


RESEARCH

Open Access



Forecasting infectious disease emergence subject to seasonal forcing

Paige B. Miller^{1,2*} , Eamon B. O'Dea^{1,2}, Pejman Rohani^{1,2,3} and John M. Drake^{1,2}

*Correspondence:

paige.miller@uga.edu

¹University of Georgia, Odum School of Ecology, 140 E. Green Street, Athens, USA

²Center for the Ecology of Infectious Diseases, University of Georgia, Athens, USA

Full list of author information is available at the end of the article

Abstract

Background: Despite high vaccination coverage, many childhood infections pose a growing threat to human populations. Accurate disease forecasting would be of tremendous value to public health. Forecasting disease emergence using early warning signals (EWS) is possible in non-seasonal models of infectious diseases. Here, we assessed whether EWS also anticipate disease emergence in seasonal models.

Methods: We simulated the dynamics of an immunizing infectious pathogen approaching the tipping point to disease endemicity. To explore the effect of seasonality on the reliability of early warning statistics, we varied the amplitude of fluctuations around the average transmission. We proposed and analyzed two new early warning signals based on the wavelet spectrum. We measured the reliability of the early warning signals depending on the strength of their trend preceding the tipping point and then calculated the Area Under the Curve (AUC) statistic.

Results: Early warning signals were reliable when disease transmission was subject to seasonal forcing. Wavelet-based early warning signals were as reliable as other conventional early warning signals. We found that removing seasonal trends, prior to analysis, did not improve early warning statistics uniformly.

Conclusions: Early warning signals anticipate the onset of critical transitions for infectious diseases which are subject to seasonal forcing. Wavelet-based early warning statistics can also be used to forecast infectious disease.

Keywords: Disease forecasting, Seasonality, Critical transition

Background

Improvement of methods for forecasting infectious disease dynamics is of tremendous value to public health and the global economy [1]. Disease forecasting is challenging because transmission is subject to complex interactions among a variety of independent actors, non-linearity, noise, and seasonality driven by age [2] and spatial structure [3], susceptible depletion [4, 5], environmental variability [6], and behavior [5, 7]. Despite these challenges, a number of approaches have been developed along distinctly different lines: (i) sequential Monte-Carlo methods (e.g., [8]), (ii) data assimilation techniques inspired by numerical weather prediction [9, 10], and (iii) “Wisdom of the Crowd” approaches [11] among others. Monte-Carlo methods aim to estimate unobserved states in Markov chains (such as the size of the infectious population) while simultaneously estimating unknown parameters [8]. In contrast, ensemble Kalman adjusted filters are commonly used in data assimilation and take advantage of incoming information by iteratively

updating an ensemble of models to contract variance [9, 10]. Lastly, “Wisdom of the Crowd” approaches incorporate experts’ experience and knowledge [11]. These studies demonstrate that accurately forecasting the trajectory of infectious disease incidence is possible. In contrast, we are in the earlier stages of theory and method development for anticipating the emergence of infectious disease [12].

Emerging infectious diseases of the past century include three major influenza pandemics, Ebola, HIV, and Zika. The increasing frequency of emergence by novel pathogens is usually attributed to changes in socio-economic, environmental and ecological factors [13]. All of these are now combated with highly developed diagnostics, therapeutics, and behavioral interventions. If they could have been anticipated, behavioral interventions could have been deployed earlier, preventing loss of life and reducing spread even while medical countermeasures were under development. Additionally, some countries have reported resurgence of vaccine preventable diseases (e.g., pertussis, measles) despite high rates of vaccination [14]. Many candidate hypotheses have been advanced to explain the resurgence of vaccine preventable diseases including changing vaccination schedule and composition, changing immune acquisition, and evolution of the pathogen (reviewed in [14]). The economic and health benefits of the increased preparedness that could be achieved by more advanced warning of disease emergence and re-emergence would be substantial.

In the past decade, researchers have studied an alternative, model-independent forecasting approach which summarizes the patterns of fluctuations in a system to infer whether it is close to a critical transition [15–18]. A critical transition is a sudden and large change in the state of a dynamical system [19]. Statistical signals primarily result from a phenomenon known as *critical slowing down* (CSD), which is a signature of a second-order phase transition and refers to the increased relaxation time for a near-critical system to reach equilibrium following a perturbation compared with a system that is far from critical [20]. Critical slowing down occurs in a wide range of natural tipping points, including large-scale climate shifts [15], population collapse [17], and lake eutrophication [21]. O’Regan and Drake [22] proposed that statistical early warning signals (EWS) may also precede the onset of sustained transmission in a contagion process where the approach to disease emergence is marked by a slow drift to criticality. Because of the key assumption of slow drift, critical slowing down has primarily been considered in non-periodic systems [23]. But, many diseases exhibit periodic forcing due to seasonal variations in climate [24, 25], human behavior [26], or immune function [27]. Because seasonal forcing results in periodic sojourns near criticality, we wondered if standard approaches to detecting critical slowing down might therefore fail in systems such as resurgent childhood infections forced by the annual calendar of school terms [5, 28].

We performed a simulation study to investigate how periodic forcing might interfere with the detection of critical slowing down in directly-transmitted disease systems. A recent review of resilience in ecological systems noted that early warning signals are difficult to detect in systems with intrinsic cycles [23]. In this case, [23] recommend first removing periodic trends in the time series through seasonal detrending and then performing analysis on the residuals. However, prior to reaching criticality, transmission systems may not exhibit a reliably periodic pattern. Thus, it’s possible that periodic detrending or differencing may corrupt the signal of CSD and unintentionally reduce the reliability of the early warning signals. Nonetheless, the robustness of differencing and

detrending have never been validated in the context of early warning signals of epidemic transitions. To fill this gap, we studied the detectability of CSD prior to disease emergence in simulations. To address the issue of periodic forcing, we analyzed our data both with and without first removing seasonal trends via seasonal detrending and differencing.

