Estimating the distance to an epidemic threshold

Eamon B. O'Dea Andrew W. Park John M. Drake

January 13, 2018

Abstract

Predicting infectious disease emergence or eradication depends on monitoring the distance to the epidemic threshold. One approach to such monitoring, the early warning signals approach, is to quantify the slowing down of dynamics that is characteristic of an approach to a threshold. However, in the susceptible-infected-recovered (SIR) model, the vital dynamics of the host population may occur slowly even when transmission is far from threshold levels. The extent to which fluctuations of an individual variable can provide an estimate of the distance to the threshold, then, depends on the relative weighting of transmission and vital dynamics in the fluctuations. Here we show analytically how this weighting depends on the covariance of the perturbations to a system with two degrees of freedom. Although these results are exact only in the limit of long-term observation of a large system, we find that they still provide useful insight into the behavior of estimates from simulations with a range of population sizes, environmental noise, and observation schemes. Having established some guidelines about when estimates are accurate, we then illustrate how multiple distance estimates can be used to estimate the rate of approach to the threshold.

keywords: slowing down; early warning; infectious disease model; multivariate statistics

Introduction

Many infectious disease epidemics occur with sufficient regularity that their anticipation is straightforward. For example, seasonal influenza has a pronounced winter seasonality in most of the world, with annual outbreaks [1]. Some systems are more episodic but still well-understood, such as measles in sub-Saharan Africa where regional inter-epidemic periods between 1–4 years have been observed in recent times [2]. In contrast, emerging and re-emerging infectious diseases are rarely anticipated, even though the root causes are often discerned soon after the event. Many childhood infectious diseases naturally spread effectively, including measles, chickenpox and rubella. This means that in unvaccinated populations, one infectious individual may infect many others, measured by the pathogen's basic reproduction number, R_0 [3]. Outbreaks are prevented in these cases by maintaining a very high proportion of vaccinated individuals, generating herd immunity in which the effective reproduction number is below 1, meaning small chains of transmission are quickly broken [4]. Reduced vaccine uptake rates can move the infectious disease system from controlled (sub-critical, with effective $R_0 < 1$) to super-critical when outbreaks may occur [5]. Alternatively, other features of the system may be slowly changing, similarly enhancing the transmission of the pathogen. Host demographic changes, particularly rising birth rates, can increase the supply of susceptible individuals to the population, and pathogens frequently evolve at high rates, whereby fitter strains (higher R_0) may be favored by selection [6]. Predicting a dynamical system's movement from sub- to super-critical before it happens has enormous potential to remove the element of surprise associated with emerging infectious diseases, to prioritize mitigation strategies to reverse, stop, or slow the transition, and in worst cases to simply be better prepared for the inevitable. Recent work has also illustrated that following a transition from sub- to super-critical there is a characterizable bifurcation delay—a waiting time until the outbreak actually occurs following suitable conditions being met [7]. Consequently, estimates of how far a system is from the epidemic threshold could help public health officials make judgments about policy, infer on which side of the threshold the population lies, and track the movement of a system towards a threshold (providing early warnings) and even away from a threshold as a way of evaluating the effectiveness of any external changes to the system aimed at controlling infectious disease outbreaks.

A potentially robust basis for estimating the distance to a threshold is the general slowing down of a system's dynamics as a threshold is approached. To be more precise, the average decay rate of deviations from a fixed point of the system becomes increasingly smaller as the parameters of the system approach the point at which that fixed point becomes unstable. Wissel [8] pointed out that this phenomenon, known as *critical slowing down* or sometimes simply as *slowing down*, could be used to determine whether the parameters of a system were approaching a threshold that, when crossed, could result in the system changing in an abrupt and drastic manner. Such changes have come to be called *critical transitions* [9]. Recently a great deal of interest has developed in the possibility of devising model-independent methods to anticipate critical transitions in complex systems using early-warning signals [10]. In general, early-warning signals are statistical properties of observations of systems that can be expected to change in characteristic ways as a threshold is approached. Perhaps the most common examples are increasing autocorrelation and variance of model variables. These signals can often be derived from the increasingly slow decay of perturbations due to slowing down, and many other early-warning signals are in one way or another quantifications of slowing down. The beauty of early-warning signals is that their basis in generic properties of dynamical systems means they have the potential to be reliable even when the system is complex and unidentifiable. Examples of complex and poorly identified systems abound in ecology and epidemiology. With application to such systems in mind, O'Regan and Drake and O'Regan and others [11, 12] demonstrated the application of early-warning signals based on slowing down to forecasting infectious disease emergence and eradication. Further development and integration of these methods into surveillance systems may provide a novel and broadly applicable method of evaluating the control of infectious diseases from existing surviellance data streams.

To explain some of the current challenges in further developing approaches to estimating the distance to the threshold, we will make reference to some elements of dynamical systems theory. Following Wiggins [13], a general dynamical system may be written as a system of equations for a vector field $\dot{x} = f(x, \theta)$, where the overdot indicates a derivative with respect to time, x is a vector of real numbers that determine the point of the system in its phase space, and θ is a vector of real numbers that are parameters of the system. A solution to the system is a function x of time that over some time interval satisfies $\dot{x} = f(x(t), \theta)$. A fixed point x^* of the system is a solution that does not change with time (i.e., it satisfies $0 = f(x^*, \theta)$). Such a point is also referred to as a steady state or an equilibrium of the system. A fixed point is called asymptotically stable if solutions that start at points near the fixed point move closer to it over time. Because the starting points are nearby, deviations $z = x - x^*$ are small and can be accurately modeled by solutions to the linear system $\dot{z} = Fz$, where F denotes the matrix of first derivatives of f with respect to x (i.e., the Jacobian matrix). The general solution of such a system is $z(t) = \exp(Ft)z(0)$. If the real parts of all of the eigenvalues of F are negative, this solution will shrink to zero and it follows that x^* is asympotically stable. If the real parts of any of the eigenvalues are positive, the solution will not shrink to zero and x^* is not asymptotically stable. Thus, as long as the real parts of the eigenvalues of F are not zero, their signs tell us whether or not any fixed point is stable.

The relationship between the speed of a system's dynamics and the distance to the threshold arises in the common case that the eigenvalues of F are continuous functions of the parameters θ of the system and none of the eigenvalues have zero real parts. In this case for a stable fixed point to become unstable, one of the eigenvalues must cross zero. Thus as the parameters approach the threshold where stability is lost, one of the eigenvalues must approach zero in its real part. We call such an eigenvalue an *informative eigenvalue* since its value is informative of how far the system's parameters are from a threshold. We call the magnitude of such an eigenvalue a distance to the threshold. If an informative eigenvalue can be monitored over time, one can determine whether the system is approaching a threshold or not and even make a forecast of when the threshold will be crossed. An informative eigenvalue can be measured by monitoring the decay of small perturbations away from the fixed point along the eigendirection of the informative eigenvalue. Identifying trends in such a decay rate is the goal of early-warning signals based on slowing down.

Despite the simplicity of this goal, it is currently not clear exactly how it can be achieved when systems have multi-dimensional phase space. When one of the eigenvalues of F gets closer to zero, only a small number of the model's observable variables may become less resilient to perturbations. The implication is that early-warning signals such as increasing variance and autocorrelation will not be present in all of the model's variables. Several authors have provided examples of such a case. Kuehn [14] showed that in a susceptible–infected–susceptible (SIS) model of an epidemic on an adaptive contact network, only one of the three model variables had a clear increase in variance as the epidemic threshold was crossed. Boerlijst and others [15]

even showed that, depending on the types of perturbations a system experiences, the autocorrelation of some variables may either increase or decrease as a threshold is approached. Consequently, a recent review [16] identified the selection of appropriate variables in multivariate systems for detection of slowing down as an important problem in need of solution. Dakos [17] has recently used an eigendecomposition of F to derive a simple rule about which state variables have a decay rate that is most affected by the dominant eigenvalue of F. However, this approach only provides a partial answer to the question of variable selection because it does not account for the covariance of the perturbations to the system, which can be as important as the eigenvectors of F on the decay rate of a state variable. Furthermore, another consequence of a models having multiple dimensions is that the informative eigenvalue may not necessarily be the dominant eigenvalue. When its real part gets close enough to zero, the informative eigenvalue will of course become dominant but, as we shall demonstrate, that may not happen until it is very small. So although slowing down is often explained to be a consequence of the dominant eigenvalue approaching zero, methods to estimate the dominant eigenvalue of F from a multivariate time series may not reliably estimate the distance to the threshold. There does not seem to be any general approach for estimating the distance to the threshold in multidimensional systems.

In this work, we derive an explicit relationship between the eigenvalues of F and the autocorrelation function of each of the variables in a multivariate system. The resulting equations lead us to a simple condition for determining the types of perturbations under which estimation of a variable's autocovariance function can be translated into an estimate of the distance to the threshold. We demonstrate the application of this method to the susceptible–infected–removed (SIR) model for directly transmitted infectious diseases. We find that for parameters relevant to many vaccine-preventable diseases, the autocorrelation of the number infected almost always is indicative of the distance to the epidemic threshold, while the autocorrelation of the number susceptible is not. We examine the sensitivity of the accuracy of these estimates to environmental noise, small population size, the frequency of observation, and observation of case reports instead of the actual number infected. We also show a simple example of estimating the change in the distance to the threshold over the length of a time series. These results demonstrate the general feasibility of developing statistical systems for forecasting disease emergence and documenting the approach to elimination.

