Detection of imperfect population synchrony in an uncertain world

BERNARD CAZELLES*† and LEWI STONE‡
*Laboratoire d’Ecologie, CNRS UMR 7625, Université Pierre et Marie Curie, 7 quai Saint Bernard, CC 237, 75252 Paris, France, †UFR de Biologie, Université Paris 7 – Denis Diderot, and ‡The Porter Super-Center for Ecological and Environmental Studies, and Department of Zoology Tel Aviv University, Ramat Aviv Tel Aviv 69978, Israel

Summary
1. We propose the method of ‘phase analysis’ for studying the spatio-temporal fluctuations of animal populations, and use this method for helping identify remarkable synchronized population fluctuations that may sometimes be found over very large spatial domains.
2. The method requires decomposing the observed time series of population fluctuations into two components – one that quantifies the changing phase of the signal, and the other that quantifies the changing amplitude.
3. Two populations are considered to be ‘phase synchronized’ if there is locking or synchrony between their phase components, while their associated amplitudes may nevertheless remain largely uncorrelated.
4. Since environmental noise often masks population synchrony, a null hypothesis approach is used to detect whether the phase variables are locked more than would be expected by chance alone.
5. The technique is thus particularly appropriate for ecological analyses where it is often important to study evidence of weak interactions in irregular non-stationary and noisy time series. Because climatic patterns (and predicted climate changes) almost certainly influence population dynamics, the approach appears particularly relevant for analysing the potential links between climatic fluctuations and population abundance.

Key-words: climatic influences, cyclic phase difference, Moran effect, phase synchronization, population cycles, population synchrony.


Introduction
For a century or more ecologists have been greatly interested in studying long-term population cycles, be they periodic, quasi-periodic or chaotic. Some of the more unusual population cycles have periods extending over many years and are thus difficult to explain in terms of simple seasonal patterns. Furthermore, many populations are able to synchronize their oscillations over large spatial domains (Royama 1992), sometimes extending over entire continents. Thus populations that may appear spatially dislocated are able to rise and fall in abundance in an unusually precise synchronized manner. The Canadian 10-year hare–lynx cycle is one of the better known examples of this phenomenon (Elton & Nicholson 1942; Keith 1963; Royama 1992), but it is also known to occur for a great variety of taxa in different locations across the globe (reviewed in Bjørnstad, Ims & Lambin 1999; Koenig 1999). The causes of population cycles may have a straightforward explanation in terms of predator–prey relationships. Usually, however, they are more difficult to understand and can, for example, arise from complicated processes such as combinations of direct and delayed density-dependent and density-independent factors. The ability of different populations to synchronize their oscillations in unison is also difficult to explain and has become a matter of great debate and speculation.

Patterns of spatial synchrony have been attributed to various factors, such as dispersal between spatially separated subpopulations or interactions with nomadic predators (Bjørnstad et al. 1999; Hudson & Cattadori 1999; Koenig 1999; Ims & Andreassen 2000). Dispersal tends to link subpopulations and in turn enhances the possibility of joint synchrony. Because dispersal is often locally restricted and always distance-dependent, such subpopulations are expected to exhibit a strong
Phase analysis: concepts and tools

THE NEED FOR A BROADER DEFINITION OF SYNCHRONY

The traditional measures of synchronization are based on cross-correlation statistics in the time domain and coherence in the frequency domain, both of which depend on the assumption that the signals under investigation are linear and stationary (Box & Jenkins 1976; Chatfield 1989). But population time series are short, irregular, probably non-linear, and often strongly non-stationary. These characteristics make it difficult and frequently inappropriate to use even sliding versions of traditional correlation or spectral techniques when analysing the mutual dependencies in such data.

The nature of the relation between the abundance of two populations, or between a population and an environmental signal that are synchronized, can be much more complex than a simple linear relation. For instance, Myers (1998) emphasized that in some cases, outbreaks of forest Lepidoptera remain in synchrony even when all populations do not reach a high density simultaneously together in each cycle. Conversely, sometimes visual indications of synchrony may be misleading. Milner et al. (1999), studying the influence of climatic fluctuations on the survival of Soay sheep, showed that despite a visual impression of a strong association between population abundance and a climatic index, the correlation between the two variables was in fact not significant. Similarly, Higgins et al. (1997b), who used a stochastic non-linear model to describe Dungeness crab populations, questioned the value of using a linear correlation alone when analysing population dynamics.