Additionally, we propose and examine two new statistics based on components of the frequency domain as alternatives to seasonal detrending and differencing. In the frequency domain of time series without periodic cycles, CSD manifests as spectral reddening, or the dominance of low frequencies prior to a transition [29]. Spectral reddening of the Fourier spectra was shown to be a reliable indicator of climate tipping points [29] and housing market changes [30, 31]. Our concern is that for a periodic time series, the periodic signal will overwhelm that of spectral reddening such that CSD is difficult to detect. Wavelet analysis, a non-parametric approach to identifying periodic components of the system over time, might allow us to identify the frequencies that are most sensitive to CSD. In this paper, we show that signals of CSD can still be found in systems with periodic behavior and present methods for forecasting emergence of seasonally forced infectious diseases based on spectral reddening of the wavelet spectrum.

Methods

SIR model

To simulate the dynamics of an immunizing pathogen with seasonal transmission (e.g., pertussis, measles) [7], we considered a seasonally-forced stochastic Susceptible-Infected-Recovered (SIR) model. Transitions between states occur with rates shown in Table 1 where $N = S + I + R$ and the mean field equations are

$$\begin{aligned} \frac{d(S/N)}{dt} &= -\beta(t)(S/N)(I/N) - \xi(S/N) + \mu(1 - S/N) \\ \frac{d(I/N)}{dt} &= \beta(t)(S/N)(I/N) + \xi(S/N) - \gamma(I/N) - \mu(I/N) \\ \frac{d(R/N)}{dt} &= \gamma(I/N) - \mu(R/N). \end{aligned} \tag{1}$$

Parameter definitions and values are given in Table 2. We modeled the contagion process as a Markov chain using Gillespie’s direct method [32]. A Markov chain is a stochastic process with the property that the future state of the system is dependent only on the present state of the system and conditionally independent of all past states (i.e., the system has a memoryless property).

Our model is slowly forced through a transition from subcritical (average R_0 during a year less than 1) to supercritical (average R_0 during a year greater than 1) by linearly increasing the average transmission rate, β_0 , over time. Increasing transmission, as formulated in this model, can be envisioned as changes in behavior between individuals

Table 1 Transition rules for stochastic SIR model

Event	$(\Delta S, \Delta I, \Delta R)$	Rate
Birth of a susceptible	$(1, 0, 0)$	μN_0
Death of a susceptible	$(-1, 0, 0)$	μS
Infection	$(-1, 1, 0)$	$\frac{\beta S I}{N} + \xi S$
Death of an infective	$(0, -1, 0)$	μI
Recovery of an infective	$(0, -1, 1)$	γI
Death of a removed	$(0, 0, -1)$	μR

Table 2 SIR model parameters

Symbol	Definition	Value
β	Transmission rate	varied
ξ	Importation rate	$6 \cdot 10^{-5}$ per year
$1/\gamma$	Infectious period	14 days
$1/\mu$	Lifespan	70 years
N_0	Initial population size	$2 \cdot 10^5$ individuals

in a population leading to higher contact rates over time or pathogen adaptation. We assumed frequency-dependent transmission but note that in this case, the difference between density-dependent and frequency-dependent transmission is negligible because the population size fluctuates minimally (due to demographic stochasticity).

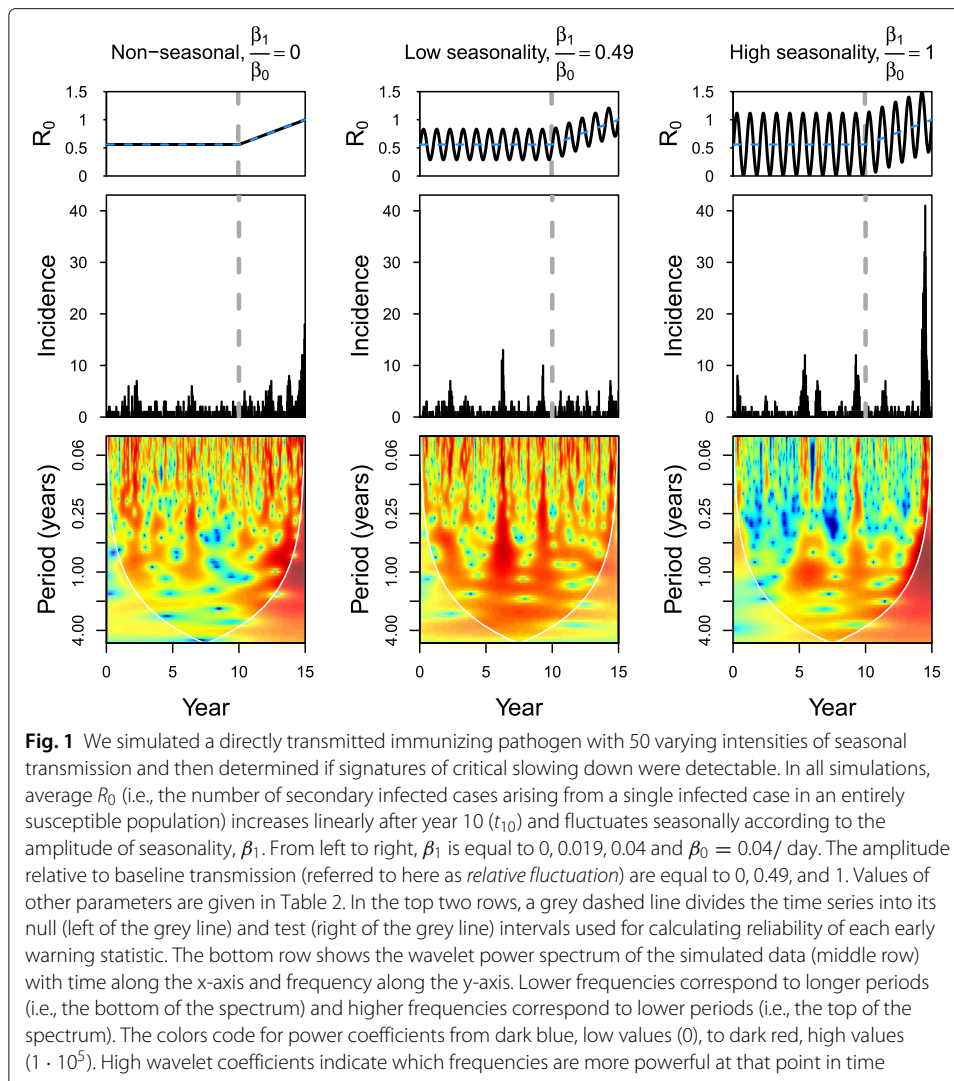
In our model, susceptible individuals are infected through contact with infected individuals within the population at a time-varying rate $\beta(t)$. To represent a realistic scenario such that low levels of contact with outside populations also occur, we assumed that susceptible individuals can also become infected at a constant rate ξ [22]. This assumption signifies that the equilibrium size of the infectious population is non-zero, allowing for stuttered chains of transmission even when the system is sub-critical (i.e., when $R_0 < 1$). Sinusoidal transmission follows the function of time $\beta(t) = \beta_0(t) + \beta_1 \sin(2\pi t/365)$ where t is in units of days. We assume the time-localized average rate of transmission (β_0) begins to increase linearly ten years later (i.e., $t_{\text{start}} = 10 \cdot 365$ days), so that

