Ecology, 96(6), 2015, pp. 1724–1732

Detecting cyclicity in ecological time series

Stilianos Louca[1,4] and Michael Doebeli[2,3]

[1]*Institute of Applied Mathematics, University of British Columbia, 121-1984 Mathematics Road, Vancouver, British Columbia V6T 1Z2 Canada*
[2]*Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia V6T 1Z4 Canada*
[3]*Department of Mathematics, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia V6T 1Z4 Canada*

Abstract. Cyclic population dynamics are of central interest in ecology. Reliably identifying and quantifying the cyclicity of populations is valuable for the understanding of regulatory mechanisms and their variability across spatiotemporal scales. Cyclicity can be detected using periodogram analysis of time series. The statistical significance of periodogram peaks is commonly evaluated against the null hypothesis of uncorrelated fluctuations, also known as white noise. Here, we show that this null hypothesis is inadequate for cycle detection in ecosystems with non-negligible correlation times. As an alternative null hypothesis we propose the so-called Ornstein-Uhlenbeck state-space (OUSS) model, which generalizes white noise to allow for temporal correlations. We justify its use on mechanistic principles and demonstrate its advantages using numerical simulations of simple population models. We show that merely contrasting cyclicity against white noise greatly increases the false cycle detection rate and can lead to wrong conclusions even for simple systems. A comparative statistical analysis of the Global Population Dynamics Database using both null hypotheses suggests that a significant number of populations might have been misinterpreted as cyclic in the past. Our proposed methods for cycle detection are available as an R package (peacots).

Key words: fluctuations; null hypothesis; Ornstein-Uhlenbeck process; population cycles; state-space model; statistical significance; temporal correlation; white noise.

Introduction

The detection of cyclicity has been a long-standing problem in disciplines ranging from astronomy (Brault and White 1971, Scargle 1982) to geophysics (Båth 2012) and ecology (Bulmer 1974, Kendall et al. 1998). Reliably quantifying the cyclicity of populations is a valuable tool for the identification and understanding of regulatory mechanisms and stressors, as well as their variability across spatiotemporal scales. These are particularly important issues in conservation biology (Murray et al. 2008, Salvidio 2009) and climate change research (Ims et al. 2008, Kausrud et al. 2008).

The cyclicity of a stochastic process is commonly identified with the existence of a clear global maximum in its power spectrum, a so-called spectral peak, at a nonzero frequency. Spectral peaks can be detected in the periodogram of a time series, which is a stochastic estimator for the true power spectrum of the unknown generating process. Periodograms are typically calculated using the discrete Fourier transform (Platt and Denman 1975). The statistical significance, or false alarm probability (FAP), of a periodogram peak is the probability of the peak appearing by mere chance, rather than because of true cyclicity, given some noncyclic null hypothesis for the underlying process. More precisely, it is the probability that the maximum of a periodogram produced under the null hypothesis is at least as strong as the observed peak. The null hypothesis is typically chosen to be uncorrelated (white) noise (Scargle 1982, Horne and Baliunas 1986, Kendall et al. 1998, Glynn et al. 2006, Murray et al. 2008), which has a flat power spectrum (Fig. 1a). With this null hypothesis, the FAP is given by

$$P_{\mathrm{WN}} = 1 - \left[1 - e^{-s_{\mathrm{p}}/s}\right]^N \tag{1}$$

where s is the power of the white noise process (WN) at any frequency, s_{p} is the power of the periodogram peak, and N is the number of considered frequencies in the periodogram (Scargle 1982, Chatfield 1996). Periodogram peaks are accepted as indicators of cyclicity if P_{WN} falls below a fixed significance level, typically 0.05. The white noise power s can be estimated from the sample variance (Horne and Baliunas 1986), which is equivalent to using the average periodogram power.

However, taking uncorrelated noise as a null hypothesis for ecological fluctuations is only sensible if their temporal correlation is much shorter than the time intervals between observations, an idealized case sometimes referred to as perfect compensation (Royama 1977, Berryman and Turchin 2001). Ecosystems can easily exhibit comparably long correlation times, even if

Manuscript received 24 January 2014; revised 3 November 2014; accepted 26 November 2014. Corresponding Editor: A. M. Kilpatrick.

[4] Email: louca@math.ubc.ca

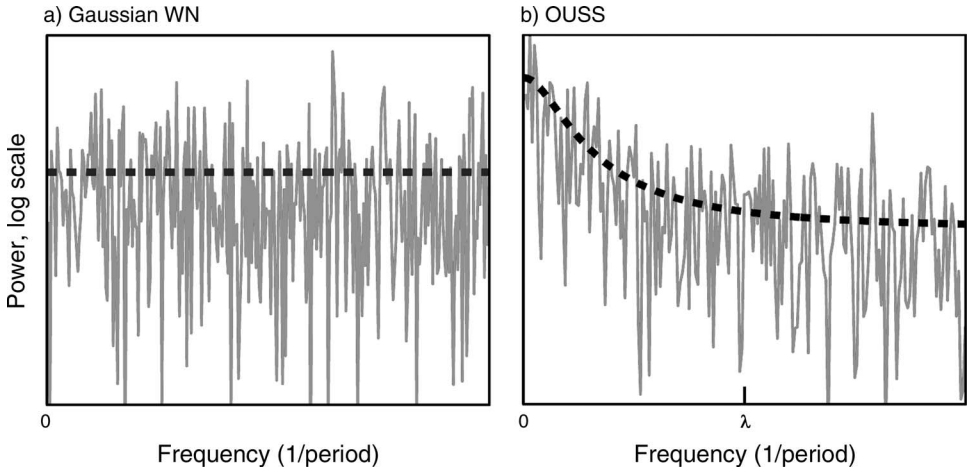

FIG. 1. Periodograms (power over oscillation frequency; continuous curves) of time series generated by (a) Gaussian white noise and (b) an Ornstein-Uhlenbeck state-space (OUSS) process. The resilience of the OUSS process (λ, indicated on the frequency axis), determines the typical frequencies below which the power spectrum increases (Eq. 4). OUSS measurement errors were comparable to natural fluctuations ($\varepsilon = \sigma$), resulting in a long-tailed periodogram. The dashed curves represent the theoretical expectation.

they merely undergo random fluctuations about a stable equilibrium. These correlations can, for example, result from slow reactions of the abiotic environment, slow return times to equilibria after disturbance, or large generation times of the focal species (Steele 1985, Ricker 1997, Vasseur and Yodzis 2004, Myers and Cory 2013). In the periodogram, these long-term correlations are typically reflected in an increased power at lower frequencies, as has been observed in several ecological and environmental variables (Gutierrez and Almirall 1989, Sugihara 1995, Halley 1996). In these cases, periodogram peaks are more likely to appear at lower frequencies and to dominate the rest of the periodogram to an extent very unlikely to occur with uncorrelated noise. Hence, ruling out white noise as the source of a periodogram peak cannot be considered a strong argument for true cyclicity.

An Alternative Null Hypothesis

As an alternative null hypothesis of ecological dynamics, we propose the so-called one-dimensional Ornstein-Uhlenbeck (OU) process (Uhlenbeck and Ornstein 1930, Gardiner 1985). The OU process is defined as a Brownian motion (or random walk) in a quadratic potential well and offers a simple description of a stochastic system subject to linear stabilizing forces. It typically appears as a weak-noise approximation in stochastic population models exhibiting a stable equilibrium in the deterministic limit (Pineda-Krch et al. 2007, Baxendale and Greenwood 2011). Formally, the OU process is defined by a stochastic differential equation

$$\frac{dX}{dt} = \lambda(\mu - X) + \sqrt{2\lambda\sigma^2}\, W \qquad (2)$$

where μ is the deterministic equilibrium and λ is the deterministic resilience of the system. The term W denotes Gaussian white noise, resulting in fluctuations of X about its equilibrium with a standard deviation σ.