Methods

Model

The model that motivated the development of the following methods is the susceptible–infected–removed (SIR) model with demography. We let X denote the number of susceptible individuals, Y the number of infected (and infectious) individuals, Z the number of removed individuals (recovered or vaccinated), and $N = X + Y + Z$ the total population size. Typically, we assume that these numbers are the integer-valued random variables of a Markov process having the parameters defined in table 1 and the transitions defined in table 2. We also consider models where the death rate or the force of infection (i.e., the per capita rate at which susceptibles become infected) is subject to variation over time due to fluctuations in the environment over time. We follow Bretó and Ionides [18] in modeling such variation as multiplicative gamma white ("temporally uncorrelated") noise. Bretó and Ionides [18] show that the model remains Markovian with such noise with the modified propensities for the death and transmission events given by the expressions in table 3.

There are several biological assumptions implicit in our model. We use the standard assumption of frequency-dependent transmission, which has been shown to be a more appropriate model than the common alternative assumption of density-dependent transmission for a number of infectious diseases [19]. Another assumption is that the average death rate of individuals is constant throughout their lifetimes. This assumption is also common and is reasonable for populations in some developing countries. A key feature of our model is the inclusion of the η term in the force of infection (tables 2 and 3), which relaxes the assumption that the population is closed to infection from other populations or environmental reservoirs. We include such a term to allow our model to represent populations in which an infectious disease is repeatedly introduced but unable to persist within the population.

Although the model is stochastic, the expected value of the model's variables is deterministic. The rate of change in the expected value when the system is in a given state can be approximated summing over all possible updates in tables 2 and 3 and weighting each update by its propensity [20]. Calculating the rate of change in the expected value of X, Y, and Z in this way leads to the following system of differential

Table 1: Model parameters

Symbol	Definition	Default value
η	importation rate	$400\,N_0^{-1}\text{year}^{-1}$
β	transmission rate	varied
γ	recovery rate	$365/22$ year^{-1}
μ	death rate	0.02 year^{-1}
N_0	initial population size	10^7 individuals
	observation frequency	52 year^{-1}
τ_f	magnitude of environmental noise in force of infection	0
τ_d	magnitude of environmental noise in death rate	0

Table 2: Transitions of the SIR model

Name	$(\Delta X, \Delta Y, \Delta Z)$	Propensity
birth	$(\ 1,\ 0,\ 0)$	$N_0\mu$
death of X	$(-1,\ 0,\ 0)$	$X\mu$
death of Y	$(\ 0,-1,\ 0)$	$Y\mu$
death of Z	$(\ 0,\ 0,-1)$	$Z\mu$
transmission	$(-1,\ 1,\ 0)$	$XY\beta/N_0 + X\eta$
recovery	$(\ 0,-1,\ 1)$	$Y\gamma$

equations

$$\dot{X} = N_0\mu - \bar{\lambda}X - \bar{\mu}X, \tag{1}$$

$$\dot{Y} = \bar{\lambda}X - \gamma Y - \bar{\mu}Y, \tag{2}$$

$$\dot{Z} = \gamma Y - \bar{\mu}Z, \tag{3}$$

where the overdot indicate a time derivative and where

$$\bar{\lambda} = \begin{cases} \tau_f^{-1}\ln(1+\tau_f(\beta Y/N_0+\eta)), & \tau_f > 0, \\ \beta Y/N_0 + \eta, & \tau_f = 0, \end{cases} \tag{4}$$

$$\bar{\mu} = \begin{cases} \tau_d^{-1}\ln(1+\tau_d\mu), & \tau_d > 0, \\ \mu, & \tau_d = 0. \end{cases} \tag{5}$$

The equations for $\bar{\lambda}$ and $\bar{\mu}$ in the case of non-zero environmental noise are the infinitesimal means derived in Bretó and Ionides [18]. By setting these equations equal to zero and solving for X and Y, we can find the approximate fixed point of the system for a given set of model parameters.

The equations for the fixed point of the differential equations allow us to explain what we mean by epidemic threshold. For the sake of clarity, we consider the equations only when τ_d and τ_f are zero. In that case the exact equation for the Y-coordinate of the fixed point, which we denote Y^*, is

$$Y^* = \frac{N_0}{2}\left[\frac{\mu}{\beta}(R_0-1) - \frac{\eta}{\beta}\right] + \frac{N_0}{2}\sqrt{\left[\frac{\mu}{\beta}(R_0-1) - \frac{\eta}{\beta}\right]^2 + 4\frac{\mu\eta}{(\gamma+\mu)\beta}} \tag{6}$$

where $R_0 = \beta/(\gamma+\mu)$. R_0 is known as the basic reproduction number and we consider the epidemic threshold to be the surface in parameter space where $R_0 = 1$ and $\eta = 0$. To see why, note that when $\eta = 0$, equation

Table 3: Modified transitions of the SIR model allowing for environmental heterogeneity

Name	$(\Delta X, \Delta Y, \Delta Z)$	Propensity
death	$(-k_1, -k_2, -k_3)$	$\binom{X}{k_1}\binom{Y}{k_2}\binom{Z}{k_3}\sum_{j=0}^{k_1+k_2+k_3}\binom{k_1+k_2+k_3}{j}(-1)^{k_1+k_2+k_3-j+1}\tau_d^{-1}\ln(1+\mu\tau_d(X+Y+Z-j))$
transmission	$(-k,\ \ k,\ \ 0)$	$\binom{X}{k}\sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j+1}\tau_f^{-1}\ln(1+(\beta Y/N_0+\eta)\tau_f(X-j))$

(6) has a non-zero value only when $R_0 > 1$; only when $R_0 > 1$ will the introduction of an infection into a susceptible population lead to an epidemic according to the system of differential equations. Accordingly, one can interpret R_0 as the average number of new infections caused by an infected individual in a susceptible population. From the point of view of fixed points, the epidemic threshold separates the region of parameter space where a fixed point occurs with $Y^* = 0$, a disease-free equilibrium, from the region where a fixed point occurs with some $Y^* > 0$, an endemic equilibrium.

When there is a small rate at which individuals can be infected from other populations or an environmental source ($0 < \eta \ll 1$), there is no longer a disease-free equilibrium but the concept of an epidemic threshold is still relevant. One can see from equation (6) that Y^* is slightly increased by the addition of η to the force of infection. When R_0 is not too close to 1, Y^* can be well-approximated by making a linear approximation to the square root function starting at the point where its argument is equal to its first term, which yields

$$Y^*/N_0 \approx \max\left(0, \frac{\mu}{\beta}(R_0 - 1) - \frac{\eta}{\beta}\right) + \frac{\eta}{|\beta - \gamma - \mu - \eta(\gamma + \mu)/\mu|} \tag{7}$$

$$\approx \max\left(0, \frac{\mu}{\beta}(R_0 - 1)\right) + \frac{\eta}{|\beta - \gamma - \mu|}. \tag{8}$$

Using this approximation, one can see that when $R_0 \ll 1$, Y^*/N_0 scales with η. When $R_0 \gg 1$, Y^*/N_0 is potentially much larger and effectively independent of η. Also, a large epidemic is only possible when $R_0 \gg 1$. Therefore, we can still consider $R_0 = 1$ as an epidemic threshold for small η. To clarify that Y^*/N_0 is not too large when $R_0 \approx 1$ we need a second approximation. In this case, Y^*/N_0 can be bounded by applying the triangle inequality to obtain

$$Y^*/N_0 < \max\left(0, \frac{\mu}{\beta}(R_0 - 1) - \frac{\eta}{\beta}\right) + \sqrt{\mu\eta/[(\gamma + \mu)\beta]}. \tag{9}$$

This bound on Y^*/N_0 scales with $\sqrt{\eta}$. Accordingly, Y^*/N_0 is intermediate in size to those cases when R_0 is far from 1. Thus, although when $\eta > 0$ the model never passes through an epidemic threshold, as R_0 passes through one Y^*/N_0 behaves similarly to the case where the model parameters do pass through a threshold point (i.e., when $\eta = 0$). Thus it is still of interest to establish how close the parameters are to the point of the epidemic threshold.

