $y$ 's are the corresponding observations of a single dependent variable. Kernel regression is a general nonparametric regression method based on local weighted averaging of the data, which provides a consistent estimate of the conditional mean of  $y$  given  $X$  (Härdle 1990; Cheng and Tong 1992). In this article, the  $X$ 's are values of the system state variables—current and lagged case totals, susceptibles, and/or seasonal covariates—and the  $y$ 's are the number of cases at a future time. The fitted kernel regression model thus provides forecasts of future cases. The degree of predictability in the data can then be estimated from the residuals, much as the residual mean square can be used to estimate the error variance in linear regression.

The predicted  $y$  value for any  $X$  (which need not be one of the data values) is a weighted average of the  $y_i$  in which most weight is given to those with  $X_i$  near  $X$ ; the general formula for a kernel regression model is

$$\hat{y} = \frac{\sum_i w_i(X) y_i}{\sum_i w_i(X)}, \quad (\text{A1})$$

where  $\hat{y}$  is the predicted value of  $y$  (i.e., the conditional mean of  $y$  given  $X$ ). The weights  $w_i$  are given by  $k(d_i/h)$ , where  $d_i$  is the distance from  $X$  to  $X_i$ ,  $k$  is a weighting function (the “kernel”) that falls off with distance, and  $h$  (called the *bandwidth*) is a constant that controls the rate at which more distant points are down-weighted. We used  $k(d) = 1/(1 + d + 0.5d^4 + 0.3d^6)$ , which falls off roughly as  $\exp(-d^2)$  but is faster to compute.

The bandwidth  $h$  was chosen by cross validation to minimize the mean square prediction error. Cross valida-

tion is a process to estimate the prediction accuracy of (A1) as a function of the bandwidth; it is akin to fitting a model to one data set and testing its accuracy on independent data (e.g., by splitting the available data into two halves). However, cross validation is more efficient because all data points are used in both the fitting and testing steps. Given a trial value of  $h$ , equation (A1) is applied repeatedly to reduced data sets, each obtained by deleting a window of temporal neighbors centered at one data point  $X_j$ . Equation (A1) then produces a prediction of  $y_j$  based on this reduced data set, which is compared with the actual  $y_j$  in the data set. The squared prediction errors are summed over all  $j$  to produce the overall measure of prediction error for the given value of  $h$ . Numerical function minimization is then used to find the value of  $h$  at which the overall prediction error is minimized.

In all applications of cross validation in this article, the window of temporal neighbors omitted from (A1) extended 24 mo on either side of the time for which predictions were made. A window of that size is necessary to avoid spurious results arising from the strong temporal autocorrelations in the data (Grenfell et al. 1995a).

## APPENDIX B

### Fitting the Models

The SNP, FNN, and SC\_FNN models were fitted for a prediction time of 1 mo ahead. For FNN and SC\_FNN, the parameters  $(\beta_i, \gamma_{ij}, \mu_i)$  were estimated by nonlinear least squares for each given value of  $k$ , and the value of  $k$  was chosen by the Bayes Information Criterion (BIC), using methods described elsewhere (Ellner et al. 1992; Nychka et al. 1992). After some trial-and-error using only the first half of each data series, the SNP model was constrained to use 24 past monthly values in the linear autoregression but only the four most recent values in the polynomials. Methods used to estimate parameter values and select the polynomial order for SNP are described elsewhere (Gallant and Tauchen 1992).

In kernel regression the only fitted parameter is the bandwidth. In all kernel regression fits we selected the bandwidth by cross validation as described in appendix A.

For SEIR, we initially followed Tidd et al. (1993) and implemented the model as a finite-population Monte Carlo simulation, with a small amount of immigration (as in Kendall et al. 1994) to prevent an unrealistically high number of months with zero cases. However, we found that substantially more accurate forecasts were obtained using the differential equations (2). We used the values of  $m$ ,  $a$ , and  $g$  (which can be estimated directly) and of  $\beta_0$  from Tidd et al. For each data series we com-

puted forecasting accuracy for  $\beta_1 = 0.20, 0.22, 0.24, 0.26,$  and  $0.28$ , which spans the range of empirical estimates and runs from simple periodic dynamics to chaos. Forecasting accuracy was improved, especially at lower values of  $\beta_1$ , by adding small random variations about the seasonal trend in contact rate intensity,  $b(t) = \beta_0(1 + \beta_1\varphi(t) + \sigma z(t))$ , where  $z(t)$  was a first-order Gaussian autoregressive process with unit variance and autocorrelation of 0.5 at time lag of 2 wk. For each value of  $\beta_1$  the value of  $\sigma$  was increased from 0 in increments of 0.01 until the model dynamics were qualitatively similar to the data (i.e., a mix of 2-yr-periodic, 3-periodic, and aperiodic dynamics). These values of  $\sigma$  are thus the minimal modification of the deterministic model necessary to generate realistic simulations. The largest value of  $\sigma$  was 0.03 at  $\beta_1 = 0.20$ . For each of the five cities, we report results only for the value of  $\beta_1$  that gave the highest average forecasting accuracy for 1–24 mo ahead using embedding parameters  $L = 3, D = 6$ :  $\beta_1 = 0.20$  for New York City and London,  $\beta_1 = 0.22$  for Detroit,  $\beta_1 = 0.24$  for Baltimore, and  $\beta_1 = 0.26$  for Milwaukee.