The temporal correlation of the OU process decays exponentially at a rate equal to the resilience of the system. Its power at frequency f is given by

$$S(f) = \frac{S_o\lambda^2}{(2\pi f)^2 + \lambda^2} \qquad (3)$$

(where $S_o = 2\sigma^2/\lambda$ is the power at zero frequency), which decays quadratically as $f \to \infty$ (Fig. 1b). When the resilience of the system is high compared to the considered frequencies ($\lambda \gg f$) or, equivalently, when the system's correlation time is much shorter than the considered time scales ($1/\lambda \ll 1/f$), Eq. 3 becomes identical to the white noise spectrum. This is because the OU process becomes Gaussian white noise in the limit of instantaneously decaying fluctuations ($\lambda \to \infty$). We argue that the OU process is a more suitable null hypothesis for noncyclic ecological dynamics because it can account for the temporal correlations between fluctuations about an equilibrium. Analogous arguments have been made for the use of higher-order autoregressive null models for the detection of density dependence in ecological time series (Berryman and Turchin 2001).

In practice, the power spectrum of an ecological variable is estimated from the periodogram of a finite time series. Furthermore, measurements inevitably introduce errors that can be hard to quantify and that compromise the suitability of models such as the OU process for describing the data. In order to accommodate the possibility of random measurement errors, we extend the OU process by uncorrelated errors that are added to the OU process $X(t)$ at each sample point.

Hence, under the null hypothesis, a time series taken at times $t_1, \ldots, t_M$ is assumed to be given by the values $X(t_1) + Y_1, \ldots, X(t_M) + Y_M$, where $Y_1, \ldots, Y_M$ are uncorrelated random variables with zero mean and variance ε^2. Note that the exact distribution of the Y_k is irrelevant for the purposes of periodogram analysis. This model is called the OU state-space (OUSS) model and has recently been proposed by Dennis and Ponciano (2014) as a model for population time series.

For a long regular time series, the obtained periodogram is asymptotically exponentially distributed, independently at each frequency (Brillinger 2001). If δ is the time step, the expected periodogram value at frequency $f = mf_o$ (where $f_o = 1/(M\delta)$ and $m = 0, \ldots, M/2$) is asymptotically given by the function

$$s(f) = \frac{s_o(1-\rho)^2}{1+\rho^2-2\rho\cos(2\pi f\delta)} + \delta\varepsilon^2 \tag{4}$$

that we refer to as the OUSS power spectrum. Here, $\rho = e^{-\lambda\delta}$ is the Pearson correlation between two consecutive measurements and

$$s_o = S_o\delta\frac{\lambda(1+\rho)}{2(1-\rho)} \tag{5}$$

is the power at zero frequency originating in the OU process. The last term in Eq. 4 is the additional power in the time series due to measurement errors. We refer to Appendix A for more details. If the time step δ is very small compared to the considered period $1/f$ and the correlation time $1/\lambda$, then $s(f)$ becomes equal to $S(f)$. On the other hand, if δ is much greater than $1/\lambda$ (i.e., consecutive measurements are uncorrelated), then $s(f)$ becomes flat and equal to a white-noise-like spectrum $\delta(\sigma^2 + \varepsilon^2)$.

We define the statistical significance of the periodogram peak (power s_p at frequency f_p), denoted by P_{OUSS}, as the probability that the periodogram maximum (power $\hat{s}_p$ at frequency $\hat{f}_p$) of any OUSS time series with the same parameters turns out to satisfy

$$\frac{\hat{s}_p^2}{\bar{s}(\hat{f}_p)} \geq \frac{s_p^2}{\bar{s}(f_p)}. \tag{6}$$

Here, $\bar{s}$ is the expected periodogram power and is equal to Eq. 4 only in the infinite time-series limit (the correction for finite time series is given in Appendix A). In the simpler case of white noise, condition 6 reduces to the standard condition $\hat{s}_p \geq s_p$. In the general OUSS case, the modified condition accounts for the nonconstant power spectrum, which results in a lower likelihood of spurious peaks appearing at high frequencies. Consequently, periodogram peaks observed at high frequencies will be statistically more significant.

The parameters s_o, λ, and ε can be estimated from the periodogram through maximum-likelihood fitting. The statistical significance of the periodogram peak can then be calculated using numerical Bernoulli experiments, in which random periodograms of the OUSS model are emulated using exponentially distributed numbers with mean equal to $\bar{s}$. Details are given in Appendix B. The performance of this approach is discussed in Appendix C.

Theoretical Comparison

To explore the differences between the two cyclicity tests, i.e., using the OUSS or WN as a null hypothesis, we used numerical simulations of an OUSS model with a cyclic component of varying amplitude. The OUSS model provides a generic description of fluctuating ecological variables stabilized by linear forces and subject to measurement error. More precisely, we considered the OU process in Eq. 2 and chose a periodically oscillating equilibrium μ of the form

$$\mu(t) = A\sin(2\pi t/T) \tag{7}$$

where A and T are the oscillation amplitude and period, respectively. Such an oscillating equilibrium can originate, for example, in seasonal forcing. In the deterministic limit, X oscillates with an amplitude α, which is generally smaller than A due to the finite resilience λ. In the presence of noise, X fluctuates around its deterministic trajectory with an approximate variance σ^2 and correlation time $1/\lambda$. Time series obtained from this process were superimposed by Gaussian error terms of variance ε^2, yielding an OUSS process superimposed on a cyclic signal. We considered the standard deviation σ as a reference scale of natural fluctuations to which we related the other two parameters α and ε. For example, if $\alpha \ll \sigma$, then X is practically noncyclic, and if $\varepsilon \gg \alpha$, then measurement error completely overshadows the cyclic signal.

We generated a large number of time series of the cyclic OUSS model, while choosing the parameters within a wide plausible range. We examined cycle detection rates using both null hypotheses at a nominal significance threshold of 0.05, for varying α/σ and ε/σ. We considered time series of two different qualities (200 points across 25 correlation times or 40 points across 15 correlation times). Technical details are given in Appendix D.

The results, summarized in Fig. 2, allow for some key observations: For low-quality time series, both methods have similar cycle detection rates in the presence of cycles ($\alpha \gtrsim \sigma$), reaching over 80% when cycles dominate stochastic fluctuations ($\alpha \gtrsim 2\sigma$), even in the presence of strong measurement error (Fig. 2a, c, e). In the absence of cycles ($\alpha \ll \sigma$), the OUSS test results in far fewer Type I errors (~5%) than the WN test (~23%). For high-quality time series, the differences between the two tests becomes much more apparent. While the OUSS test becomes better at distinguishing between the presence or absence of cycles (~5% detection rate for $\alpha = 0$, 98% for $\alpha = 2\sigma$; Fig. 2d, f), the WN test has a cycle detection rate of over 79%, even in the complete absence of cycles (Fig. 2b, f). The decreased specificity