Relating eigenvalues to autocovariance

When the dynamics are characterized by small fluctuations around a fixed point, the degree of autocorrelation of these fluctuations may be indicative of the distance to the threshold. A first step in demonstrating this relationship is to derive a probability density function for the fluctuations. Let $z(t)$ denote a vector of deviations from the fixed point that is in units of the square root of the system's size. Let $p(z)$ be the probability density function of these deviations. In the limit of a large system size, this function may be approximated as the solution to the Fokker–Planck equation

$$\frac{\partial p(z,t)}{\partial t} = \sum_{ij} -f_{ij}\frac{\partial(z_j p)}{\partial z_i} + \frac{1}{2}\sum_{ij} d_{ij}\frac{\partial^2 p}{\partial z_i \partial z_j}, \tag{10}$$

where the matrix F (with elements f_{ij}) determines the expected trajectory of z toward zero and the matrix D (with elements d_{ij}) describes the covariance of a Gaussian white noise process that acts on z. The matrices F and D follow directly from the transition probabilities. For the SIR model in the previous subsection, we take N_0 as the system size, $z = ((X - X^*)/\sqrt{N_0}, (Y - Y^*)/\sqrt{N_0})$ and obtain

$$F(X^*, Y^*) = \begin{pmatrix} -\bar{\lambda} - \bar{\mu} & -\frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}Y}X^* \\ \bar{\lambda} & \frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}Y}X^* - \gamma - \bar{\mu} \end{pmatrix}, \tag{11}$$

where $\bar{\lambda}$ and $\frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}Y}$ are evaluated at $Y = Y^*$. For the covariance matrix, we obtain

$$D(X^*, Y^*) = \begin{pmatrix} \mu + m_{X,Y} + m_{X,\emptyset} & -m_{X,Y} + m_{XY,\emptyset} \\ -m_{X,Y} + m_{XY,\emptyset} & \gamma Y^*/N_0 + m_{X,Y} + m_{Y,\emptyset} \end{pmatrix}, \tag{12}$$

where

$$m_{X,Y} = \begin{cases} X^*\bar{\lambda}/N_0 + X^*(X^*-1)[2\bar{\lambda} - \tau_f^{-1}\ln(1+2\tau_f(\beta Y^*/N_0+\eta))]/N_0, & \tau_f > 0, \\ X^*\bar{\lambda}/N_0, & \tau_f = 0, \end{cases} \tag{13}$$

$$m_{X,\emptyset} = \begin{cases} X^*\bar{\mu}/N_0 + X^*(X^*-1)[2\bar{\mu} - \tau_d^{-1}\ln(1+2\tau_d\mu)]/N_0, & \tau_d > 0, \\ X^*\bar{\mu}/N_0, & \tau_d = 0, \end{cases} \tag{14}$$

$$m_{Y,\emptyset} = \begin{cases} Y^*\bar{\mu}/N_0 + Y^*(Y^*-1)[2\bar{\mu} - \tau_d^{-1}\ln(1+2\tau_d\mu)]/N_0, & \tau_d > 0, \\ Y^*\bar{\mu}/N_0, & \tau_d = 0, \end{cases} \tag{15}$$

$$m_{XY,\emptyset} = \begin{cases} X^*Y^*[2\bar{\mu} - \tau_d^{-1}\ln(1+2\tau_d\mu)]/N_0, & \tau_d > 0, \\ 0, & \tau_d = 0. \end{cases} \tag{16}$$

A solution to equation (10) is a Gaussian density function with a mean of zero and a covariance matrix Σ (with elements σ_{ij}) that depends on F and D. van Kampen [20] provides a detailed introduction to these methods.

For these Gaussian solutions, the autocovariance function of the deviations may be written in terms of the eigenvalues of F. The relationship is particularly simple when the eigenvectors of F are used as the basis of the coordinates. Thus let $\tilde{z} = W^{-1}z$, where W is a matrix of the eigenvectors of F, and let $\tilde{\Sigma}$ denote the covariance matrix of $\tilde{z}$. Then, using the decomposition of Kwon and others [21], it follows that

$$\tilde{\sigma}_{ij} = -\tilde{d}_{ij}/(\lambda_i + \lambda_j), \tag{17}$$

where λ_i denotes an eigenvalue of F and we assume that all of these eigenvalues are distinct. The autocovariance matrix is defined as $\Sigma_\tau = \langle z(t-\tau)z(t)^\intercal \rangle$, where the angular brackets denote expected value over time or realizations of the system. It follows from the stationarity of the solution that $\Sigma_\tau = \exp(F\tau)\Sigma$. In the eigenvector basis, we have $\tilde{\sigma}_{\tau,ij} = \exp(\lambda_i\tau)\tilde{\sigma}_{ij}$. Thus the behavior of the autocovariance along an eigendirection as a function of the lag τ is a simple and identifiable function of the corresponding eigenvalue. If λ_i is real, then $\tilde{\sigma}_{\tau,ii}$ decays exponentially toward zero at the rate λ_i. If λ_i has an imaginary component, then the real and imaginary parts of $\tilde{\sigma}_{\tau,ii}$ oscillate around zero with a frequency given by the imaginary component of λ_i and an amplitude that decays exponentially at the rate given by the real component of λ_i. Since $\Sigma_\tau = W\tilde{\Sigma}_\tau W^\intercal$, $\sigma_{\tau,ii}$ will be a linear combination of the elements of $\tilde{\Sigma}_\tau$. Therefore, the elements of the autocovariance matrix Σ_τ are linear combinations of functions from which the eigenvalues of F are identifiable.

Solving for the space of suitable noise parameters

The relationship between the eigenvalues and the autocovariance established in the previous subsection clarifies the question of when the autocovariance of a variable contains sufficient information to estimate the distance to a threshold. Any threshold corresponds to an eigenvalue crossing zero. Recall that we call such an eigenvalue an informative eigenvalue and that the magnitude of its real part can be considered the distance to the threshold. If it is known that the imaginary part of the eigenvalue will also be zero at the threshold, then the magnitude of the imaginary part can be considered a second component of the distance. Note that in the case that an informative eigenvalue is complex it will be a part of a conjugate pair. Estimation of the decay rate and frequency of oscillation of a variable's autocovariance function can provide an estimate of the distance to the threshold when they are close to the real and imaginary parts of an informative eigenvalue. This condition on the autocovariance function for an estimate to be accurate, together with equations $\Sigma_\tau = W\tilde{\Sigma}_\tau W^\intercal$ and equation (17), can be translated into conditions on the eigenvectors W of F and the covariance matrix D of the perturbations. Thus, we now have a general link between the parameters of models and the potential for a model variable to provide an estimate of the distance to the threshold. In the electronic supplementary material, we provide an explicit calculation of the values of D that permit a distance estimate for each variable.

Obtaining distance estimates from time series

We estimate the distance to the threshold from a time series as follows. The main idea is to suppose that the autocorrelation will exponentially decay with increasing lags at a rate equal to the real part of the informative eigenvalue and that any oscillations in the autocorrelation function have a frequency equal in magnitude to the imaginary part of the informative eigenvalue. The first step is then to estimate the autocorrelation of the time series for a series of lags, which we demonstrate using the acf function in R. Because sometimes the autocorrelation can have cycles with a period of several years, we used lags from 0 to 30 observations less than the length of the time series. Next we use a nonlinear least-squares optimizer to fit two models for the estimated autocorrelation $\hat{\sigma}_{ii,\tau}/\hat{\sigma}_{ii}$:

$$\hat{\sigma}_{ii,\tau}/\hat{\sigma}_{ii} = \mathrm{e}^{g\tau} + e_\tau \tag{18}$$

$$\hat{\sigma}_{ii,\tau}/\hat{\sigma}_{ii} = \sqrt{1+a^2}\mathrm{e}^{g\tau}\sin(\omega\tau + \mathrm{atan2}(1,a)) + e_\tau, \tag{19}$$

where e_τ is an error term, g is the decay rate parameter, ω is the frequency parameter, and a is a phase angle parameter, and atan2 is the inverse tangent function with arguments in the order y, x. We use the nlsLM function to fit these models. This function is available in the minpack.lm package [22], and it provides an R interface to the Levenberg–Marquardt optimizer in the MINPACK library. We used nlsLM instead of the nls function that comes with R because it was less sensitive to the choice of initial values of the parameters for the optimization of the model fit. For initial values, we set a to zero, g to the least-squares slope of the log of the absolute value of the estimated autocorrelation versus the lag, and ω to the frequency that maximized the spectral density of the estimated autocorrelation. We fit the data with and without an oscillation component in the model and use the following AIC-like score to evaluate the models: twice the number of parameters plus the residual sum of squared errors. We take the estimates from the model with the lower score. We use these estimates to calculate the distance to the threshold as $\sqrt{\omega^2+g^2}$.

Simulation experiments

In the following, we apply our theory and estimation methods to the SIR model. To generate data for estimation, we simulated time series of the number of individuals in each state according to our Markov process model using the Euler scheme of He and others [23]. The pomp [24, 25] R package was used to implement the model. Our typical procedure was to simulate data with most of the parameters fixed at the default values in table 1 and for several choice of transmission rate. Simple sensitivity analyses of the distance estimates were carried out by allowing one or two of the parameters to vary from the default values. The full set of parameters used for each set of distance estimates is reported in the results. The initial values of the states were set to the equilibrium values and the model was run for 10 simulation years before sampling to allow the initially sampled states to vary according to the stationary distribution of the process. The sampling scheme was 1040 observations at a frequency of 1 observation per week. This corresponds to about 20 years of weekly observations, which is a realistic size for an epidemiological data set. Sampled time series of both the number infected and the number susceptible were used to generate an estimate of the distance to the threshold by the method described above. The true value for each estimate was calculated by plugging the simulation parameters into equation (11), solving for the fixed points, and calculating the eigenvalues of F. If there were two real eigenvalues, the informative eigenvalue was identified as the eigenvalue that would cross zero if the parameters were moved through the bifurcation point where $R_0 = 1$ and $\eta = 0$.

We also conducted simulations with a linearly increasing transmission rate to evaluate the performance of estimates of changes in distance. To ease comparison with estimates from our other simulations, we used a similar amount of data for the individual distance estimates used to calculate the change in distance. The simulations were sampled for twice as long, the time series were split into two windows of 1040 weekly observations, and an estimate was obtained for each window.

Software and reproducibility

Code to reproduce our results is available in an online repository [26].