Values of  $S_t$  and  $r_t$  for the semimechanistic model (4) were estimated from case report data using methods described elsewhere (Bobashev 1997; G. Bobashev, S. P. Ellner, D. W. Nychka, and B. T. Grenfell, unpublished manuscript). The  $S_t$  estimate differs from the actual number of susceptibles by an unknown shift of location and scale;  $r_t$  differs from actual recruitment by the same shift of scale, which is approximately equal to the fraction of cases that are reported. The method guarantees that the first line of equation (4) is valid. For SC\_FNN, the parameters of  $g$  were estimated by least squares.

## APPENDIX C

### Computing Forecasts from the Models

For FNN, SC\_FNN, and SNP the fitted model gives forecasts 1 mo ahead. Forecasts farther ahead were obtained by simulating the model repeatedly and averaging over simulations to obtain the conditional mean. For SNP this involves repeated random draws from the fitted one-step-ahead transition density. For FNN we iterated forward equation (1) with  $T = 1$ , using random shocks  $e_t$  generated by random sampling with replacement from the residuals of the 1-mo-ahead model. Forecasts for SC\_FNN were also generated by simulating the model, but the recruitment rate  $r_t$  was treated as a known covariate, since it is an exogenous factor not predictable from epidemic data. Both  $x_t$  and  $S_t$  were forecast by iterating equation (4) with the fitted  $g$ .

Forecasts for the SEIR model were obtained by a refinement of the method in Tidd et al. (1993). We gener-

ated a 200-yr-long series of simulated case totals, which were then log-transformed and scaled in the same way as the real data, and converted into an “atlas” of simulated state vectors  $\{X_t^s, t = 0, 1, \dots\}$ . A kernel regression model was then fitted to these simulated data with bandwidth chosen by cross validation (app. A). Predictions are then obtained directly from equation (A1), with the “data”  $(X_i, y_i)$  being output from the simulation model, and  $X$  being the point in the real data for which a prediction is being made. Spot-checks with longer atlas lengths (up to 500 yr) indicated that only minuscule improvements result from an atlas longer than 200 yr.

The forecasting accuracy of each method depends on the choice of embedding parameters in the state vector, namely, the time delay  $L$ , and the number of lags  $D = m + 1$  (see eq. [1]). Tidd et al. (1993) used  $L = 1$  mo and  $D = 4$  or  $6$ , but Grenfell et al. (1994) found that  $L = 4$  greatly improved the forecasting accuracy on measles data for England and Wales. Data self-forecasting accuracy (estimated by cross validation; see app. A) was used as the criterion for choosing embedding parameters (as in Sugihara and May 1990), on the principle that good embedding parameters for the data should also be good for an accurate model. For our data sets,  $L = 3$  or  $4$  with  $D = 5$  or  $6$  gave much more accurate forecasts. In most of the cities, forecasts with  $L = 3$  or  $4$  were better, and forecasts with  $L = 1$  were worse, than forecasts based solely on the seasonal trend. The highest average forecasting accuracy was obtained at  $L = 3, D = 6$ , which in fact has been the most popular choice in past studies of these data. We therefore report results here using  $L = 3, D = 6$ , for the SEIR and data-atlas forecasts; all results were very similar for any of the better embedding parameters ( $L = 3$  or  $4, D = 5$  or  $6$ ). For the FNN model, the six-dimensional state vector consisted of four past values plus the two seasonal covariates (sin and cos terms).