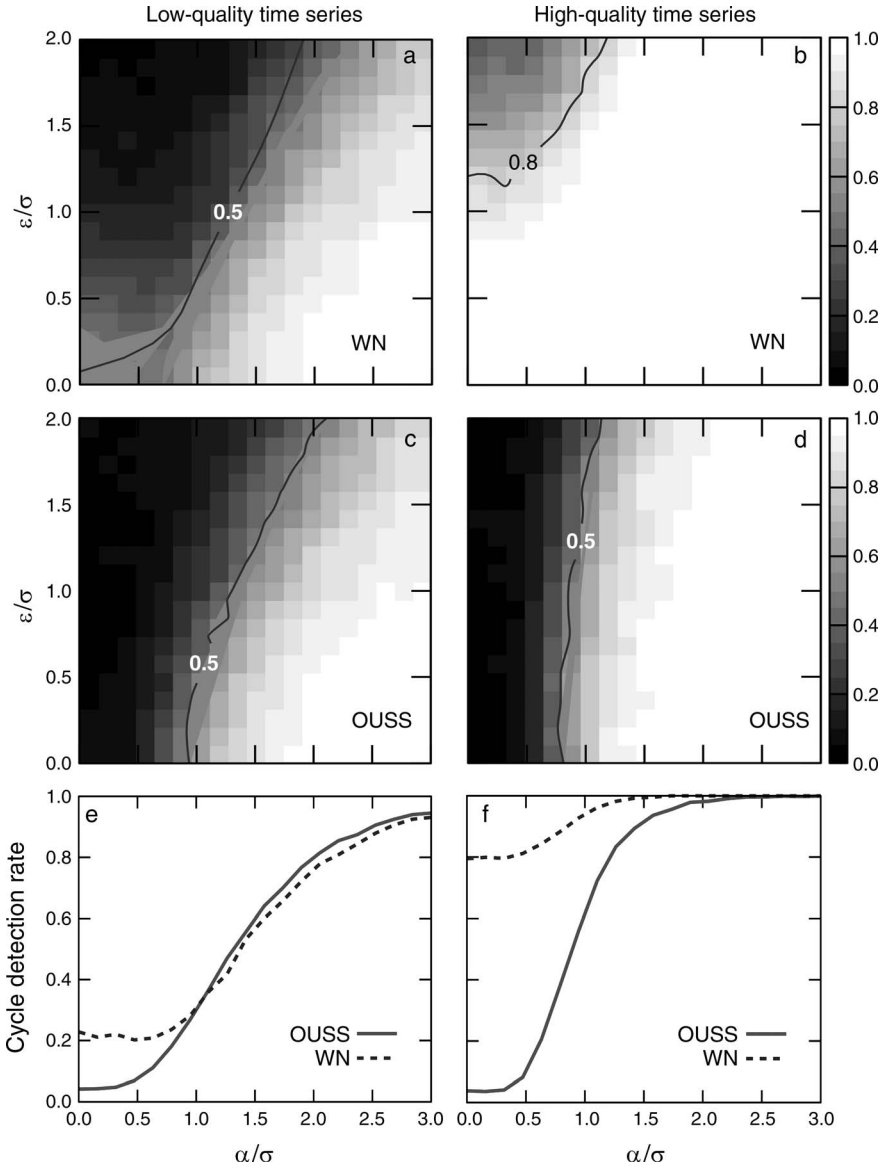

FIG. 2. (a–d) Cycle detection rates using (a, b) the white noise (WN) null hypothesis as well as (c, d) the OUSS null hypothesis at a significance threshold of 0.05, for varying cycle amplitudes α and measurement error amplitudes ε. Lighter shades correspond to higher cycle detection rates. (e, f) Cycle detection rates for varying α and averaged over all ε. In each plot, the far-left region ($\alpha \ll \sigma$) corresponds to the noncyclic limit. Time series were generated using an OUSS model with periodic equilibrium. Technical details are given in Appendix D.

of the WN test with superior time series highlights the pitfalls associated with such a test and reveals the fundamental nature of its flaws. A further, unsurprising observation is that measurement errors tend to reduce the cycle detection rate, albeit not significantly and less so for higher-quality time series (Fig. 2b, d).

An Example on Population Models

To further illustrate our argument, we compared the OUSS and WN tests for time series generated by two simple population models. Details are given in Appendix E. The first model corresponds to a noncyclic density-regulated population, exhibiting a stable equilibrium disturbed by environmental noise. More precisely, we considered the stochastic logistic growth model

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) + \frac{N}{K}\sqrt{2r\sigma^2}\,W \tag{8}$$

with carrying capacity K and intrinsic growth rate r. The rightmost term in Eq. 8 is Gaussian white noise, corresponding to environmental and demographic stochasticity. The noise amplitude scales linearly with population size. For weak noise, the population density N fluctuates about its equilibrium value K approximate-

ly as an OU process with variance σ^2 and correlation time $1/r$. To each time-series point, we added a random Gaussian error term of variance ε^2, thus obtaining a logistic growth state space (LGSS) model.

Simulations and periodogram analysis of the model demonstrated that generated time series can easily be falsely identified as cyclic by the WN test, verifying our previous conclusions. Fig. 3a, b exemplifies this by showing a generated time series and its corresponding periodogram. The latter is clearly not generated by white noise, and testing the observed periodogram peak against white noise would falsely identify the time series as cyclic ($P_{WN} \approx 0.014$). In contrast, the periodogram and its peak seem much more likely to be generated by an OUSS process ($P_{OUSS} > 0.3$).

We examined a large number of time series of the noncyclic LGSS model, while randomly choosing σ/K and ε/σ within a wide plausible range. We found that 20.8% of the peaks had a significance $P_{WN} < 0.05$, while only 4.5% had a significance $P_{OUSS} < 0.05$. This demonstrates that even for very simple population dynamics, using WN as a null hypothesis can lead to a false cycle detection rate that is much higher than the chosen significance level. In contrast, a comparison of time series to the OUSS process results in a much more frequent recognition of noncyclic correlated fluctuations.

We then modified the logistic growth model to describe a cyclic population driven by the periodic variation of its carrying capacity. More precisely, we considered the stochastic model

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K + A\sin(2\pi t/T)}\right) + \frac{N}{K}\sqrt{2r\sigma^2}W \qquad (9)$$

where A and T are the oscillation amplitude and oscillation period of the carrying capacity, respectively. In the deterministic limit and for small amplitudes ($A \ll K$), N oscillates about its carrying capacity with a decreased amplitude α. Due to the noise term in Eq. 9, the population cycles are superimposed by random but correlated fluctuations, similarly to the cyclic OUSS process. Finally, we added random error terms to the time series, as described for the previous two state-space models.

Similar to the noncyclic model, we examined the cycle detection rate for a large number of random time series of the cyclic LGSS model. Sensitivities were high and comparable for both tests: 66.4% and 70.8% of the cases were detected as cyclic using the WN and OUSS test, respectively (64.8% of the cycles were detected by both tests). This shows that the greater generality of the OUSS process does not necessarily result in a reduction of sensitivity. In fact, the detection rate of the OUSS test increases for high-frequency cycles, because spurious periodogram peaks have a decreased likelihood of appearing at higher frequencies in an OUSS process (Fig. 4).

Cycles in Natural Populations

Kendall et al. (1998) tested nearly 700 time series in the Global Population Dynamics Database (GPDD) for cyclicity against the null hypothesis of white noise, concluding that approximately 29% of the examined time series were cyclic (database *accessible online*).[5] In analogy to their work, we tested more than 1700 time series from the GPDD for cyclicity against both white noise as well as the more general OUSS process. For consistency with Kendall et al. (1998), time series were detrended for the WN test to suppress low-frequency modes. For the OUSS test, time series were not detrended because (1) the OUSS model is constructed to describe the increased power at low frequencies and (2) detrending time series generated by an OUSS process will not yield an OUSS process anymore (but see *Discussion* for a proposed alternative to detrending). Methodological details are given in Appendix F. Our source code is provided in Supplement 1.

We found that 28.9% of the observed periodogram peaks had a significance $P_{WN} < 0.05$, in line with Kendall et al.'s (1998) findings. On the other hand, the OUSS test only classified 17.6% of cases as cyclic. This suggests that the number of cyclic populations might have been greatly overestimated by previous studies (by up to 64%), with many of the populations that were interpreted as cyclic merely undergoing correlated fluctuations about their equilibria. Fig. 5a shows the joint distribution of P_{WN} and P_{OUSS} across all examined time series. As can be seen, the two methods agree well for periodogram peaks away from the low-frequency region, as revealed by the nearly diagonal distribution of P_{WN} and P_{OUSS} for these cases. In contrast, the two tests differ significantly for time series whose periodograms peak at low frequencies. An example is given by Fig. 6 for a raccoon population time series, whose periodogram peak is much better explained by the OUSS model than by white noise ($P_{OUSS} > 0.5$, $P_{WN} < 0.05$).