$$\beta_0(t) = \begin{cases} \beta_k, & t < t_{\text{start}} \\ \beta_k + b(t - t_{\text{start}}), & t \geq t_{\text{start}} \end{cases} \quad (2)$$

where $\beta_k = 0.04$ is the initial transmissibility and $b = 1.7242 \times 10^{-5}/\text{day}$ is the daily rate at which transmission increases. These parameters cause the critical value of R_0 to be reached at year fifteen (five years after t_{start}). Before t_{start} , the average R_0 , calculated as $\beta/(\gamma + \mu)$, is approximately 0.56 ($\beta_0 = 0.04$). To explore the effect of seasonality, we varied β_1 linearly from 0 to 0.04 with 50 levels and simulated each level 250 times. For sinusoidal seasonal forcing, we present and discuss intensity of *relative fluctuations*, i.e. β_1/β_0 , for ease of interpretation. For example, if $\beta_1/\beta_0 = 1$ then seasonality would cause R_0 to fluctuate between 0 and twice that of baseline transmission (Fig. 1).

We note that seasonality in transmission is often modeled as $\beta = \beta_0(1 + \beta_1 \cos(2\pi t/p))$ where β_1 can be interpreted as the amplitude of fluctuations relative to the average rate of transmission (see [33]). In this formulation, though, if β_0 were to increase (i.e., as formulated in our model) then the size of oscillations around the increasing average transmission, β_0 , would increase as well. By contrast, our parameterization induces a constant amplitude for seasonality. The relative amplitude of fluctuation in our parameterization is simply given by β_1/β_0 .

For validation of our results with sinusoidal forcing, we also analyzed the impact of seasonal transmission on EWS using square wave forcing, term-time forcing [28], and monthly averaged rates [5]. These other seasonality functions were defined by a set of 52 repeating deviations from the mean. The magnitude of seasonality for these models was controlled by scaling the deviations such that the deviation below the mean was at most equal to a given fraction of the mean.



We allowed a burn-in time of ten years before analyzing simulated incidence to remove the effects of transient behavior. For comparison with typical incidence data, cases were aggregated as number of individuals entering the removed class each week denoted by X_t . We assumed no under-reporting. Because we assumed individuals within the population could become infected from those outside the population, sparking transmission with initially infected individuals was not necessary. Thus, our initial conditions for state variables were $S(0) = N(0)$, $I(0) = 0$, $R(0) = 0$. All simulations were performed in R using packages *spaero* [34] and *pomp* [35]. Code to reproduce the simulations is available online at <https://github.com/project-aero>.

Calculation of early warning signals

We defined the null and test intervals to compare trends in EWS when the system is and is not approaching criticality. The null interval is defined when β_0 remains constant as compared with the test interval when β_0 is increasing and therefore approaching criticality (Fig. 1). Following separation into the null and test intervals, we calculated early

warning statistics both with and without first removing seasonal trends via (1) seasonal decomposition by Lowess [36] and (2) seasonal differencing where the time series at X_t is subtracted from X_{t-52} when observations are measured weekly.

We analyzed two statistics for anticipating disease emergence based on the wavelet spectrum: (1) *wavelet filtered reddening* and (2) *wavelet spectral reddening*. We defined wavelet filtered reddening, $\bar{W}_{t,j_1j_2}^2$, as any quantity proportional to the variance of the time series after it is filtered to include only its components in a wavelet transform's frequency domain with scales ranging from s_{j_1} to s_{j_2} . We calculated it according to

$$\bar{W}_{t,j_1j_2}^2 = \frac{0.34T\delta_t}{\lambda} \sum_{j=j_1}^{j_2} \frac{|W_t(s_j)|^2}{s_j}, \tag{3}$$

where T is the total number of observations in the times series, δ_t is the difference in time between observations, s_j is a wavelet scale, $W_t(s_j)$ is a wavelet transform of the observation time series X with localized time index t and scale s_j , and λ is the ratio of the Fourier period of the wavelet function to its scale. The Fourier period of the wavelet function is the period at which the Fourier transform of the wavelet function peaks.

To see that (3) is proportional to the average variance of the time series within a frequency band, note that the right hand side of (3) differs from that of Eq. 24 of Ref. [37] by a constant factor only. Equation (3) may also be viewed as a partial summation over elements of the bias-corrected power spectrum of Ref. [38] multiplied by a constant factor. Thus in our plot of the wavelet power spectrum, we plot the values of $[0.34T\delta_t|W_t(s_j)|^2] / [s_j\lambda\sigma^2]$ for all scales s_j and times t , where σ^2 is the estimated variance of the time series X . Because we look at trends in the wavelet power spectrum, the constant factors in our equations have no effect on our results and are included simply because they are included in the calculations of the software package we used.

Next we explain exactly how we calculated the wavelet transform. The wavelet transform was calculated to satisfy

$$W_t(s) = \sum_{t'=1}^T (X_{t'} - \bar{X}) \left(\frac{\delta_t}{s}\right)^{0.5} \psi_0^* \left[\frac{(t' - t)\delta_t}{s}\right], \tag{4}$$

\bar{X} is the mean of the time series, and ψ_0^* is the complex conjugate of the Morlet wavelet function. The Morlet wavelet function satisfies

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\eta^2/2} \tag{5}$$

where ω_0 is the nondimensional frequency. We used typical parameters to calculate the wavelet transform. We set ω_0 to 6, and the sequence of scales used started at $2\delta_t$ and increased geometrically with 12 equal steps per octave until it reached about one third the duration of the time series. These parameters correspond to the defaults in the R package `biwavelet` [39], which we used for all wavelet calculations. To choose the parameters j_1 and j_2 in Eq. (3), we conducted a grid search and used AUC to determine which j_1 and j_2 resulted in good separation of the distributions of $\bar{W}_{t,j_1j_2}^2$ calculated from emergence and non-emergence simulations.