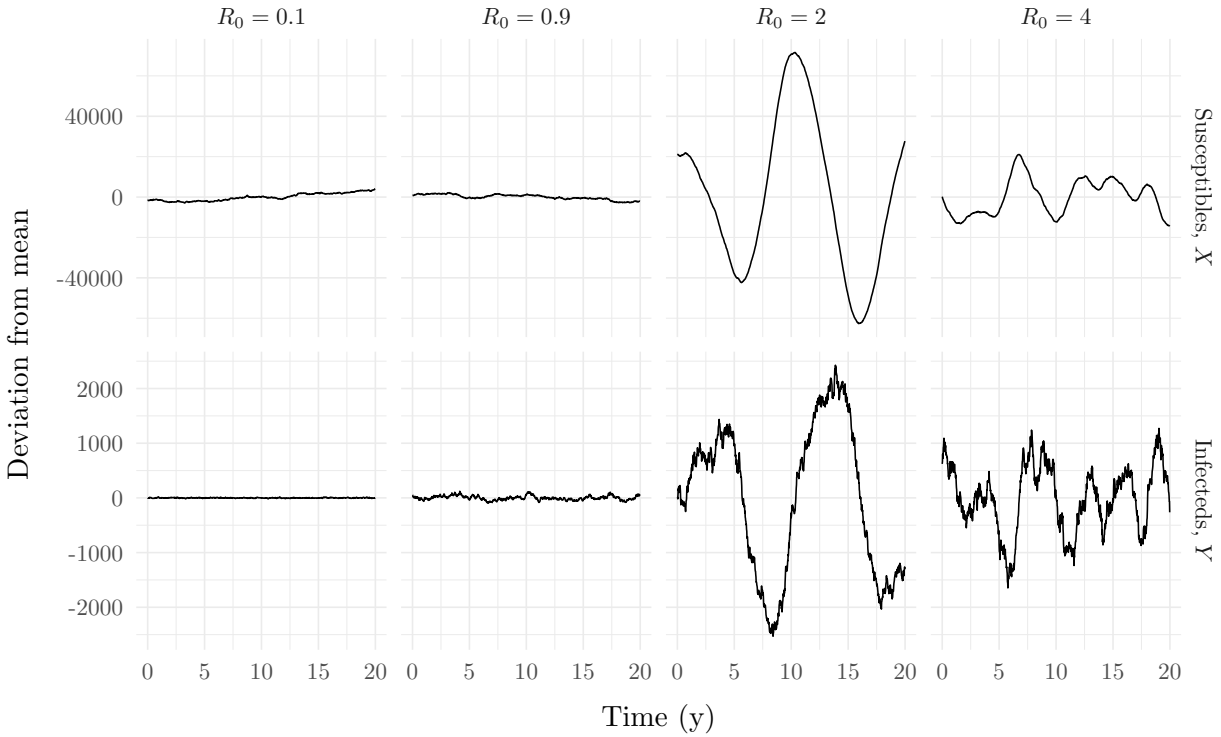

Figure 1: Example of simulated time series that could provide a distance estimate. Deviations are plotted instead of the simulated counts to align the time series vertically and because the deviations are what is used to calculate the estimate. The main idea is to estimate the rate at which deviations decay and oscillate. As these examples suggest, these rates tend to increase as the parameters of the system move away from $R_0 = 1$. Parameters for the simulations are in table 1, with β set to $R_0(\gamma + \mu)$.

Results

Main determinants of the accuracy of distance estimates

We first present some general considerations regarding when the distance to the epidemic threshold can be estimated from the fluctuation dynamics of the SIR model. Figure 1 shows representative examples of the kinds of time series that we suppose could become available for statistical analysis. For two of the parameter values, cycles are visible in both the number of susceptibles, X, and the number of infecteds, Y; which is a consequence of the eigenvalues of the Jacobian, F, being complex. This behavior is typical of parameters for which $R_0 > 1$. We have explained in Methods that in this case, any white-noise perturbation may allow for the distance estimate to be obtained from either variable. When $R_0 < 1$, typically there are two real eigenvalues. Without any knowledge about the model, we would expect that the ability to obtain an estimate depends on the covariance of the perturbations. With the knowledge that the observations come from an SIR model, we could expect that the dynamics of the X and Y variables will be largely independent of each other. The number infected will generally be too small to affect the fluctuations in the number susceptible. Thus the rate at which susceptible perturbations decay will depend mostly on the per-capita death rate μ, whereas the rate at which infected perturbations decay will depend on the sum of the per-capita rates at which Y grows and shrinks, $\beta X^*/N_0 - \gamma - \mu$. Thus the variable Y is generally the one that should be observed to estimate the distance to the threshold when the disease is not widespread. In the electronic supplementary material, we derive explicit equations for the autocorrelations that supports this conclusion.

Having provided some general insights into why distance estimates may be obtained from Y and not X when $R_0 < 1$, we next consider a more specific answer for a specific set of parameters. We use the approach described in Methods to find the set of noise parameters that allow the distance to the threshold

to be estimated with a given accuracy from each variable. These sets appear as regions in space in figure 2. Consistently with the conclusions of the previous paragraph, the regions are much larger for the number infected, Y, than for the number susceptible, X. The regions for Y include the perturbations that result from the intrinsic noise present in simulations of the model with finite population sizes. In contrast, a large part of the lower-error region identified for X is in fact not feasible because the covariance matrix constraint of $d_{12}^2 \leq d_{11}d_{22}$ is not satisfied.

Having shown that, in principle, distance estimates are often possible to obtain from the SIR model from at least one of the variables, we next turn to the question of whether estimates may be obtained in practice from a simulated time series of realistic length using our estimation method. Figure 3 shows that for time series of about 1000 observations our estimation method was generally successful when the perturbations are predicted to be suitable. The perturbations were simply intrinsic noise, so the low accuracy of estimates based on X when $R_0 = (0.1, 0.5, 0.9)$ is consistent with figure 2. As expected, when $R_0 > 1$ estimates from both X and Y were similarly accurate. Therefore Y permits a distance estimate for all R_0 considered. Although there is perhaps more overlap in the estimates from different R_0 than is desirable from 20 years of data, the estimates are sufficiently accurate to be useful.

Sensitivity analyses

Although we found that observation of the number infected could provide accurate distance estimates when the system is subject to intrinsic noise only, figure 2 indicates that deviations from intrinsic noise can change the situation. Thus we next examined the distance estimates with increasing amounts of environmental noise in the death rate. As expected, the $R_0 = 0.1$ panel of figure 4 shows that for sufficiently large amounts environmental noise, the estimates of distance based on Y become inaccurate. However, the amount of noise required to seriously compromise accuracy is large. Figure 5 shows that environmental noise is slow to change the decay of the autocorrelation at short lags, and thus the robustness of the accuracy is a consequence of our distance estimates being sensitive to the rapid decay at short lags. Note that a typical approach when calculating early warning indicators is to look for trends in the autocorrelation at a single lag, and for large enough lags trends in such indicators could mistakenly identify increasing noise in the death rate for an approach to an epidemic threshold. However, we have also noticed that for 20 year time series the increases in the autocorrelation due to environmental noise are typically below their long-term expected value (figure 6), which serves to reduce the effect relative to that expected from figure 5 and also contributes to the robustness of our distance estimates. If the estimated autocorrelations from the 20 year time series looked like those in figure 5, the estimates based on Y would be much more similar to those based on X when the standard deviation in the death rate was 1 and the population size was 10^9.

From the definition of the m terms in equation (12), it can be seen that environmental variation in the death rate increases the covariance of perturbations of X and Y. In contrast, environmental variation in the force of infection makes that covariance more negative. To more thoroughly examine the effects of deviations from pure intrinsic noise we also evaluated distance estimates from simulations with environmental noise in the force of infection. Estimates in He and others [23] for the standard deviation of such noise from measles case reports range from 0.038 to 0.096. Estimates from the analysis of polio case reports in Martinez-Bakker and others [27] for the standard deviation of noise in the force of infection range from $1.6 \cdot 10^{-3}$ to 0.68 (personal communication with P. Rohani). Figure 7 plots distance to threshold estimates from simulations having noise levels that bracket these noise parameter estimates. In contrast to the case of noise in the death rate, distance estimates are not affected by high levels of noise when $R_0 = 0.1$, but they are affected when R_0 is equal to 2 or 16. We inspected the time series in these cases and noticed that the variation in the force of infection leads to frequent interruption of the regular cycles that characterize the dynamics in the presence of intrinsic noise. Bretó and others [28] have noted that environmental variation in the transmission rate can cause such irregularities. Although one could certainly attempt to adapt the distance estimation procedure to such dynamics, our goal was simply to describe a basic procedure and identify some of its limitations.