Discussion

The described estimator for the false alarm probability (FAP) is only accurate in the ideal case of regular and infinitely long time series, with a priori known parameters, s_o, λ, and ε of the hypothesized OUSS process. For realistic data sets, none of these assumptions are met, and hence the proposed significance test is only an approximate one. In fact, for low-quality time series, we observed a significant overestimation of the FAP at values below 0.1. We corrected this bias by transforming the estimated FAPs with a conversion table generated using Monte Carlo simulations. Details are given in Appendix C. We note that the statistics of the maximum-likelihood estimator for s_o, λ, and ε, as well as the FAP estimator, are not completely understood. These issues also exist for the simpler null hypothesis of

[5] http://www3.imperial.ac.uk/cpb/databases/gpdd

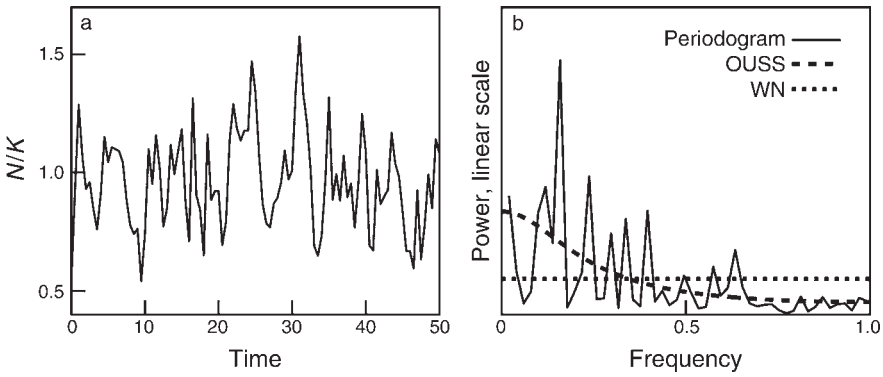

FIG. 3. (a) Sample time series and (b) corresponding periodogram of the noncyclic population model in Eq. 8, where N represents population density and K represents carrying capacity. The dashed curves in (b) show the estimated WN and OUSS power spectra. The observed peak in (b), $f_p \approx 0.16$, has a statistical significance $P_{WN} \approx 0.014$ against the null hypothesis of WN, and $P_{OUSS} \approx 0.32$ against the null hypothesis of an OUSS process. The time series spans 100 points, with $\sigma/K = 0.2$ and $\varepsilon/\sigma = 0.5$.

white noise (Schwarzenberg-Czerny 1996, Frescura et al. 2008, Süveges 2012).

An evaluation of the cyclicity of time series is often complicated by the presence of a dominating low-frequency maximum (LFM) that masks lower, but potentially interesting peaks. LFMs can arise through long correlation times (as emphasized in this study), but can also result from long-term trends, for example due to environmental changes such as climate change (Kausrud et al. 2008). By construction, both the WN and OUSS tests only evaluate the significance of the global periodogram maximum. For the WN test, time series are typically detrended in order to extract the stationary part of the underlying process (Kendall et al. 1998, Clarke et al. 2009). However, detrending can introduce spurious low-frequency periodicities (Hatanaka and Howrey 1969, Nelson and Kang 1981), and the interpretation of a rejection of the OUSS null hypothesis based on a detrended time series is unclear because detrending does not preserve the OUSS structure.

To avoid an overestimation of the OUSS power due to suspected long-term trends and to evaluate the significance of periodogram peaks masked by an LFM (if the latter is suspected of resulting from long-term trends), we suggest ignoring periodogram frequencies within the LFM. Such "low-frequency trimming" is compatible with an OUSS test, because both the OUSS parameters as well as the FAP can be estimated from any particular subset of available frequencies. Care should be taken, however, with the interpretation of such a test. Rejecting the OUSS null hypothesis based on a low-frequency trimmed periodogram only means that it is unlikely that an OUSS process (with the estimated parameters) could have generated such an extreme peak within the trimmed frequency band. Fig. 7 shows an example time series of the Canadian lynx, where low-frequency trimming is sensible. The characteristic 10-year cycle (Elton and Nicholson 1942) is masked by an LFM, which can be avoided by ignoring all frequencies below 1/(20 years). An indiscriminate re-evaluation of the GPDD using the same low-frequency trimming for the OUSS test identified 21.9% of time series as cyclic, an increase of about 24% over the simple OUSS test (Fig. 5b).

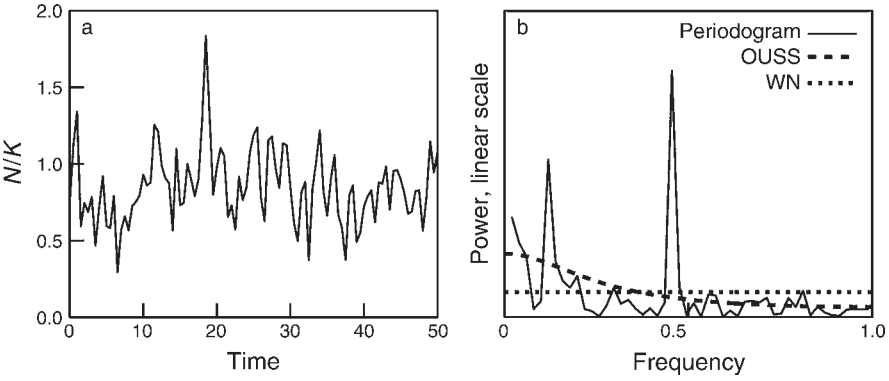

FIG. 4. (a) Sample time series and (b) corresponding periodogram of the cyclic population model (Eq. 9). The dashed curves in (b) show the estimated WN and OUSS power spectra (linear scale). The periodogram peak in (b) has statistical significances $P_{WN} \approx 0.003$ and $P_{OUSS} \approx 0.0004$ against the null hypothesis of an OUSS process. The time series spans 100 points, with $\sigma/K = 0.2$, and $\alpha/\sigma = 1$ and $\varepsilon/\sigma = 0.5$.

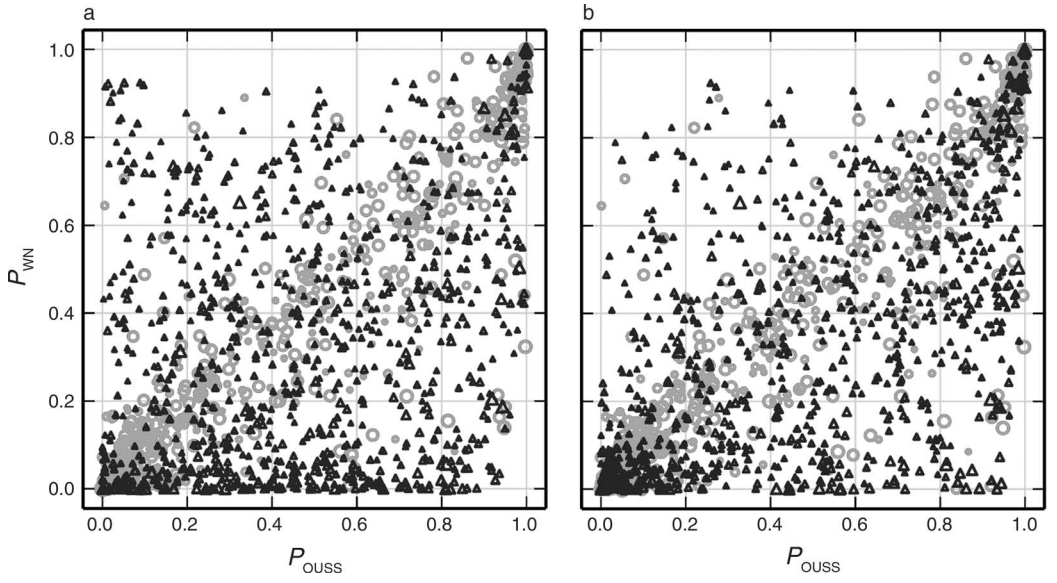

FIG. 5. Scatterplot of the statistical significances of periodogram peaks (P) in time series analyzed from the GPDD, using the OUSS test (x-axis) and the WN test (y-axis). (a) The OUSS test was applied to the original time series. Black triangles denote time series for which the periodogram peak is located within the lowest three available frequencies; gray circles denote the remaining time series. Point sizes are proportional to the length of the time series. (b) The OUSS test was applied to the low-frequency trimmed time series; threshold 1/(20 yr). Technical details are given in Appendix F.