Our second wavelet statistic, wavelet spectral reddening, was defined as the median scale of the wavelet spectral density at a given time index t . We denote this statistic as

$s_{\text{median}}(t)$, and we calculate it according to

$$s_{\text{median}}(t) = \min \left\{ s_j \text{ such that } \bar{W}_{t,1j}^2 / \bar{W}_{t,1j_{\text{max}}}^2 > 0.5 \right\}, \tag{6}$$

where the indices of the increasing wavelet scales s_j run from 1 to j_{max} . A spectral red- dening statistic of the Fourier power spectrum was calculated in an analogous manner in Ref. [31]. This statistic is intended to quantify any increasing dominance of the low fre- quency components of the power spectrum associated with the decreasing stability of an equilibrium.

The other early warning statistics were computed according to the formulas in Table 3, in which the weights $w_{t,t'}$ depend on the smoothing kernel and bandwidth. We used both a uniform kernel and a Gaussian kernel. The equation for the weights using the uniform kernel is

$$w_{t,t'} = \begin{cases} 1/N_t, & |t - t'| < b, \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

where the normalization constant $N_t = \sum_{j=\max(1,t-b+1)}^{\min(T,t+b-1)} 1$. The equation for the weights using the Gaussian kernel is

$$w_{t,t'} = f(t - t', b) / N_t \tag{8}$$

where f satisfies $f(x, b) = \exp(-x^2 / (2b^2)) / [b\sqrt{2\pi}]$ and $N_t = \sum_{t'=1}^T f(t - t', b)$.

Following [17, 22, 40, 41] the other early warning statistics were computed according to the formulas in Table 3. Moving window statistics of time series X (centered on index t) were calculated with a uniform kernel and the bandwidth (parameter b in Eq. 7) was 100. If data were detrended using STL [36], the equations for the moving window statistics were slightly modified. The STL estimate of the trend was used as mean_t in Table 3. Then, the STL estimate of the irregular component of the time series was substituted in for the residuals $(X_t - \text{mean}_t)$ in Table 3.

We determined the performance of each EWS using the AUC statistic as follows. We calculated the association between each indicator time series and time using Kendall's rank correlation coefficients. Since we had 250 simulations, we generated a distri- bution of correlation coefficients for each indicator when the system was and was not approaching disease emergence (i.e., from the test or null interval). The amount of overlap between the distributions was calculated with the AUC statistic: $\text{AUC} = [r_{\text{test}} - (n_{\text{test}} + 1) / 2] / (n_{\text{test}} n_{\text{null}})$ where r_{test} is the sum of the ranks of test coefficients in a combined set of test and null coefficients (where the lower numbers have lower ranks),

Table 3 Formulas for moving window statistics

Statistic	Formula
Mean _t	$\sum_{t'=1}^T w_{t,t'} X_{t'}$
Variance _t	$\sum_{t'=1}^T w_{t,t'} (X_{t'} - \text{mean}_{t'})^2$
(Variance convexity) _t	$\text{variance}_t - \text{variance}_{t-1}$
Autocovariance _t	$\sum_{t'=(\text{lag}+1)}^T w_{t,t'} (X_{t'} - \text{mean}_{t'}) (X_{t'-\text{lag}} - \text{mean}_{t'-\text{lag}})$
Autocorrelation _t	$\text{autocovariance}_t / (\text{variance}_t \times \text{variance}_{t-\text{lag}})^{0.5}$
(Decay time) _t	$-\text{lag} / (\log \min(\max(\text{autocorrelation}_t, 0), 1))$
(Index of dispersion) _t	$\text{variance}_t / \text{mean}_t$
(Coefficient of variation) _t	$(\text{variance}_t)^{0.5} / \text{mean}_t$
Skewness _t	$\sum_{t'=1}^T w_{t,t'} (X_{t'} - \text{mean}_{t'})^3 / (\text{variance}_t)^{1.5}$
Kurtosis _t	$\sum_{t'=1}^T w_{t,t'} (X_{t'} - \text{mean}_{t'})^4 / (\text{variance}_t)^2$

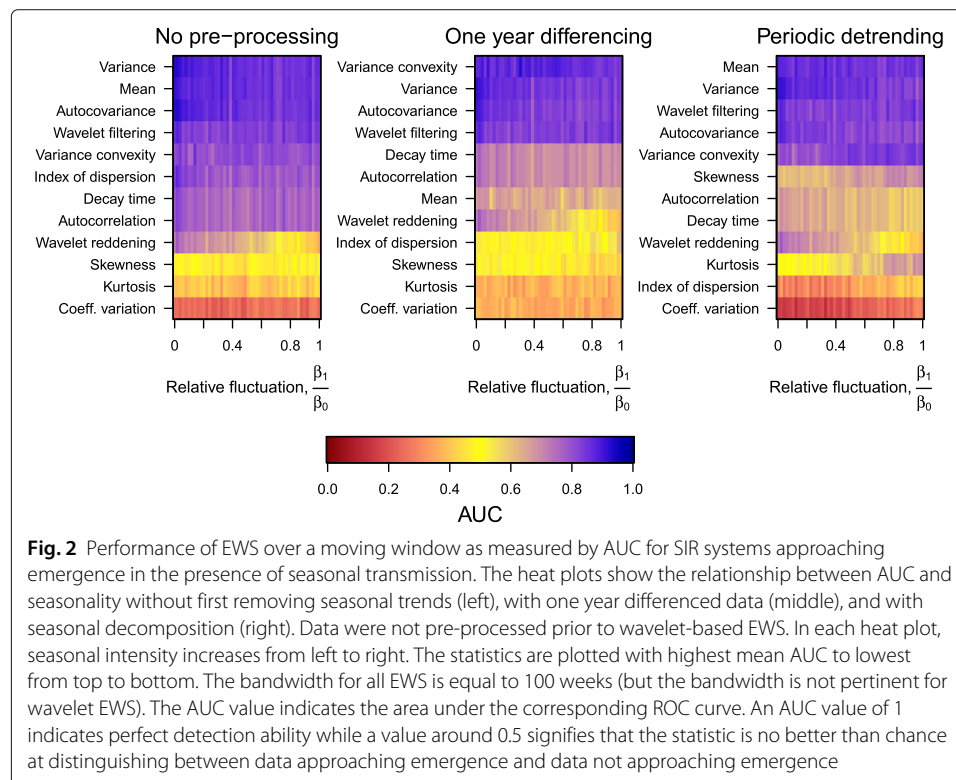
n_{test} is the number of test coefficients, and n_{null} is the number of null coefficients. The AUC of an EWS is the probability that a randomly chosen test coefficient is higher than a randomly chosen null coefficient [42].