### Literature Cited

- Abarbanel, H. D. I., R. Brown, and M. B. Kennel. 1991. Variation of Lyapunov exponents on a strange attractor. *Journal of Nonlinear Science* 1:175–199.
- . 1992. Local Lyapunov exponents computed from observed data. *Journal of Nonlinear Science* 2: 343–365.
- Anderson, R. M., and R. M. May. 1991. *Infectious diseases of humans: dynamics and control*. Oxford University Press, Oxford.
- Bailey, B. A. 1996. Local Lyapunov exponents: predictability depends on where you are. Pages 343–357 in W. Barnett, A. Kirman, and M. Salmond, eds. *Nonlinear dynamics in economics: proceedings of the Tenth International Symposium in Economic Theory and*

- Econometrics. Cambridge University Press, Cambridge.
- Bailey, B. A., S. Ellner, and D. W. Nychka. 1997. Chaos with confidence: asymptotics and applications of local Lyapunov exponents. Pages 115–134 in C. D. Cutler and D. T. Kaplan, eds. *Nonlinear dynamics and time series: building a bridge between the natural and statistical sciences*. American Mathematical Society, Providence, R.I.
- Bartlett, M. S. 1957. Measles periodicity and community size. *Journal of the Royal Statistical Society A, General* 120:48–70.
- Bobashev, G. 1997. Spatial modeling of measles epidemic dynamics. Ph.D. thesis. North Carolina State University, Raleigh.
- Bolker, B. M., and B. T. Grenfell. 1993. Chaos and biological complexity in measles dynamics. *Proceedings of the Royal Society of London B, Biological Sciences* 251:75–81.
- Chambers, J. M., and T. J. Hastie, eds. 1992. *Statistical models in S*. Wadsworth & Brooks/Cole, Pacific Grove, Calif.
- Cheng, B., and H. Tong. 1992. On consistent nonparametric order determination and chaos. *Journal of the Royal Statistical Society B, Methodological* 54:427–450.
- Costantino, R. F., J. M. Cushing, and R. A. Desharnais. 1995. Experimentally induced transitions in the dynamics behavior of insect populations. *Nature (London)* 375:227–229.
- Daniels, S., K. Simkiss, and R. H. Smith. 1991. A simple larval diet for population studies on the blowfly *Lucilia sericata* (Diptera: Calliphoridae). *Medical and Veterinary Entomology* 5:283–292.
- Deissler, R. J., and J. D. Farmer. 1992. Deterministic noise amplifiers. *Physica D* 55:155–165.
- Dennis, B., R. A. Desharnais, J. M. Cushing, and R. F. Costantino. 1995. Nonlinear demographic dynamics: mathematical models, statistical methods, and biological experiments. *Ecological Monographs* 65:261–281.
- Ellner, S. 1991. Detecting low-dimensional chaos in population dynamics data: a critical review. Pages 63–90 in J. Logan and F. Hain, eds. *Chaos and insect ecology*. Virginia Agricultural Experimental Station, Virginia Polytechnic Institute and State University, Charlottesville.
- Ellner, S., and P. Turchin. 1995. Chaos in a noisy world: new methods and evidence from time-series analysis. *American Naturalist* 145:343–375.
- Ellner S., D. W. Nychka, and A. R. Gallant. 1992. LENNS, a program to estimate the dominant Lyapunov exponent of noisy nonlinear systems from time series data. Institute of Statistics Mimeo Series no. 2235. Statistics Department, North Carolina State University, Raleigh.
- Ellner, S., A. R. Gallant, and J. Theiler. 1995. Detecting nonlinearity and chaos in epidemic data. Pages 229–247 in D. Mollison, ed. *Epidemic models: their structure and relation to data*. Cambridge University Press, Cambridge.
- Ellner, S. P., B. E. Kendall, S. M. Wood, E. McCauley, and C. J. Briggs. 1997. Inferring mechanism from time-series data: delay differential equations. *Physica D: Nonlinear Phenomena* 110:182–194.
- Gallant, A. R., and G. Tauchen. 1992. A nonparametric approach to nonlinear time series analysis: estimation and simulation. Pages 77–92 in D. Brillinger, P. Caines, J. Geweke, E. Parzen, M. Rosenblatt, and M. S. Taqqu, eds. *New directions in time series analysis*. Pt. 2. Springer, New York.
- Grenfell, B. T., A. Kleczkowski, S. Ellner, and B. M. Bolker. 1995a. Non-linear forecasting and chaos in ecology and epidemiology: measles as a case study. Pages 345–371 in H. Tong, ed. *Chaos and forecasting: proceedings of the Royal Society Discussion Meeting*. World Scientific, Singapore.
- Grenfell, B., B. M. Bolker, and A. Kleczkowski. 1995b. Seasonality and extinction in chaotic metapopulations. *Proceedings of the Royal Society of London B, Biological Sciences* 259:97–103.
- . 1995c. Seasonality, demography and the dynamics of measles in developed countries. Pages 248–268 in D. Mollison, ed. *Epidemic models: their structure and relation to data*. Cambridge University Press, Cambridge.
- Gurney, W. S. C., R. M. Nisbet, and J. H. Lawton. 1983. The systematic formulation of tractable single-species population models incorporating age structure. *Journal of Animal Ecology* 52:479–495.
- Hanski, I., P. Turchin, E. Korpimäki, and H. Henttonen. 1993. Population oscillations of boreal rodents: regulation by mustelid predators leads to chaos. *Nature (London)* 364:232–235.
- Härdle, W. 1990. *Applied nonparametric regression*. Cambridge University Press, Cambridge.
- Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray. 1993. Chaos in ecology: is Mother Nature a strange attractor? *Annual Reviews of Ecology and Systematics* 24:1–33.
- Higgins, K., A. Hastings, J. N. Sorvela, and L. W. Botsford. 1997. Stochastic dynamics and deterministic skeletons: population behavior of Dungeness crab. *Science (Washington, D.C.)* 276:1431–1435.
- Keeling, M. J., and B. T. Grenfell. 1997. Disease extinction and community size: modeling the persistence of measles. *Science (Washington, D.C.)* 275:65–67.