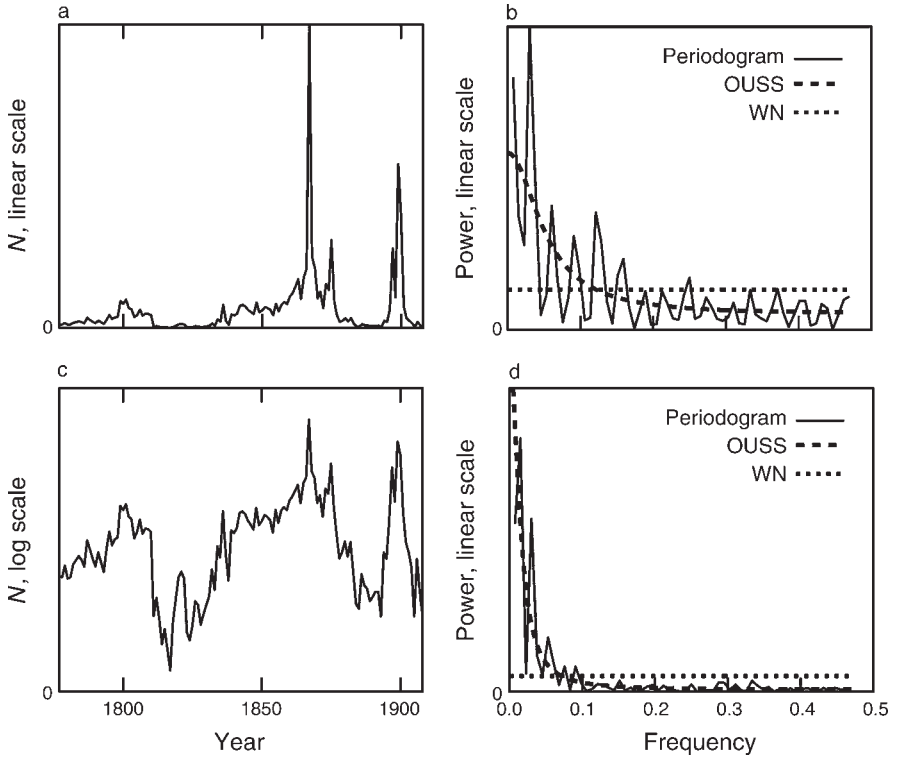

FIG. 6. (a) Time series and (b) corresponding periodogram for the North American raccoon (*Procyon lotor*) by Hudson Bay Company (Poland 1892). The dashed curves in (b) show the fitted WN and OUSS power spectra. The peak in (b), $f_p \approx 1/(33 \text{ yr})$, has a statistical significance $P_{WN} \approx 0.034$ against the null hypothesis of white noise, and $P_{OUSS} \approx 0.59$ against the null hypothesis of an OUSS process. (c) Log-transformed time series and (d) corresponding periodogram, respectively, taken from the Global Population Dynamics Database (GPDD), entry 9711.

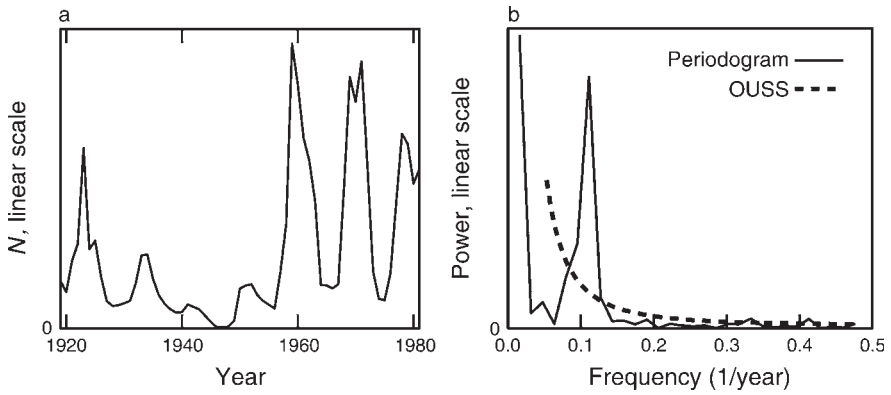

FIG. 7. (a) Time series and (b) corresponding periodogram for the Canadian lynx (*Lynx canadensis*) by Ontario Trappers Association (Novak et al. 1987). The dashed curve in (b) shows the maximum-likelihood fitted OUSS power spectrum after low-frequency trimming at threshold 1/(20 yr). The peak in (b), $f_p \approx 1/(9 \text{ yr})$, has a statistical significance $P_{\text{OUSS}} \approx 0.00014$ against the null hypothesis of an OUSS process. Taken from the GPDD, entry 407.

We emphasize that our goal was not a thorough test of the cyclicity of all populations in the GPDD, but rather to demonstrate the importance of a suitable null hypothesis using realistic data. Expert knowledge might be required for appropriate data preprocessing and interpretation in individual cases. For example, frequency trimming should only be used on a case-by-case basis, since it is always associated with a loss of information on the structure of the underlying power spectrum. Furthermore, some population sizes might be best described by an OUSS process on a logarithmic scale, rather than a linear scale (see, e.g., Fig. 7; Dennis and Ponciano 2014). Repeating the GPDD analysis using log-transformed data identified even more populations as OUSS processes (10.0% or 17.3% classified as cyclic without or with frequency trimming, respectively).

On more general lines, time series can contain several putative cyclic components of interest, emerging from multiple oscillatory mechanisms (Bjørnstad and Grenfell 2001, Hammer 2007, Elderd et al. 2013). While secondary peaks might not dominate the observed dynamics, they can reveal important information on less obvious processes. By construction, both the WN and the OUSS test cannot evaluate the statistical significance of secondary periodogram peaks (i.e., local, but not global maxima). We suggest evaluating such peaks based on the ratio of their power to the expected periodogram power of the OUSS process at that frequency, i.e., $s_p/\bar{s}(f_p)$. This roughly corresponds to a comparison of the peak power to the average power of nearby frequencies. The "local P value" of a secondary peak would then be the probability that a periodogram generated by the fitted OUSS process would have a power-to-expectation ratio at least as high as $s_p/\bar{s}(f_p)$, at any frequency. Asymptotically (i.e., for long regular time series), this probability is given by Eq. 1 after replacing s with $\bar{s}(f_p)$. We have implemented all of the proposed statistical methods as an R (R Core Team 2014) package (peacots; periodogram peaks in correlated time series; package *available online*).[6] The package can calculate the periodogram of a time series, maximum-likelihood fit the OUSS null model to the periodogram, and calculate the statistical significance of the periodogram maximum. The package also includes options for low-frequency trimming, as well as calculating the local statistical significance of secondary periodogram peaks. Example code making use of peacots is available in Supplement 3. A thorough usage manual is included in the package distribution.