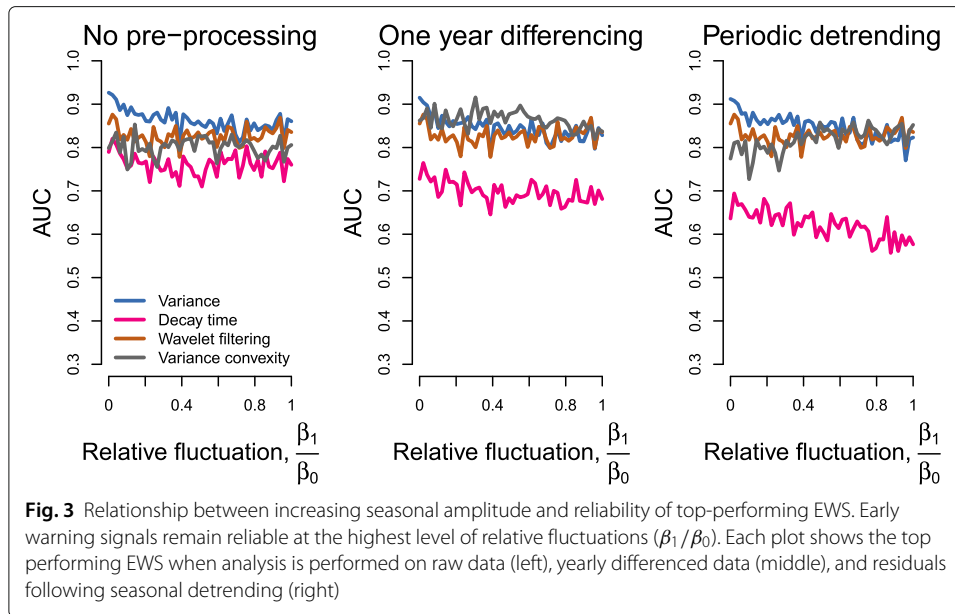
Results

We documented signals of critical slowing down when transmission was subject to periodic variation, addressing a current gap in studies of early warning signals for critical transitions (Fig. 2). The AUCs of most early warning statistics were negatively associated with increasing seasonal amplitude (Figs. 2 and 3). When transmission was not subject to seasonal forcing, the most reliable statistics (mean, variance-based, and wavelet filtering), achieved AUCs above 0.85 regardless of how data were pre-processed. For simulations with the highest amplitude of seasonal transmission, AUCs for the mean, variance, autocovariance, and wavelet filtering decreased by 0.05, 0.07, 0.06, and 0.02 compared to the non-seasonal simulations, respectively.

Skewness and kurtosis were poor indicators of disease emergence compared with variance-based statistics but detrending the time series prior to analysis improved the AUCs slightly when transmission was subject to moderate or high levels of seasonal forcing (Right panel, Fig. 2). For other EWS (i.e., decay time, autocorrelation, and index of dispersion), pre-processing the time series prior to analysis worsened or had little impact on the AUCs (Middle and left panels, Fig. 2).

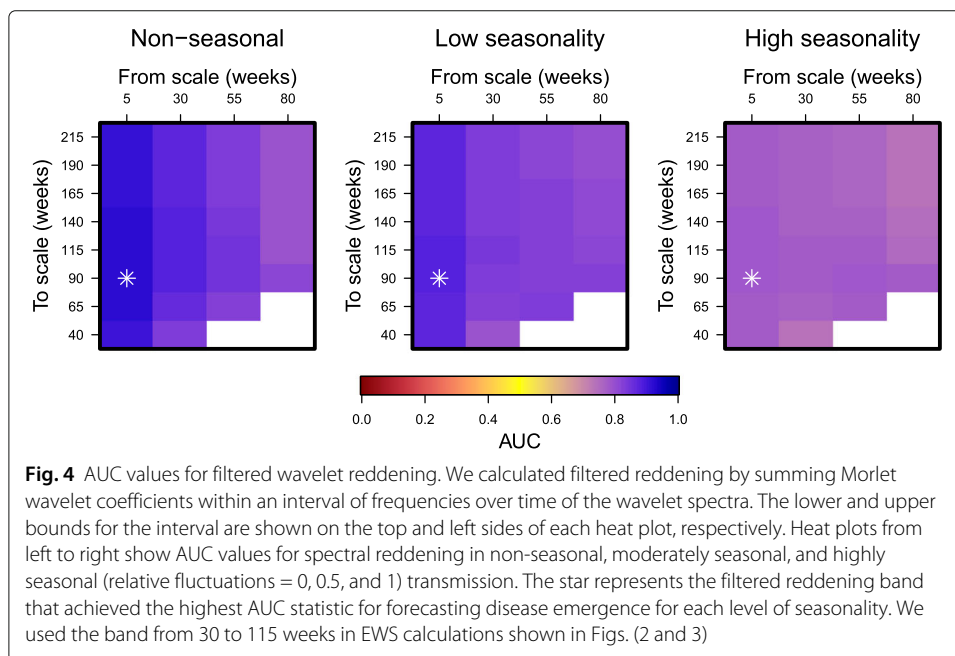
In simulations with non-seasonal transmission, wavelet reddening and filtering performed similarly to variance-based early warning statistics (Figs. 2 and 3). However, as the amplitude of seasonality increased, the AUC of wavelet reddening decreased drastically





(from 0.72 to 0.35) with increasing seasonal amplitude whereas the AUCs of wavelet filtering remained relatively constant as seasonal amplitude increased. For all levels of seasonality, the AUCs of wavelet filtering were highest when the frequency band ranged from 5 to 90 weeks (Fig. 4).