To that end, we examined with simulation how the accuracy of distance estimates declined with population size. Figure 8 shows that we did not find the accuracy to be generally sensitive to population size. This robustness is remarkable because for small populations the marginal distribution of the time series becomes highly non-Gaussian. However, its seems the Gaussian solution based on large population sizes can still provide a good approximation to the autocorrelation when the population size is fairly small. In short, it

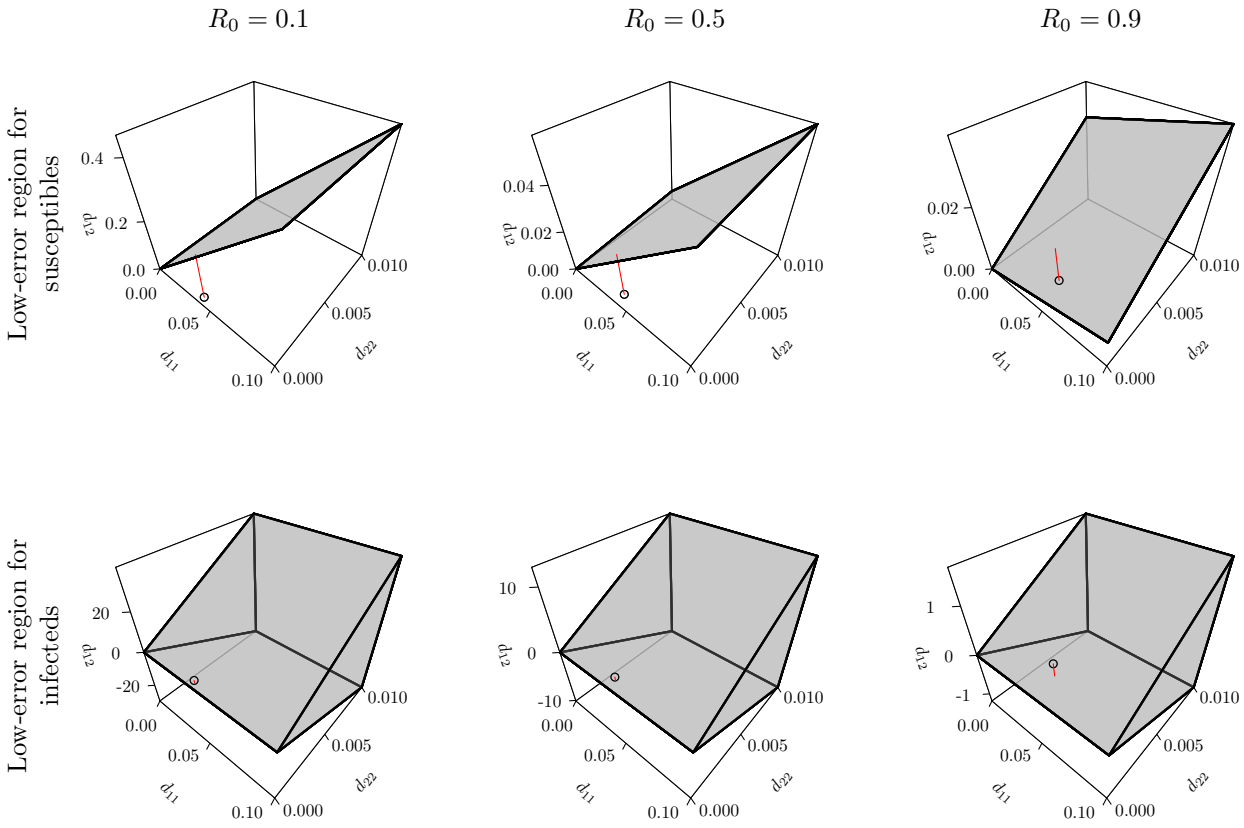

Figure 2: In our SIR model for most $R_0 < 1$, the set of noise parameters under which the distance to the epidemic threshold may be estimated is much larger for the number infected than for the number susceptible. Each panel plots as a prism, for a range of variances in white-noise perturbations to the number susceptible and infected (subscripts 1 and 2, respectively) the region containing the covariances for the perturbations that result in the autocorrelation of the model variable being within 0.1 of the reference that would result in a perfect distance estimate. The axis labels denote elements of the matrix D in equation (10). The points in each panel correspond to the intrinsic perturbations calculated by the linear noise approximation equation (12). The red line segments are drawn vertically from these points to the nearest boundary of the low-error region. The points in the panels for the infecteds fall within the low-error region. Parameters were as in table 1, with β set to 1.66, 8.31 and 14.9.

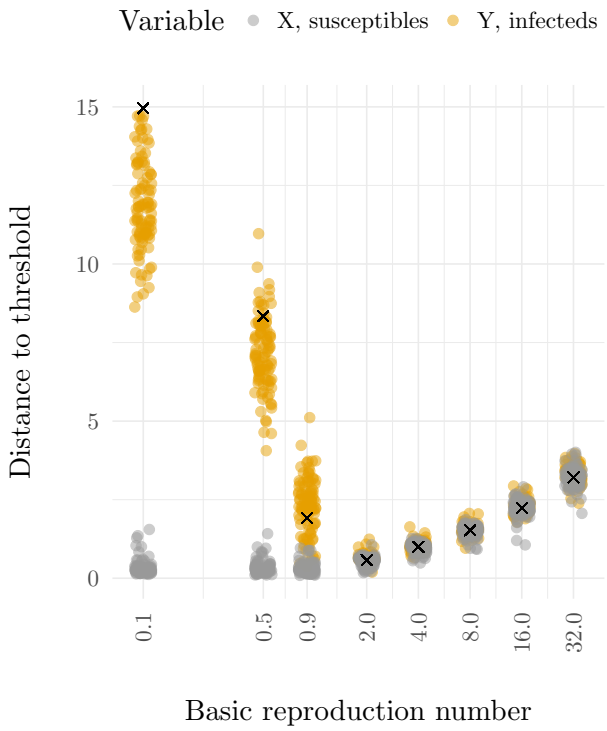

Figure 3: Distance estimates can be obtained from both X and Y in stochastic simulations of our SIR model. However, the distance estimates based on X are not always reliable. The X and Y variables were sampled weekly for 20 years. Parameters for the simulations are in table 1, with β set to $R_0(\gamma + \mu)$. The distance to the threshold is defined here as the modulus of the informative eigenvalue. The true value is marked with an 'x'. There are 100 estimates for each variable for each R_0. The x-values of the points have been jittered to make the density of points at each y-value clearer.

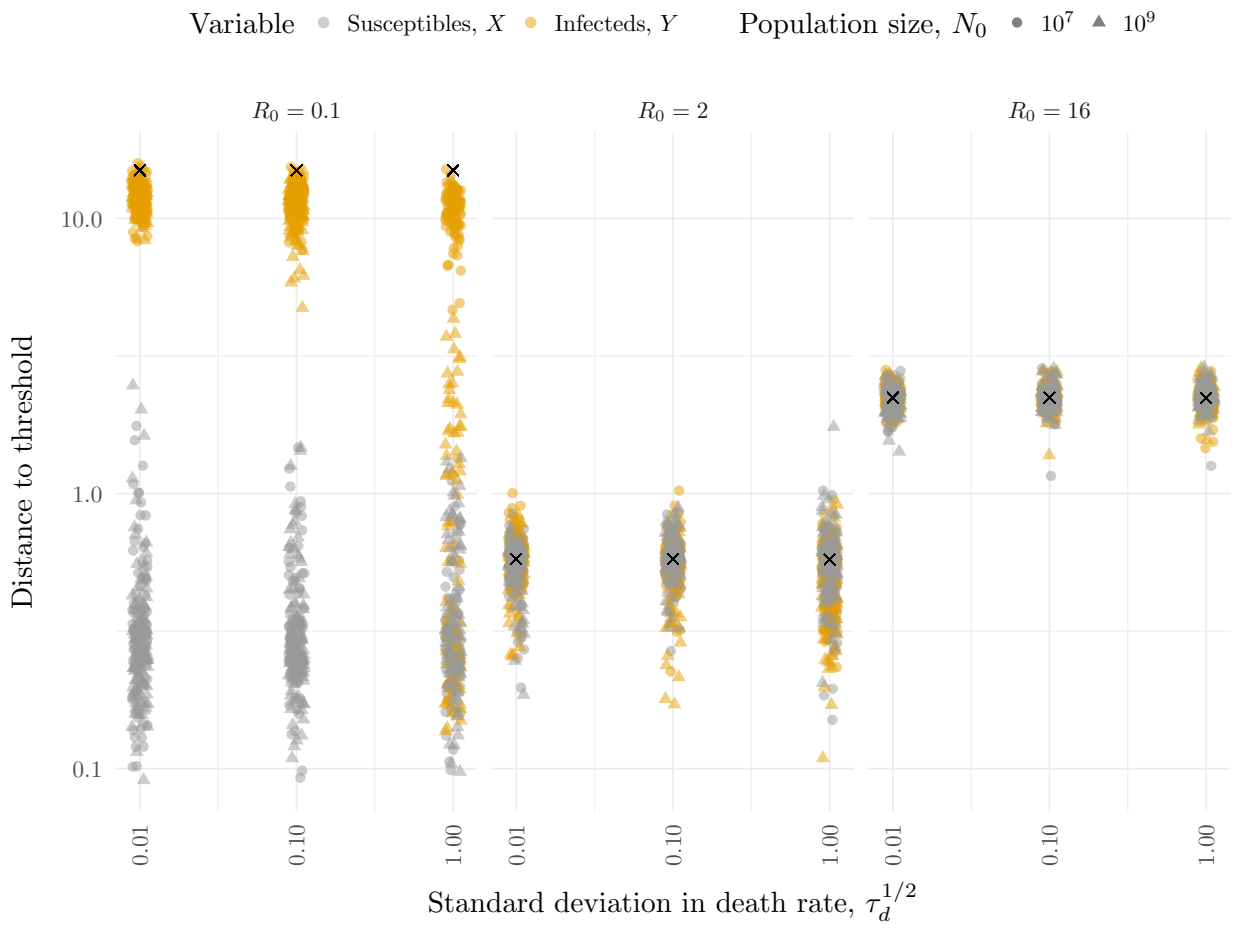

Figure 4: The accuracy of distance estimates is not affected by stochastic variation in the death rate parameter until the variation becomes extreme. The true distance is marked with an '×'. Parameter η was fixed to $4 \cdot 10^{-5}$. Other parameters were in table 1. Parameter β was set to $R_0(\gamma + \mu)$.