CONCLUSIONS

Correctly assessing the cyclicity of populations is essential if we want to understand the effects of environmental gradients and climate change on population dynamics (Kendall et al. 1998, Murray et al. 2008, Salvidio 2009). This involves the use of appropriate statistical tools for the evaluation of putative spectral peaks that take into account important features of the focal ecosystem, such as correlated fluctuations (Vasseur and Yodzis 2004). To address the latter, we have demonstrated the superiority of the Ornstein-Uhlenbeck state-space model as a null hypothesis for measured ecological fluctuations over the conventionally used white noise. Our basic statistical analysis of the Global Population Dynamics Database showed that the differences implied for cycle detection can be of practical importance.

While we have focused on cycle detection using periodogram analysis, we expect our results to be generalizable to other similar statistical tests (such as correlogram analysis) that use white noise as a null hypothesis. On these grounds, we advocate a corresponding modification of the statistical methods applied in future studies of ecological cycles, in particular for systems with long correlation times. To facilitate this

[6] http://cran.r-project.org/web/packages/peacots

advancement, we provide the statistical methods as an easy-to-use R package.

Acknowledgments

This work was supported by the PIMS IGTC for Mathematical Biology and NSERC (Canada). We thank two unknown reviewers for their constructive comments.

Literature Cited

Båth, B. 2012. Spectral analysis in geophysics. Elsevier Science, Amsterdam, Netherlands.

Baxendale, P. H., and P. E. Greenwood. 2011. Sustained oscillations for density dependent Markov processes. Journal of Mathematical Biology 63:433–457.

Berryman, A., and P. Turchin. 2001. Identifying the density-dependent structure underlying ecological time series. Oikos 92:265–270.

Bjørnstad, O. N., and B. T. Grenfell. 2001. Noisy clockwork: time series analysis of population fluctuations in animals. Science 293:638–643.

Brault, J. W., and O. R. White. 1971. The analysis and restoration of astronomical data via the fast Fourier transform. Astronomy and Astrophysics 13:169–189.

Brillinger, D. R. 2001. Time series: data analysis and theory. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.

Bulmer, M. G. 1974. A statistical analysis of the 10-year cycle in Canada. Journal of Animal Ecology 43:701–718.

Chatfield, C. 1996. The analysis of time series: an introduction. Chapman and Hall/CRC, Boca Raton, Florida, USA.

Clarke, B., E. Fokoue, and H. Zhang. 2009. Principles and theory for data mining and machine learning. Springer, Berlin, Germany.

Dennis, B., and J. Ponciano. 2014. Density-dependent state-space model for population-abundance data with unequal time intervals. Ecology 95:2069–2076.

Elderd, B. D., B. J. Rehill, K. J. Haynes, and G. Dwyer. 2013. Induced plant defenses, host–pathogen interactions, and forest insect outbreaks. Proceedings of the National Academy of Sciences USA 110:14978–14983.

Elton, C., and M. Nicholson. 1942. The ten-year cycle in numbers of the lynx in Canada. Journal of Animal Ecology 11:215–244.

Frescura, F. A. M., C. A. Engelbrecht, and B. S. Frank. 2008. Significance of periodogram peaks and a pulsation mode analysis of the Beta Cephei star V403 Car. Monthly Notices of the Royal Astronomical Society 388:1693–1707.

Gardiner, C. W. 1985. Handbook of stochastic methods for physics, chemistry and the natural sciences. Springer, Berlin, Germany.

Glynn, E. F., J. Chen, and A. R. Mushegian. 2006. Detecting periodic patterns in unevenly spaced gene expression time series using Lomb–Scargle periodograms. Bioinformatics 22: 310–316.

Gutierrez, E., and H. Almirall. 1989. Temporal properties of some biological systems and their fractal attractors. Bulletin of Mathematical Biology 51:785–800.

Halley, J. M. 1996. Ecology, evolution and 1 f-noise. Trends in Ecology and Evolution 11:33–37.

Hammer, Ø. 2007. Spectral analysis of a Plio-Pleistocene multispecies time series using the Mantel periodogram. Palaeogeography, Palaeoclimatology, Palaeoecology 243: 373–377.

Hatanaka, M., and E. P. Howrey. 1969. Low frequency variation in economic time series. Kyklos 22:752–766.

Horne, J. H., and S. L. Baliunas. 1986. A prescription for period analysis of unevenly sampled time series. Astrophysical Journal 302:757–763.

Ims, R. A., J. A. Henden, and S. T. Killengreen. 2008. Collapsing population cycles. Trends in Ecology & Evolution 23:79–86.

Kausrud, K. L., et al. 2008. Linking climate change to lemming cycles. Nature 456:93–97.

Kendall, B. E., J. Prendergast, and O. N. Bjørnstad. 1998. The macroecology of population dynamics: taxonomic and biogeographic patterns in population cycles. Ecology Letters 1:160–164.

Murray, D. L., T. D. Steury, and J. D. Roth. 2008. Assessment of Canada lynx research and conservation needs in the southern range: another kick at the cat. Journal of Wildlife Management 72:1463–1472.

Myers, J. H., and J. S. Cory. 2013. Population cycles in forest Lepidoptera revisited. Annual Review of Ecology, Evolution, and Systematics 44:565–592.

Nelson, C. R., and H. Kang. 1981. Spurious periodicity in inappropriately detrended time series. Econometrica 49:741–751.

Novak, M., M. Obbard, J. Jones, R. Newman, A. Booth, A. Satterthwaite, and G. Linscombe. 1987. Furbearer harvests in North America, 1600–1984. Ontario Trappers Association, Toronto, Ontario, Canada.

Pineda-Krch, M., J. H. Blok, U. Dieckmann, and M. Doebeli. 2007. A tale of two cycles—distinguishing quasi-cycles and limit cycles in finite predator–prey populations. Oikos 116: 53–64.

Platt, T., and K. L. Denman. 1975. Spectral analysis in ecology. Annual Review of Ecology and Systematics 6:189–210.

Poland, H. 1892. Fur-bearing animals in nature and in commerce. Gurney and Jackson, London, UK.

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Ricker, W. E. 1997. Cycles of abundance among Fraser River sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic Sciences 54:950–968.

Royama, T. 1977. Population persistence and density dependence. Ecological Monographs 47:1–35.

Salvidio, S. 2009. Detecting amphibian population cycles: the importance of appropriate statistical analyses. Biological Conservation 142:455–461.

Scargle, J. D. 1982. Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. Astrophysical Journal 263:835–853.

Schwarzenberg-Czerny, A. 1996. Fast and statistically optimal period search in uneven sampled observations. Astrophysical Journal Letters 460:L107.

Steele, J. H. 1985. A comparison of terrestrial and marine ecological systems. Nature 313:355–358.

Sugihara, G. 1995. From out of the blue. Nature 378:559–560.

Süveges, M. 2012. False Alarm Probability based on bootstrap and extreme-value methods for periodogram peaks. Page 16 *in* ADA7-Seventh Conference on Astronomical Data Analysis. Volume 1. Cargése, Corsica, France.

Uhlenbeck, G. E., and L. S. Ornstein. 1930. On the theory of the Brownian motion. Physical Review 36:823–841.

Vasseur, D. A., and P. Yodzis. 2004. The color of environmental noise. Ecology 85:1146–1152.