Early warning signals performed similarly regardless of the function of seasonal forcing (sinusoidal, term-time [28], monthly averaged [5], and square-wave seasonal forcing) (Additional file 1). For each type of seasonal forcing, the most reliable early warning statistics were variance, variance convexity, autocovariance, and wavelet filtering (Additional file 1). Conventional early warning signals were sensitive to the choice of



bandwidth. When the bandwidth was 150 weeks long, rather than 100, early warning signal reliability decreased slightly (Additional file 1). The choice of window shape (i.e., uniform or Gaussian) was less important and resulted in similar AUC values for moving window statistics.

Discussion

A principal challenge in infectious disease epidemiology is predicting epidemics. Prediction of epidemics is made difficult by a combination of non-linearities driving the system. Early warning signals for disease emergence are potentially useful for predicting emerging infectious diseases because they do not require a detailed understanding of the system's drivers. For infectious diseases without a seasonal trend (e.g., STDs), generic statistical signals are present before critical transitions [22]. However, seasonal variation alters the spread and persistence of many diseases [33] and has been noted as a limitation to using early warning signals [23]. Understanding how seasonal forcing, a common feature of disease dynamics, affects the trends in generic leading indicators of stability is therefore important for accurately inferring whether a disease is at risk of emergence.

We conducted a simulation study of early warning signals for emergence of infectious diseases. To understand the relationship between seasonality and the predictability of disease emergence, we varied the amplitude of seasonal transmission. We showed numerically that critical slowing down anticipates transitions in periodic systems. Additionally, we proposed and measured the performance of two new leading indicators, wavelet spectral reddening and filtering, compared to conventional early warning signals. Wavelet filtering, calculated by summing specific coefficients of wavelets across the time series, was among the top performing EWS. Our second statistic, wavelet reddening, was less reliable than wavelet filtering as the level of seasonal transmission increased. The advantage to using wavelet-based statistics is that they do not require a choice of bandwidth. Although wavelet filtering requires a choice of frequency band, we showed that the optimal choice of the frequency band was invariant to three levels of seasonal forcing. Furthermore, the AUCs for wavelet filtering were generally highest with the largest frequency bands. This suggests that all periods might be used and therefore this statistic would be parameter-free.

Our study's purpose was to characterize the impact of seasonality on early warning signals for disease emergence of an immunizing pathogen such as childhood infections. To do this, we varied only the amplitude and held constant the frequency of the periodic component of transmission. We showed numerically that early warning signals can be reliable even in systems with highly seasonal transmission. A recent study [43] proposed special methods for calculating early warning signals when the time scale of the system is similar to that of the forcing. In our model, the time scale of the system could be characterized by the decay time of the number of cases per week and the time scale of the forcing is the period of the seasonal cycle in R_0 . To characterize the timescale of the model as a function of R_0 , we calculated the ensemble decay time based on the ensemble of X_t comprising the simulation replicates of our model. We used simulations with a relative fluctuations of zero and an initial β of 0.04. We found that the estimated decay time averaged 0.017 years in the null interval and rose to 0.14 years over the course of the test interval. Therefore, within the test interval the timescale of the system was similar to that of the annual

forcing. There are many differences between our model and the model analyzed in reference [43]. For example, stochastic effects (e.g., imported cases) are more important in our model and the periodic forcing does not affect our observations as directly. Further analysis of models similar to ours could clarify when seasonality of infectious diseases could pose a major problem for EWS of disease emergence.

For our model, estimating and removing a periodic trend prior to EWS analysis did not improve prediction uniformly among statistics. This was not entirely surprising because the seasonal signal was not apparent in many time series. Rather, sporadically spaced small outbreaks comprised the dynamics. Therefore, periodic detrending and differencing introduced artificial patterns in the time series. In summary, even in systems where transmission is highly seasonal, and a seasonal trend is apparent in a part of the observation window, seasonal detrending is often disadvantageous for forecasting approaches using early warning signals.

In our study, we held constant the period of forcing and other parameters that sometimes exhibit periodicity (e.g., rate of imported infectious individuals, recovery rate, or demographic rates). We also kept the rate of emergence constant among all simulations. A recent study examining the effects that rate of forcing has on the strength of EWS found that trends in EWS were more difficult to assess in systems with faster rates of change [44]. Thus, we would expect that EWS be less reliable for quickly emerging diseases. Here, we focused on periodicity in transmission because it is the driving variable causing emergence and the other parameters (e.g., demographic and recovery rates) are not typically periodic in human populations. Lastly, as is typical in seasonally forced models of infectious diseases [33] and studies of periodicity in natural systems [43], we focused on seasonality using a sine function. However, we showed that our results were largely invariant to different types of seasonality functions (term-time [5], monthly-averaged [28], and square-wave forcing).

Variance and early warning signals based on spectral properties are beginning to emerge as the most reliable indicators of upcoming transitions. In O'Regan and Drake [22], approach to disease endemicity was much more difficult to forecast than elimination but variance was the best predictor in SIR and SIS systems with immigration. In vector-borne disease transmission, elimination was best predicted by variance and coefficient of variation [45]. Additionally, disturbances in food chain dynamics causing trophic cascades (i.e., the release of prey without suppression by predators) were strongly associated with changes in variance and spectral density of population sizes [21]. These experimental and theoretical results are beginning to reveal the most robust, generic indicators across a wide range of ecological and epidemiological systems.

Conclusions

Forecasting emerging infectious diseases is of tremendous value to public health and society. Advances in surveillance systems and data science have generated the expectation that early warning systems are not only feasible but necessary tools to combat the spread of infectious diseases [46]. Critical slowing down was previously shown to signal disease emergence without seasonal transmission [22]. Here, we showed that critical slowing down also precedes disease emergence with seasonal transmission. Additionally,

we found that statistics based on the wavelet spectrum are robust signals of disease emergence. Future research should aim to document critical slowing down in experimental or surveillance data to estimate the potential forecast horizon that early warning signals provide.