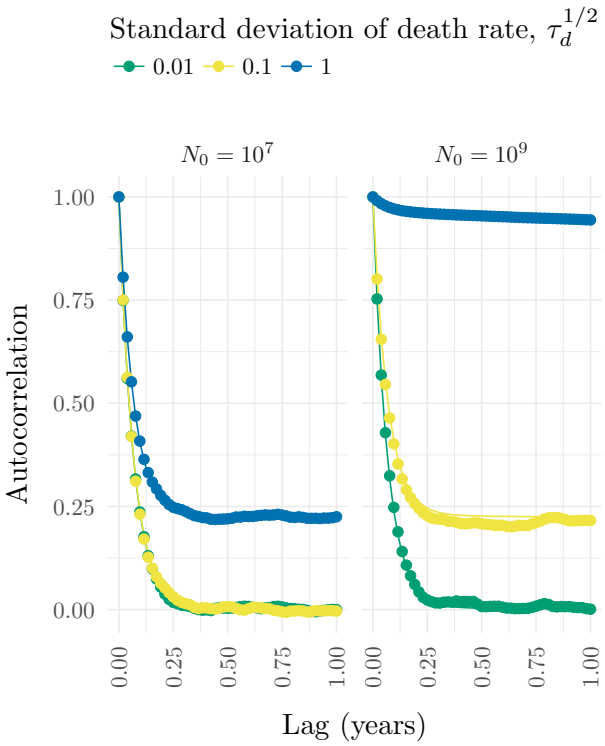

Figure 5: The sensitivity of the autocorrelation of the number infected to environmental noise in the death rate depends on the lag. Points are estimate of the autocorrelation from a long (1000 year) time series simulated according to our Markov process model. Lines represent the analytic calculations based on equation (10). Parameter η was fixed to $4 \cdot 10^{-5}$. Other parameters were as in table 1, with β set to $0.1(\gamma + \mu)$.

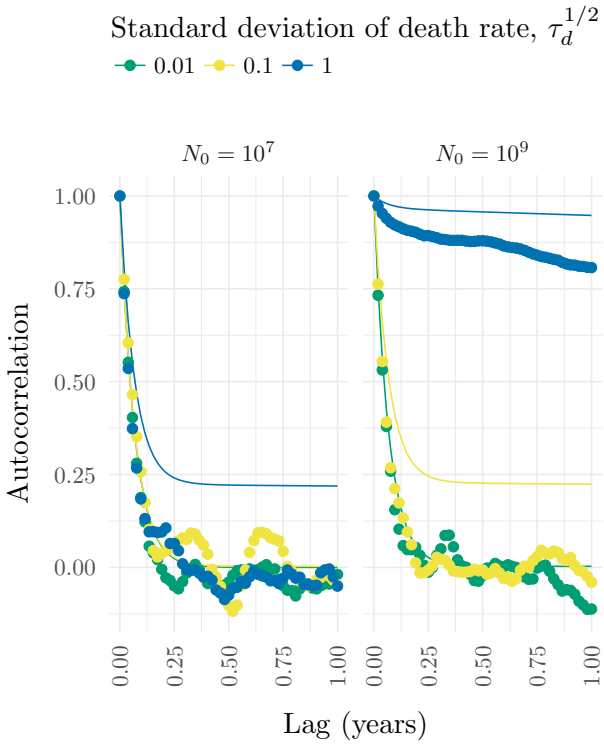

Figure 6: The increase in the autocorrelation with noise in the death rate can be less for 20 year observation periods than the long term average. Points are estimates of the autocorrelation from a 20 year time series simulated according to our Markov process model. Lines represent the analytic calculations based on equation (10). The points are much lower than the lines than is the case in figure 5, where the simulations were much longer. Parameter η was fixed to $4 \cdot 10^{-5}$. Other parameters are in table 1, with β set to $0.1(\gamma + \mu)$.

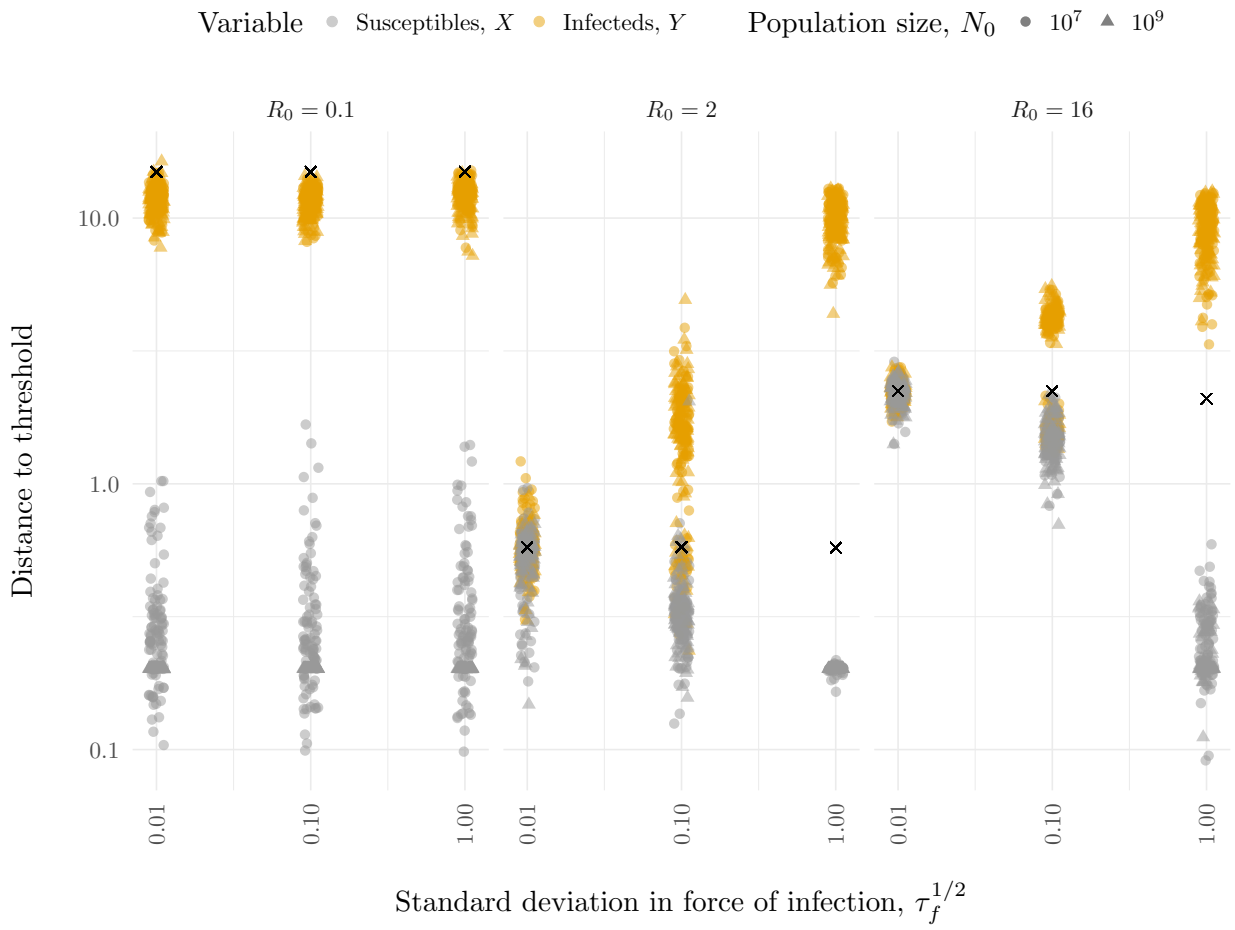

Figure 7: The accuracy of distance estimates is affected by stochastic variation in the force of infection when it is large (here, a standard deviation above 0.01) and there are cycles that it can interrupt (here, R_0 of 2 or 16). The true distance is marked with an 'x'. Parameter η was fixed to $4 \cdot 10^{-5}$. Other parameters are in table 1. β was set to $R_0(\gamma + \mu)$.