Supplemental Material

Ecological Archives

Appendices A–F and Supplements 1–4 are available online: http://dx.doi.org/10.1890/14-0126.1.sm

Stilianos Louca, Michael Doebeli (2015)
Detecting cyclicity in ecological time series
Ecology (96):1724–1732
Appendices A–F

A Periodogram of the OUSS model

Let $z_1, .., z_M$ be a real time series generated by a stationary OUSS model and measured at time intervals δ. Without loss of generality, we shall assume that the process mean is zero. Our starting point is the following definition of the periodogram

$$\widehat{s}(mf_o) = T \left| \frac{1}{M} \sum_{k=1}^{M} z_n e^{-2\pi i m(k-1)/M} \right|^2 , \tag{A.1}$$

where $f_o = 1/(M\delta)$ and $m = 0, .., M/2$. We are interested in the expected value $s = \mathbb{E}\{\widehat{s}\}$, asymptotically for large M. Taking the expectation of Eq. (A.1) yields

$$s(mf_o) = \frac{T}{M^2} \sum_{k,l} \mathbb{E}\{z_k z_l\} e^{2\pi i m(k-l)/M}. \tag{A.2}$$

Recall that $z_k = x_k + y_k$, where x_k are generated by an OU process and y_k are centralized and uncorrelated with variance ε^2. The covariance, $\mathbb{E}\{z_k z_l\}$ between any two values z_k and z_l is given by $\sigma^2 e^{-\lambda\delta|k-l|} + \delta_{kl}\varepsilon^2$, where δ_{kl} denotes the Kronecker delta (Gillespie, 1992). Hence Eq. (A.2) can be written as

$$\begin{aligned} s(mf_o) = \frac{T}{M^2} \sum_{k,l} &\left[\sigma^2 \rho^{|k-l|} + \varepsilon^2 \delta_{kl}\right] \\ &\times e^{2\pi i m(k-l)/M}, \end{aligned} \tag{A.3}$$

where $\rho = e^{-\lambda\delta}$. After a few algebraic manipulations one can write Eq. (A.3) as

$$\begin{aligned} s(f) = &\frac{s_o T(1-\rho)}{M^2\delta(1+\rho)} \left[M \frac{1-|r|^2}{|1-r|^2} \right. \\ &\left. - 2\Re\left\{ \frac{r(1-r^M)}{(1-r)^2} \right\} \right] + \delta\varepsilon^2, \end{aligned} \tag{A.4}$$

where we abbreviated

$$s_o = \frac{\delta\sigma^2(1+\rho)}{(1-\rho)},\ r = \rho e^{2\pi i f\delta},\ f = mf_o. \tag{A.5}$$

Note that s_o is the expected periodogram power at zero frequency in the limit $M \to \infty$, introduced in Eq. (5) of the main article. In that limit, Eq. (A.4) can be simplified and one obtains the asymptotic expression (Box et al., 2013, Eq. (3.2.15))

$$s(f) \sim \frac{\delta\sigma^2(1-\rho^2)}{1+\rho^2-2\rho\cos(2\pi f\delta)} + \delta\varepsilon^2. \tag{A.6}$$

If additionally $\lambda\delta \ll 1$ and $f\delta \ll 1$, and assuming $\varepsilon^2/\sigma^2 \in \mathcal{O}(1)$, one retrieves the power spectrum of the classical OU process,

$$s(f) \sim \frac{2\lambda\sigma^2}{(2\pi f)^2+\lambda^2} = S(f). \tag{A.7}$$

B Periodogram analysis

Periodograms were calculated using the Discrete Fourier Transform for the simulations and the Lomb-Scargle periodogram (Lomb, 1976) for the GPDD. Periodogram powers were calculated at frequencies $f_o, \ldots, f_o(M/2-1)$, where M is the length of the time series and f_o is the reciprocal sampling duration, or fundamental frequency. The white noise power was estimated from the mean periodogram power, which is equivalent to a least squares fit (Horne and Baliunas, 1986; Lancaster and Šalkauskas, 1986). The OUSS parameters s_o, λ and ε^2 were estimated through a maximum-likelihood fit to the periodogram, using the ALGLIB library (Bochkanov, 2013). The likelihood function was approximated by

$$L(p_1,..,p_n) = \prod_{k=1}^{n} \exp\left[-p_k/s(f_k)\right]/s(f_k), \tag{B.1}$$

where $p_1,..,p_n$ are the periodogram powers calculated from the time series for any considered frequencies $f_1,..,f_n$, and $s(f)$ is given by Eq. (A.4) in A. Eq. (B.1) assumes the periodogram powers $p_1,..,p_n$ to be uncorrelated and is thus only exact in the limit of infinitely long time series. If not mentioned otherwise, we considered all available periodogram frequencies.

False alarm probabilities (FAP) were estimated using at least 10^4 random periodograms (Bernoulli trials) of the fitted OUSS model, emulated by exponentially distributed numbers as outlined in the main article. At these sample sizes, the Bernoulli estimator has a standard deviation of at most 0.005 (and below 0.0025 if FAP < 0.05). Local P-values (i.e. of non-global periodogram peaks) were calculated using the formula

$$P = 1 - \left[1 - e^{-s_\mathrm{p}/s(f_\mathrm{p})}\right]^N, \tag{B.2}$$

where N is the number of considered frequencies, s_p is the periodogram power at the considered frequency f_p and s is given by Eq. (A.4) in the Appendix. Note that different normalization conventions for periodograms will lead to rescaled estimates for s_o and ε^2, but this has no effect on the estimated FAP.

C Bias and correction of the FAP estimator

We tested the bias of the OUSS FAP estimator by numerically calculating the type-I error rate, i.e. the rate at which OUSS time series are erroneously rejected and classified as cyclic. We generated a large number of time series of the non-cyclic OUSS model described in the main article, while choosing σ, ρ (where $\rho = \exp(-\lambda\delta)$ is the correlation between subsequent time points) and ε/σ within a wide range. More precisely, σ, ρ and ε/σ were uniformly and randomly picked in the range $[0.01, 0.2]$, $(0, 1)$ and $[0, 2]$, respectively. We considered time series of varying length up to 4000. For each time series length (TSL), we ran $25\ 000$ Monte Carlo simulations (only 5000 for TSL$= 4000$). Details on periodogram calculation and FAP estimation are given in B.

The FAP estimator is itself a random variable, whose cumulative distribution function (CDF) under the OUSS null hypothesis should ideally (i.e., in the absence of bias) be the identity function, $F(t) = t$. For example, at a 0.05 significance threshold, the null hypothesis should be rejected 5% of the times. Fig. 1 shows the obtained cumulative distribution functions for FAPs estimated using different TSLs. As can be seen, a considerable bias exists for low quality time series. For example, at a significance threshold 0.05 the type-I error rate is significantly lower than 5%. The bias disappears for longer time series. This observation comes to no surprise, as our FAP estimator assumes periodogram powers to be independent and exponentially distributed, which is only asymptotically true for long time series.

To account for this bias, the obtained CDF was applied to the estimated FAPs in all subsequent calculations. We remind the reader that for any continuous random variable X with CDF F, the random variable $F(X)$ is distributed within $[0, 1]$ and has CDF $G(t) = t$ (*diagonal*). Hence, by mapping the estimated FAPs to the corresponding CDF value, one obtains new rescaled estimators with a diagonal CDF. These rescaled estimators are, strictly speaking, not estimators for the tail of the original test statistic introduced in the main article. But since the applied CDF is monotonic, the rescaling preserves the order of the original FAPs. Hence, the rescaled (*corrected*) FAP remains suitable for testing against the OUSS process and exhibits a correct type-I error rate.

The CDF of the original FAPs is, strictly speaking, a conditional one, that is, it depends on the underlying OUSS process parameters as well as the sampling quality (TSL and time step). Nondimensionalization reduces the number of free parameters to 3, for example ρ (i.e. the correlation between subsequent time points), $\mathrm{NSR} = s_o/(\delta\varepsilon^2)$ (i.e. the power ratio of the OU process to measurement error, or *noise to signal ratio*, at zero frequency) and TSL. We thus constructed a 4-dimensional rectangular grid spanning across different TSLs (ranging from 10 to 2000), estimated ρ values (ranging from 0 to 1), estimated NSR values (ranging from 0 to 5) and different FAP values (ranging from 0.0005 to 1). Each grid point was assigned the empirical CDF value at the particular FAP and for the particular TSL, after binning $400\ 000$ Monte Carlo samples by their estimated ρ and NSR values. We used this grid to correct the FAPs estimated in all subsequent calculations. CDF values between grid points were approximated using multilinear interpolation (Andrews et al., 2013). Cases outside of the grid's domain were mapped to the closest grid value. This yielded a correct type-I error rate, e.g. a cycle detection rate of about 5% at a nominal significance level of 0.05. The precomputed grid and the appropriate FAP correction are included in the R package that we provide.