Additional file

Additional file 1: Supplementary analysis showing effects of different forms of seasonal forcing, window size, and window shape on performance of early warning signals. (PDF 151 kb)

Acknowledgements

This work was done on the Olympus High Performance Compute Cluster located at the Pittsburgh Supercomputing Center at Carnegie Mellon University, which is supported by National Institute of General Medical Sciences Modeling Infectious Disease Agent Study (MIDAS) Informatics Services Group grant 1U24GM110707.

Funding

Research funded by the National Institute of General Medical Sciences of the National Institutes of Health (Award Number U01GM110744).

Availability of data and materials

All code for simulations and analysis will be made available at <https://github.com/project-aero> following publication.

Authors' contributions

All authors jointly formulated the mathematical modelling approach. PM and EO analyzed the data. PM drafted the manuscript and all authors revised the manuscript. All other authors gave comments on the revised manuscript and approved the final version of the manuscript.

Ethics approval and consent to participate

This study does not involve human participants, human data or human tissue.

Consent for publication

This study does not contain any individual person's data in any form.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹University of Georgia, Odum School of Ecology, 140 E. Green Street, Athens, USA. ²Center for the Ecology of Infectious Diseases, University of Georgia, Athens, USA. ³Department of Infectious Diseases, University of Georgia, Athens, USA.

Received: 3 March 2017 Accepted: 23 August 2017

Published online: 06 September 2017

References

- Smith RD, Keogh-Brown MR, Barnett T, Tait J. The economy-wide impact of pandemic influenza on the UK: a computable general equilibrium modelling experiment. *BMJ (Clin Res ed)*. 2009;339(nov19_1):4571. doi:10.1136/bmj.b4571.
- Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ, Jr IL, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Germann T, Kadau K, Jr IL, Macken C, Ferguson N, Cummings D, Fraser C, Cajka J, Cooley P, Wallinga J, Teunis P, Cauchemez S, Boelle P, Donnelly C, Ferguson N, Thomas G, Pitman R, Cooper B, Trotter C, Gay N, Edmunds W, Lipsitch M, Cohen T, Cooper B, Robins J, Ma S, Riley S, Fraser C, Donnelly C, Ghani A, Abu-Raddad L, Dye C, Gay N, Wu J, Riley S, Fraser C, Leung G, Vynnycky E, Edmunds W, Hatchett R, Mecher C, Lipsitch M, Carrat F, Luong J, Lao H, Salle A, Lajaunie C, Wallinga J, Teunis P, Kretzschmar M, Fenton K, Korovessis C, Johnson A, McCadden A, McManus S, Garnett G, Hughes J, Anderson R, Stoner B, Aral S, Gregson S, Nyamukapa C, Garnett G, Mason P, Zhuwau T, Brewer D, Garrett S, Farrington C, Kanaan M, Gay N, Farrington C, Whitaker H, Edmunds W, O'Callaghan C, Nokes D, Edmunds W, Kafatos G, Wallinga J, Mossong J, Beutels P, Shkedy Z, Aerts M, Damme PV, Mikolajczyk R, Akmatov M, Rastin S, Kretzschmar M, Mikolajczyk R, Kretzschmar M, Akakzia O, Friedrichs V, Edmunds W, Mossong J, Hilbe J, Hahsler M, Grün B, Hornik K, Hahsler M, Hornik K, Wood S, Hens N, Aerts M, Shkedy Z, Kimani PK, Kojouharova M, Diekmann O, Heesterbeek J, Bansal S, Pourbohloul B, Meyers L, Eubank S, Guclu H, Kumar V, Marathe M, Srinivasan A, Fry J, Marsden P, Fu YC. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med*. 2008;5(3):74. doi:10.1371/journal.pmed.0050074.
- Keeling MJ. The effects of local spatial structure on epidemiological invasions. *Proc R Soc Lond Ser B Biol Sci*. 1999;266(1421):859–67.