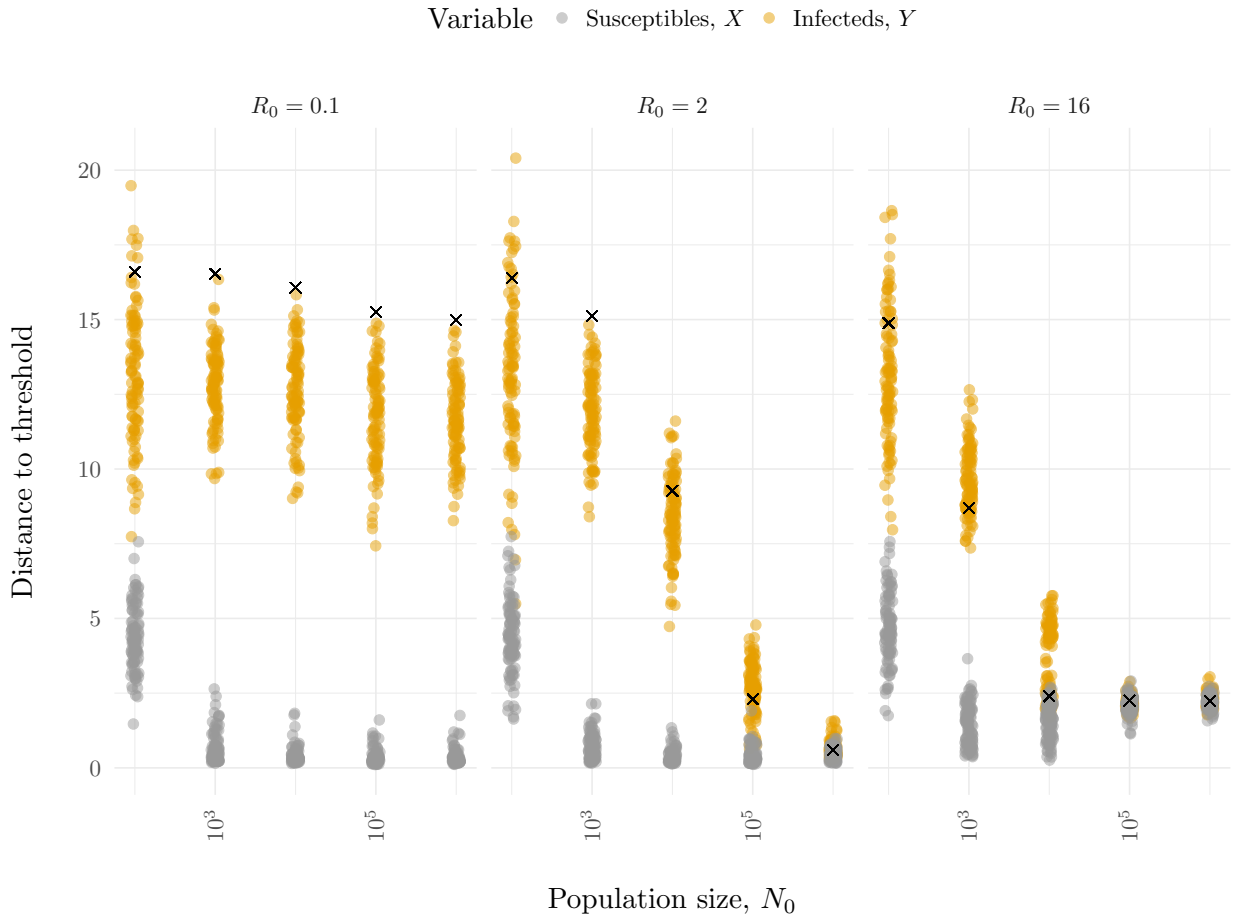

Figure 8: The accuracy of distance estimates is not sensitive to the population size. The true value is marked with an 'x'. As in figure 3, Y generally provides an accurate estimate and X is also generally accurate when the eigenvalues are complex. There are 100 estimates for each variable for each population size. Parameters besides the population size are in table 1, with β set to $R_0(\gamma + \mu)$.

seems that in some cases our methods may be reliable with data from population sizes as small as 100.

As an additional sensitivity test of the distance estimates, we next looked at the effect of the frequency of observations in the time series. To avoid confounding the effects of frequency with the effects of the length of the time series, its length was kept the same for all sampling frequencies by adjusting the stop time of the simulations. In the $R_0 = 2$ panel of figure 9, we see that when such a trade-off exists between the duration of observation and the observation frequency, a high observation frequency can be detrimental due to the short observation period. For daily observation frequency, the total duration of observation was limited to 1000 days, whereas the period length of the oscillations in the autocorrelation was about 11 years. With weekly observation frequency, the duration of observation is about 20 years and the estimates are much better. A rough guideline for accurate estimation seems to be that the duration of observation be at least as long as the period of any oscillation in the autocorrelation function. Another guideline is that the time between observations be much less than the period of oscillation of the autocorrelation. In the $R_0 = 16$ panel of figure 9, we see that the distance estimates become noticeably biased when the observation frequency goes from 0.25 per week to 1 per year. For these parameters, the autocorrelation had a period of about 3 years, so 3 observations per cycle seems much worse than 40 observations per cycle. A similar guideline on the sampling frequency holds when the autocorrelation function is not periodic. In the $R_0 = 0.1$ panel, the estimates based on Y become worse as the sampling frequency dips below 1 per week. Here, the autocorrelation shrinks by

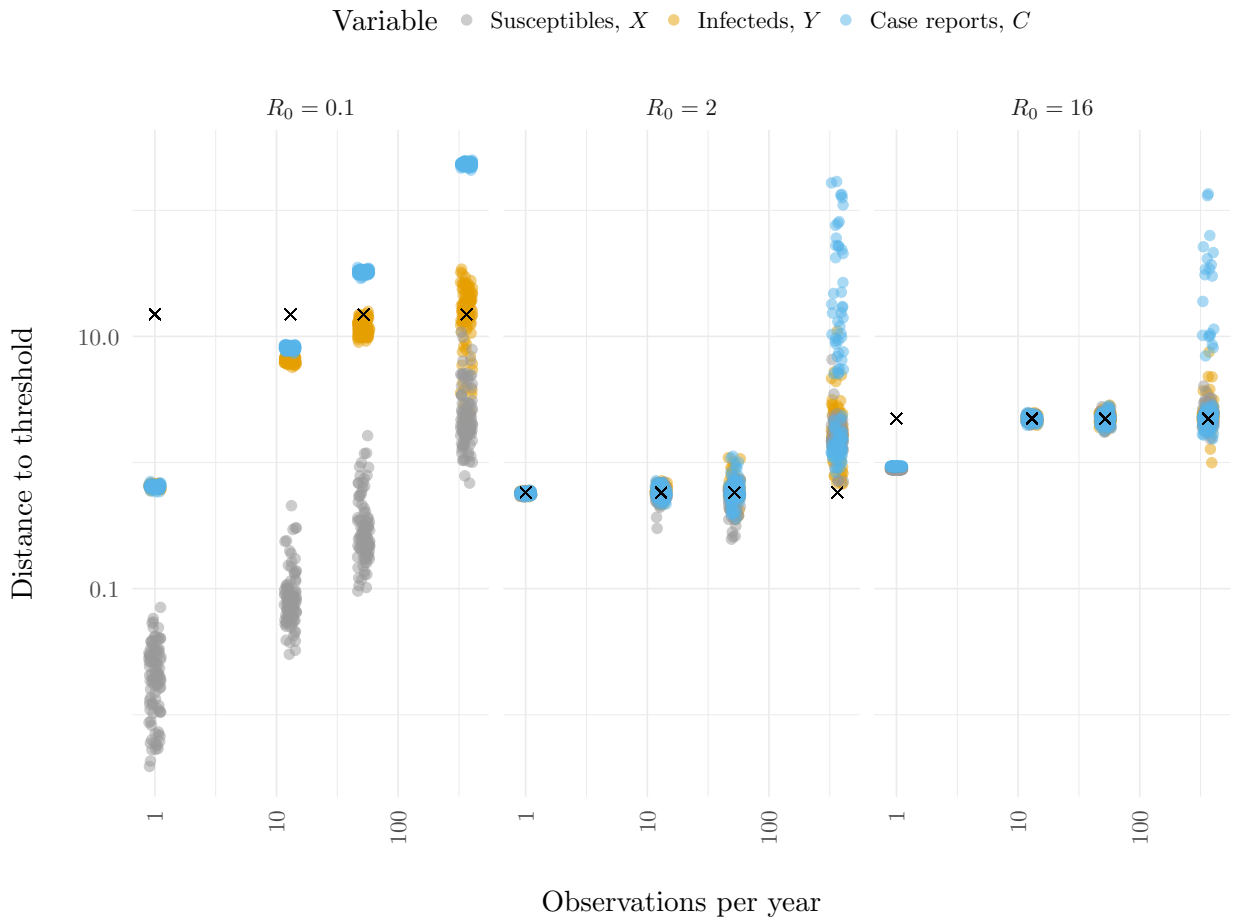

Figure 9: The accuracy of distance estimates is sensitive to the frequency of observations and whether the state of the time series consists of case reports. The true distance is marked with an 'x'. The case reports variable is the accumulated number of recovery events since the last observation. Parameters besides the observation frequency are in table 1, with β set to $R_0(\gamma + \mu)$.

a factor of e (2.7) about every 3.5 weeks. This time can be used to characterize the timescale of a decaying function and is sometimes called the return time. A third guideline, then, is that the time between samples should be less than the return time. In summary, for accurate estimates the duration of observation should be much greater than the time scale of the autocorrelation function but the time between observations should be much smaller than the time scale of the autocorrelation function.