D Theoretical comparison

Times series of the OUSS model were generated using correlated draws, i.e. by choosing

$$\begin{aligned} \widetilde{x}_{i+1} &= \rho_i \widetilde{x}_i + \sqrt{1-\rho_i^2} \cdot \sigma \; w_i, \\ z_i &= f(t_i) + \widetilde{x}_i + \varepsilon y_i, \end{aligned} \tag{D.1}$$

where w_i and y_i are independent standard-normal variables, $\rho_i = e^{-\lambda|t_{i+1}-t_i|}$ is the correlation between two consecutive time points and $f(t)$ is the deterministic asymptotic solution to the ordinary differential equation $df/dt = \lambda(\mu(t) - f(t))$. The first value $\widetilde{x}_1$ was drawn from a normal distribution with zero mean and variance σ^2. The obtained time series $z_1, z_2, ..$ follows the correct distribution of the stationary OUSS model.

For the comparison of the OUSS and WN tests we generated 50000 low quality and 50000 high quality time series of the OUSS model (40 points across 15 time units and 200 points across 25 time units, respectively). The standard deviation σ and the resilience λ were normalized to 1. The oscillation period T was randomly and uniformly chosen within the interval $[0.5, 4]$. The ratios α/σ and ε/σ were sampled uniformly on a regular grid of size 20×20 spanning the values $[0, 3]$ and $[0, 2]$, respectively. The amplitude A, defined in Eq. 7 of the main article, is connected to the parameters α and T through

$$\alpha = \frac{A\lambda}{\sqrt{\lambda^2 + (2\pi/T)^2}}. \tag{D.2}$$

Periodograms were analyzed as described in B and OUSS FAPs were corrected as described in C.

E LGSS models

For the simulations of the LGSS models (Eq. 8 and Eq. 9 in the main article) we used an explicit two-step Runge-Kutta scheme, described in detail by Milstein (1995, §3.4, Theorem 3.3) and implemented in C++. Our source code is given as a separate supplement. The carrying capacity K and the intrinsic growth rate r were both normalized to 1, hence characteristic population sizes and time units are both 1. The standard deviation σ and the ratio ε/σ were randomly and uniformly chosen within the intervals $[0.01, 0.2]$ and $[0, 2]$, respectively. For the cyclic model, the ratio α/σ and the oscillation amplitude T were chosen uniformly within $[1, 3]$ and $[0.5, 4]$, respectively. The amplitude A was calculated from the chosen α and T using

$$A = \frac{\alpha}{r}\sqrt{r^2 + (2\pi/T)^2}. \tag{E.1}$$

Cases for which $A \geq K$ were skipped. Simulations ran for 25 time units and time series comprised 50 points. The integration time step was $\delta t = 0.0002$. We generated 10^4 independent time series for the non-cyclic as well as the cyclic model. Initial population sizes were chosen randomly according to the theoretical stationary distribution in the small-noise limit (i.e. as Gaussian variables with variance σ^2 around the deterministic trajectory). Periodograms were analyzed as described in B and OUSS FAPs were corrected as described in C.

F Statistical analysis of the GPDD

Population size entries in the GPDD on a logarithmic scale were transformed back into a linear scale for consistency. Similarly to Kendall et al. (1998), we only considered time series consisting of at least 25 data points (1712 cases). Four time series were omitted from the analysis either because of inconsistent values or numerical fit problems. We used the Lomb-Scargle periodogram (Lomb, 1976) because of gaps in several time series. The zero-frequency mode was omitted from the analysis. For the WN test, time series were detrended using LOESS smoothing of degree 1 and a span of half the data, similarly to Kendall et al. (1998). For the OUSS test, time series were not preprocessed at all, and periodograms were either low-frequency trimmed at threshold $\nu = 1/20\,\mathrm{yr}$ or not pre-processed at all (see the main article for the resulting differences). In the former case, all frequencies below ν were omitted both when determining the periodogram peak as well as when estimating the OUSS parameters. For the OUSS test, periodograms whose maximum was either at the lowest mode (in the case of no frequency trimming) or exceeded in power by the immediate lower mode (in the case of frequency trimming) were not considered cyclic, because such one-sided peaks should not be considered indicators of a true power peak. For the WN test this was irrelevant because detrending typically eliminated such one-sided peaks at the lowest mode. Apart from that, periodograms were analyzed as described in B and OUSS FAPs were corrected as described in C.

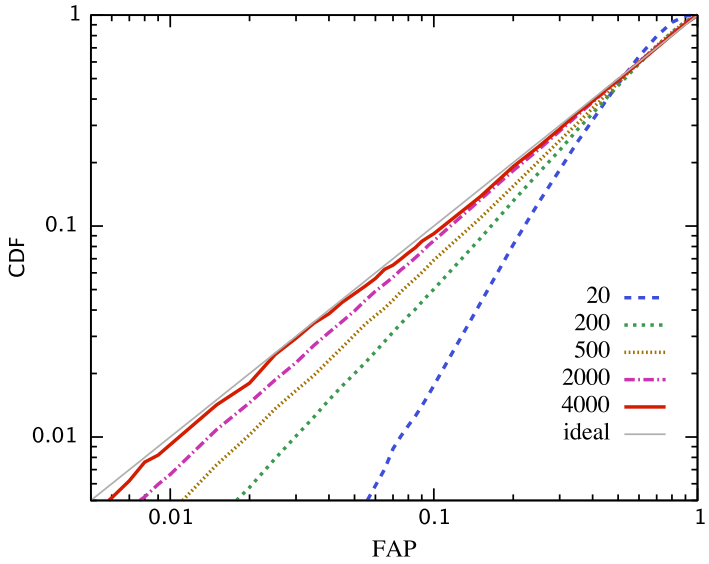

Fig. 1: Cumulative distribution function (CDF) of the uncorrected FAP estimator for randomly generated non-cyclic OUSS time series of different lengths. The grey diagonal corresponds to an ideal (i.e. unbiased) FAP estimator and is shown for comparison. The type-I error rate of an OUSS test at some significance threshold x is given by CDF(x). Technical details are given in C.

References

Andrews, T., R. Balan, J. Benedetto, W. Czaja, and K. Okoudjou. 2013. Excursions in Harmonic Analysis, Volume 2: The February Fourier Talks at the Norbert Wiener Center. Applied and Numerical Harmonic Analysis, Birkhauser.

Bochkanov, S., 2013. ALGLIB 3.8.0. Available at: http://www.alglib.net (Jan 15, 2014).

Box, G., G. Jenkins, and G. Reinsel. 2013. Time Series Analysis: Forecasting and Control. Wiley Series in Probability and Statistics, Wiley.

Gillespie, D. 1992. Markov Processes: An Introduction for Physical Scientists. Academic Press.

Horne, J. H., and S. L. Baliunas. 1986. A prescription for period analysis of unevenly sampled time series. The Astrophysical Journal 302:757–763.

Kendall, B. E., J. Prendergast, and O. N. Bjørnstad. 1998. The macroecology of population dynamics: taxonomic and biogeographic patterns in population cycles. Ecology Letters 1:160–164.

Lancaster, P., and K. Šalkauskas. 1986. Curve and Surface Fitting: An Introduction. Academic Press.

Lomb, N. R. 1976. Least-squares frequency analysis of unequally spaced data. Astrophysics and Space Science 39:447–462.

Milstein, G. 1995. Numerical Integration of Stochastic Differential Equations. Mathematics and its Applications, Kluwer, Dordrecht, The Netherlands.