4. Koopman JS, Longini IM. The Ecological Effects of Individual Exposures and Nonlinear Disease Dynamics in Populations. *Am J Public Health*. 1994;84(5):836–42.
5. Metcalf CJ, Björnstad ON, Grenfell BT, Andreasen V. Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen. *Proc Biol Sci/R Soc*. 2009;276(1676):4111–8. doi:10.1098/rspb.2009.1058.
6. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci USA*. 2009;106(9):3243–8. doi:10.1073/pnas.0806852106.
7. He D, Earn DJD. The cohort effect in childhood disease dynamics. *J R Soc Interface/R Soc*. 2016;13(120):20160156. doi:10.1098/rsif.2016.0156.
8. Ionides EL, Bretó C, King AA. Inference for nonlinear dynamical systems. *Proc Natl Acad Sci USA*. 2006;103(49):18438–43. doi:10.1073/pnas.0603181103.
9. Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. *Proc Natl Acad Sci USA*. 2012;109(3):20425–30. doi:10.1073/pnas.1208772109.
10. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. Real-time influenza forecasts during the 2012–2013 season. *Nat Commun*. 2013;4:2837. doi:10.1038/ncomms3837.
11. Li EY, Tung CY, Chang SH. The wisdom of crowds in action: Forecasting epidemic diseases with a web-based prediction market system. *Int J Med Inform*. 2016;92:35–43. doi:10.1016/j.ijmedinf.2016.04.014.
12. Brett TS, Drake JM, Rohani P. Anticipating the emergence of infectious diseases. *J R Soc Interface*. 2017;14(132). doi:10.1098/rsif.2017.0115.
13. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. Global trends in emerging infectious diseases. *Nature*. 2008;451(7181):990–3. doi:10.1038/nature06536.
14. Jackson DW, Rohani P. Perplexities of pertussis: recent global epidemiological trends and their potential causes. *Epidemiol Infect*. 2014;142(4):672–84.
15. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *Proc Natl Acad Sci*. 2008;105(38):14308–12. doi:10.1073/pnas.0802430105.
16. Guttal V, Jayaprakash C. Changing skewness: an early warning signal of regime shifts in ecosystems. *Ecol Lett*. 2008;11(5):450–60. doi:10.1111/j.1461-0248.2008.01160.x.
17. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456–9.
18. Dai L, Vorselen D, Korolev KS, Gore J. Generic Indicators for Loss of Resilience Before a Tipping Point Leading to Population Collapse. *Science*. 2012;336(6085):1175–7.
19. Kuehn C. A Mathematical Framework for Critical Transitions: Normal Forms, Variance and Applications. *J Nonlinear Sci*. 2013;23(3):457–510. doi:10.1007/s00332-012-9158-x.
20. Strogatz SH. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering. Studies in Nonlinearity*. Boulder: Westview Press; 1994.
21. Carpenter SR, Cole JJ, Pace ML, Batt R, Brock WA, Cline T, Coloso J, Hodgson JR, Kitchell JF, Seekell DA, Smith L, Weidel B. Early warnings of regime shifts: a whole-ecosystem experiment. *Science*. 2011;332(6033):1079–82. doi:10.1126/science.1203672.
22. O'Regan SM, Drake JM. Theory of early warning signals of disease emergence and leading indicators of elimination. *Theor Ecol*. 2013;6(3):333–57.
23. Scheffer M, Carpenter SR, Dakos V, van Nes EH. Generic Indicators of Ecological Resilience: Inferring the Chance of a Critical Transition. *Ann Rev Ecol Evol Syst*. 2015;46(1):145–67. doi:10.1146/annurev-ecolsys-112414-054242.
24. Pascual M, Rodó X, Ellner SP, Colwell R, Bouma MJ. Cholera Dynamics and El Niño–Southern Oscillation. *Science*. 2000;289(5485):1766–9.
25. van Panhuis WG, Choisy M, Xiong X, Chok NS, Akarasevi P, Iamsirithaworn S, Lam SK, Chong CK, Lam FC, Phommassak B, Vongphrachanh P, Bouaphanh K, Rekol H, Hien NT, Thai PQ, Duong TN, Chuang JH, Liu YL, Ng LC, Shi Y, Tayag EA, Roque VGJ, Suy LLL, Jarman RG, Gibbons RV, Velasco JMS, Yoon IK, Burke DS, Cummings DAT. Region-wide synchrony and traveling waves of dengue across eight countries in Southeast Asia. *Proc Natl Acad Sci*. 2015;112(42):13069–74.
26. London WP, Yorke JA. Recurrent outbreaks of measles, chickenpox and mumps. Seasonal variation in contact rates. *Am J Epidemiol*. 1973;98(6):453–68.
27. Dowell SF. Seasonal Variation in Host Susceptibility and Cycles of Certain Infectious Diseases. *Emerg Infect Dis*. 2001;7(3):369–74.
28. Keeling MJ, Rohani P, Grenfell BT. Seasonally forced disease dynamics explored as switching between attractors. *Physica D: Nonlinear Phenom*. 2001;148(3–4):317–35. doi:10.1016/S0167-2789(00)00187-1.
29. Kleinen T, Held H, Petschel-Held G. The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. *Ocean Dyn*. 2003;53(2):53–63. doi:10.1007/s10236-002-0023-6.
30. Tan JPL, Cheong SA, et al. Critical slowing down associated with regime shifts in the US housing market. *Eur Phys J B-Condens Matter Complex Syst*. 2014;87(2):1–10.
31. Tan J, Cheong SA. The regime shift associated with the 2004–2008 US housing market bubble. *PLoS ONE*. 2016;11(9):0162140.
32. Gillespie DT. Exact Stochastic Simulation of Coupled Chemical Reactions. *J Phys Chem*. 1977;81(25):2340–61.
33. Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P. Seasonality and the dynamics of infectious diseases. *Ecol Lett*. 2006;9(4):467–84.
34. O'Dea E. Spaero: Software for Project AERO. 2016. R package version 0.2.0. <https://CRAN.R-project.org/package=spaero>.
35. King AA, Ionides EL, Bretó CM, Ellner SP, Ferrari MJ, Kendall BE, Lavine M, Nguyen D, Reuman DC, Wearing H, Wood SN. pomp: Statistical Inference for Partially Observed Markov Processes. 2016. R package, version 1.7. <https://CRAN.R-project.org/package=pomp>.
36. Cleveland RB, Cleveland WS, Terpenning I. Stl: A seasonal-trend decomposition procedure based on loess. *J Off Stat*. 1990;6(1):3.
37. Torrence C, Compo GP. A practical guide to wavelet analysis. *Bull Am Meteorol Soc*. 1998;79(1):61–78. doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

38. Liu Y, San Liang X, Weisberg RH. Rectification of the bias in the wavelet power spectrum. *J Atmos Ocean Technol*. 2007;24(12):2093–102. doi:10.1175/2007JTECHO511.1.
39. Gouhier TC, Grinsted A, Simko V. Biwavelet: Conduct Univariate and Bivariate Wavelet Analyses. 2016. (Version 0.20.11). <http://github.com/tgouhier/biwavelet>.
40. Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, Held H, van Nes EH, Rietkerk M, Sugihara G. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53–9.
41. Dakos V, Carpenter SR, Brock WA, Ellison AM, Guttal V, Ives AR, Kéfi S, Livina V, Seekell DA, van Nes EH, Scheffer M. Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. *PLoS ONE*. 2012;7(7):e41010.
42. Fawcett T. An introduction to ROC analysis. 2006. doi:10.1016/j.patrec.2005.10.010.
43. Williamson MS, Bathiany S, Lenton TM. Early warning signals of tipping points in periodically forced systems. *Earth Syst Dynam*. 2016;7:313–26. doi:10.5194/esd-7-313-2016.
44. Clements CF, Ozgul A. Rate of forcing and the forecastability of critical transitions. *Ecol Evol*. 2016. doi:10.1002/ece3.2531.
45. O'Regan SM, Lillie JW, Drake JM. Leading indicators of mosquito-borne disease elimination. *Theor Ecol*. 2015;9(3):269–86.
46. Han BA, Drake JM. Future directions in analytics for infectious disease intelligence. *EMBO Rep*. 2016;17(6):785–9. doi:10.15252/embr.201642534.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