In addition to sampling frequency, another key characteristic of observations is whether they represent direct observation of the state of the system or cumulative flows between states. In particular, it is relatively rare for the number of infections in a population at a given point in time to be observed. A more typical type of observation is the count of the number of infected individuals that moved into the removed class, for example because these individuals visited a doctor and then avoided contact with others. We refer to such counts as case reports. Although our methods were not developed for this type of observation, we were interested in seeing whether they still might be applicable. Figure 9 shows that when the autocorrelation function was periodic (i.e., when R_0 was equal to 2 or 16), the estimates based on case reports were similar to those based on direct observation of X or Y. However, for the short time series, the estimates based on case reports tended to overestimate the distance and be more variable. In our simulations with $R_0 = 0.1$, in which the autocorrelation function did not oscillate, the estimates based on case reports tended to be higher than those based on the direct observation of Y but were less variable. Thus when the sampling frequency

was suitable for accurate estimates based on Y, using case reports led to overestimates of the distance. However, as the sampling frequency declines the estimates based on Y became low, and when there were 13 observations per year estimates based on case reports were actually slightly more accurate. With a sampling frequency of 1 per year, no meaningful difference existed between the two types of estimates. In summary, using case reports in place of the number infected did not spoil distance estimates when the autocorrelation is periodic, but tended to increase distance estimates when the autocorrelation was not periodic.

The electronic supplementary material contains the results of our estimation of the rate of change of the distance to the threshold.

Discussion

This work has presented a general solution to the problem of the selection of appropriate variables in multivariate systems for detection of slowing down as a threshold is approached. The solution is a method of calculating what type of white-noise perturbations, if any, allow slowing down to be detected based on observation of a given variable. To provide a specific example, this general solution has been applied to the SIR model and been shown to be consistent both with a model-specific analysis and with simulations. This application has also served to demonstrate and stress-test a method of estimating a distance to a threshold that is defined as one of the eigenvalues of the linearized model's matrix F. It is worth mentioning that this informative eigenvalue is not always the dominant eigenvalue. When the informative eigenvalue is not dominant it is a consequence of the vital dynamics of the host occurring on a timescale that is much longer than the dynamics of small outbreaks that occur when the infection does not spread very well in the host population. Such a difference in time scales seems likely to occur in other multivariate models of population dynamics.

Looking beyond the SIR model, the question also arises of whether our method of identifying appropriate variables will be practical for models with many more than two degrees of freedom. In the SIR model, the autocorrelation function of one of the state variables is often very similar to that of one of the eigendirections. This allowed us to select variables based on the criteria of how well the autocorrelation function matched up with that of the eigendirection corresponding to the informative eigenvalue. In general, as the number of variables grows we might expect that the autocorrelation function of each variable to become more strongly influenced by multiple eigenvalues. For this more challenging case, we wonder whether harmonic inversion methods [29] might be capable of estimating the values of each of the eigenvalues that strongly influence each variable from its autocorrelation function. Variables that allow the informative eigenvalue to be estimated in this manner could then be considered appropriate for tracking the distance to the threshold.

The distance to thresholds in systems will generally change over time, and our results concluded with a simple demonstration of how these changes might be tracked. In the context of infectious disease surveillance, an exciting prospect of this approach is the possibility that surveillance programs might be able determine that some change in the system is moving it closer to the epidemic threshold long before the threshold is crossed. Besides increasing awareness, such measurement may allow for management of the distance to the threshold in some systems, for example by guiding the allocation of resources to vaccination programs. In this way, infectious disease control goals could move beyond early detection of and rapid response to epidemics toward targeted prevention of epidemics. Further, tracking has the potential to measure the relenting and reversing of system dynamics in response to control goals. Finally, establishing the conditions under which statistical analysis of fluctuations in the number of infected individuals is more informative than similar analysis of susceptible individuals does not make a case against susceptible reconstruction methods [30] in distance-to-threshold studies because such methods estimate major trends in the susceptible population size rather than fluctuations around them. Rather, our result makes a case for analysis of the fluctuations in the number infected, whether estimated from readily available time series incidence data or from pathogen sequence data [31].

References

[1] Bedford T, *et al.*, 2015 Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* **523**, 217. doi:10.1038/nature14460

[2] Ferrari MJ, Grais RF, Bharti N, Conlan AJK, Bjørnstad ON, Wolfson LJ, Guerin PJ, Djibo A, Grenfell BT, 2008 The dynamics of measles in sub-Saharan Africa. *Nature* **451**, 679–684. doi: 10.1038/nature06509

[3] Anderson RM, May RM, 1992 *Infectious Diseases of Humans: Dynamics and Control.* New York, NY: Oxford University Press

[4] Fine P, Eames K, Heymann DL, 2011 "Herd immunity": A rough guide. *Clin. Infect. Dis.* **52**, 911–916. doi:10.1093/cid/cir007

[5] Fefferman NH, Naumova EN, 2015 Dangers of vaccine refusal near the herd immunity threshold: a modelling study. *Lancet Infect. Dis.* **15**, 922–926. doi:10.1016/S1473-3099(15)00053-5

[6] Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, Holmes EC, 2004 Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science* **303**, 327–332. doi:10.1126/science. 1090727

[7] Dibble CJ, O'Dea EB, Park AW, Drake JM, 2016 Waiting time to infectious disease emergence. *J. Roy. Soc. Interface* **13**, 20160540. doi:10.1098/rsif.2016.0540

[8] Wissel C, 1984 A universal law of the characteristic return time near thresholds. *Oecologia* **65**, 101–107

[9] Scheffer M, 2009 *Critical Transitions in Nature and Society.* Princeton, N.J.: Princeton University Press

[10] Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, Held H, van Nes EH, Rietkerk M, Sugihara G, 2009 Early-warning signals for critical transitions. *Nature* **461**, 53–59. doi:10.1038/ nature08227

[11] O'Regan SM, Drake JM, 2013 Theory of early warning signals of disease emergence and leading indicators of elimination. *Theor. Ecol.* **6**, 333–357. doi:10.1007/s12080-013-0185-5

[12] O'Regan SM, Lillie JW, Drake JM, 2015 Leading indicators of mosquito-borne disease elimination. *Theor. Ecol.* 1–18. doi:10.1007/s12080-015-0285-5

[13] Wiggins S, 1990 *Introduction to applied nonlinear dynamical systems and chaos.* New York, NY: Springer-Verlag

[14] Kuehn C, 2012 A Mathematical Framework for Critical Transitions: Normal Forms, Variance and Applications. *J. Nonlinear. Sci.* **23**, 457–510. doi:10.1007/s00332-012-9158-x

[15] Boerlijst MC, Oudman T, de Roos AM, 2013 Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models. *PLoS One* **8**, e62033. doi:10.1371/ journal.pone.0062033

[16] Boettiger C, Ross N, Hastings A, 2013 Early warning signals: the charted and uncharted territories. *Theor. Ecol.* **6**, 255–264. doi:10.1007/s12080-013-0192-6

[17] Dakos V, 2017 Identifying best-indicator species for abrupt transitions in multispecies communities. *Ecol. Indic.* doi:10.1016/j.ecolind.2017.10.024

[18] Bretó C, Ionides EL, 2011 Compound Markov counting processes and their applications to modeling infinitesimally over-dispersed systems. *Stoch. Process. Their Appl.* **121**, 2571–2591. doi:10.1016/j.spa. 2011.07.005

[19] Ferrari MJ, Perkins SE, Pomeroy LW, Bjørnstad ON, 2011 Pathogens, Social Networks, and the Paradox of Transmission Scaling. *Interdiscip. Perspect. Infect. Dis.* **2011**. doi:10.1155/2011/267049

[20] van Kampen NG, 2007 *Stochastic processes in physics and chemistry.* North Holland, 3 edn.

[21] Kwon C, Ao P, Thouless DJ, 2005 Structure of stochastic dynamics near fixed points. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13029–13033. doi:10.1073/pnas.0506347102

[22] Elzhov TV, Mullen KM, Spiess AN, Bolker B, 2015. minpack.lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. R package version 1.2-0

[23] He D, Ionides EL, King AA, 2010 Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *J. Roy. Soc. Interface* **7**, 271–283. doi:10.1098/rsif.2009.0151

[24] King AA, Nguyen D, Ionides EL, 2016 Statistical Inference for Partially Observed Markov Processes via the R Package pomp. *J. Stat. Softw.* **69**. doi:10.18637/jss.v069.i12

[25] King AA, *et al.*, 2016. pomp: Statistical inference for partially observed Markov processes. R package version 1.4.1.1

[26] O'Dea EB, 2017. e3bo/2017distance. doi:10.5281/zenodo.1119134. Zenodo

[27] Martinez-Bakker M, King AA, Rohani P, 2015 Unraveling the transmission ecology of polio. *PLoS Biol.* **13**, e1002172. doi:10.1371/journal.pbio.1002172

[28] Bretó C, He D, Ionides EL, King AA, 2009 Time series analysis via mechanistic models. *Ann. Appl. Stat.* **3**, 319–348. doi:10.1214/08-AOAS201

[29] Mandelshtam VA, Taylor HS, 1997 Harmonic inversion of time signals and its applications. *J. Chem. Phys.* **107**, 6756–6769. doi:10.1063/1.475324

[30] Bobashev GV, Ellner SP, Nychka DW, Grenfell BT, 2000 Reconstructing susceptible and recruitment dynamics from measles epidemic data. *Math. Popul. Stud.* **8**, 1–29. doi:10.1080/08898480009525471

[31] Rasmussen DA, Ratmann O, Koelle K, 2011 Inference for Nonlinear Epidemiological Models Using Genealogies and Time Series. *PLoS Comput. Biol.* **7**, e1002136. doi:10.1371/journal.pcbi.1002136