- Kendall, B. E., W. M. Schaffer, L. F. Olsen, C. W. Tidd, and B. L. Jorgensen. 1994. Using chaos to understand biological dynamics. Pages 184–203 in J. Grasman and G. van Straten, eds. *Predictability and nonlinear modeling in natural sciences and economics*. Kluwer Academic, Dordrecht.
- Kot, M., W. M. Schaffer, G. L. Truty, D. J. Graser, and L. F. Olsen. 1988. Changing criteria for imposing order. *Ecological Modelling* 43:75–110.
- Levin, S. A., B. Grenfell, A. Hastings, and A. S. Perelson. 1997. Mathematical and computational challenges in population biology and ecosystems science. *Science* (Washington, D.C.) 275:334–343.
- McCaffrey, D., S. Ellner, A. R. Gallant, and D. Nychka. 1992. Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *Journal of the American Statistical Association* 87:682–695.
- McCaughey, E., R. M. Nisbet, A. M. de Roos, W. W. Murdoch, and W. S. C. Gurney. 1996. Structured population models of herbivorous zooplankton. *Ecological Monographs* 66:479–501.
- Nychka, D. W., S. Ellner, A. R. Gallant, and D. McCaffrey. 1992. Finding chaos in noisy systems (with discussion). *Journal of the Royal Statistical Society Series B* 54:399–426.
- Nychka, D. W., B. A. Bailey, S. P. Ellner, P. Haaland, and M. O'Connell. 1996. FUNFITS: data analysis and statistical tools for estimating functions. <http://www.stat.ncsu.edu/~nychka>.
- Rand, D. A., and H. Wilson. 1991. Chaotic stochasticity: a ubiquitous source of unpredictability in epidemics. *Proceedings of the Royal Society of London B, Biological Sciences* 246:179–184.
- Rhodes, C. J., and R. M. Anderson. 1996. A scaling analysis of measles epidemics in a small population. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* 351:1679–1688.
- Schaffer, W. M., L. M. Olsen, G. L. Truty, and S. L. Fulmer. 1990. The case for chaos in childhood epidemics. Pages 138–166 in S. Krasner, ed. *The ubiquity of chaos*. American Association for the Advancement of Science, Washington, D.C.
- Schenzle, E. 1984. An age-structured model of pre- and post-vaccination measles transmission. *IMA Journal of Mathematics Applied in Medicine and Biology* 1: 169–191.
- Simkiss, K., S. Daniels, and R. H. Smith. 1993. Effects of population density and cadmium toxicity on growth and survival of blowflies. *Environmental Pollution* 81: 41–45.
- StatSci. 1995. *S-PLUS guide to statistical and mathematical analysis*. Mathsoft, Inc., Seattle, Wash.
- Sugihara, G., and R. M. May. 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* (London) 344:734–741.
- Tidd, C. W., L. F. Olsen, and W. M. Schaffer. 1993. The case for chaos in childhood epidemics. II. Predicting historical epidemics from mathematical models. *Proceedings of the Royal Society of London B, Biological Sciences* 254:257–273.
- Turchin, P. 1993. Chaos and stability in rodent population dynamics: evidence from nonlinear time-series analysis. *Oikos* 68:167–172.
- Weigend, A. S., and N. A. Gershenfeld, eds. 1994. *Time series prediction: forecasting the future and understanding the past*. Addison-Wesley, New York.
- Wolff, R. C. W. 1992. Local Lyapunov exponents: looking closely at chaos. *Journal of the Royal Statistical Society B, Methodological* 54:353–371.
- Wood, S. N. 1994. Obtaining birth and mortality patterns from structured population trajectories. *Ecological Monographs* 64:23–44.
- Yao, Q. W., and H. Tong. 1995. On prediction and chaos in stochastic systems. Pages 57–86 in H. Tong, ed. *Chaos and forecasting: proceedings of the Royal Society Discussion Meeting*. World Scientific, Singapore.