Several new concepts of synchronization have appeared in the physics literature introduced especially for the study of non-linear and chaotic systems (Brown...
The concept of phase synchronization (Rosenblum, Pikovsky & Kurths 1996) has recently been employed to explain the observed synchronization of complex oscillations in both observed and model ecological communities (Blasius, Huppert & Stone 1999; Blasius & Stone 2000a; Cazelles & Boudjema 2001), and also has relevance in many other biological disciplines (see also: Schäfer et al. 1998a,b; Tass et al. 1998; Neiman et al. 1999; Rodriguez et al. 1999; Lachaux et al. 2000; Mormann et al. 2000; Pavlov et al. 2000; Holstein-Rathlou et al. 2001; Varela et al. 2001; Bahar et al. 2002). The underlying idea is to check whether two signals tend to oscillate simultaneously at the same ‘pace’, rising and falling together with the same rhythm. For example, if both oscillating signals have maxima that peak at exactly the same times, the signals must be keeping in rhythm, and should therefore be considered phase synchronized. Note that even though the signals might peak together, the maxima are not necessarily of the same magnitude. Hence the fact that two variables are phase synchronized does not mean that they are ‘fully synchronized’, something we discuss more below.

Oscillating signals can be decomposed into two components: phase and amplitude. For example, consider the signals:

\[ x(t) = A \sin (\Omega t), \quad y(t) = A \cos (\Omega t), \]

where \( y(t) \) is proportional to the derivative \( dx/dt \). In the \( x-y \) phase plane, these variables describe motion on the circle \( x^2 + y^2 = A^2 \), of radius or ‘amplitude’ \( A \). The rate at which the point \((x, y)\) rotates about the circle is given by \( \Omega \), and the phase at any time \( t \) is \( \phi(t) = \Omega t \). The phase, of course, describes the angle the point \((x, y)\) makes with respect to the \( x \)-axis.

Two signals \( x_1(t) \) and \( x_2(t) \) are said to be phase synchronized if their phases lock together in some fixed stable relationship and therefore do not tend to drift apart. Again, the amplitudes of the two signals, which could be quite chaotic, do not necessarily have to be synchronized or even correlated. Thus to check for phase synchronization, tools are needed to obtain and keep track of the phases of the signals under investigation. One general approach has been based on the analytic

\[ \Delta\phi \]

signal concept, where the phase of the signal is obtained directly through the use of a mathematical technique known as the Hilbert transform (Pikovsky et al. 1997; Rosenblum et al. 2001). A simpler approach advocated here is based on the maxima (or minima) of the cycles in the time series (Blasius et al. 1999), where one can reasonably assume that the phase difference between any two maxima (minima) is \( 2\pi \).

In many cases the signal (the time series) can be broken up into ‘quasi-cycles’ where the maxima and/or minima of each quasi-cycle can be readily determined (Fig. 1a). Suppose the time series \( x(t) \) has maxima at times \( t_1, t_2, \ldots, t_n, \ldots \). Between any time interval \([t_n, t_{n+1}]\) the phase must increase by exactly \( 2\pi \) (Fig. 1a). Hence, we can assign to the times \( t_n \) the values of the phase \( \phi(t_n) = 2\pi(n-1) \). As a first approximation we can assume that the phase grows linearly with time in the interval \([t_n, t_{n+1}]\) (Fig. 1b) so that:

\[ \Delta\phi = \Delta\Psi \]

Fig. 1. Definition of the phase and phase difference of two time series (population abundances or population abundance and environmental signal). (a) The two time series (thin and thick solid lines); intervals between two maxima (+) where the phase \( \phi \) increases by exactly \( 2\pi \). (b) Evolution of the phase \( \phi \) (solid lines) and of the cyclic phase \( \Psi \) (dashed lines) between the two time series; the \( \Psi \) evolve only between 0 and \( 2\pi \). (c) Evolution, for the two time series, of the phase difference \( \Delta\phi \) and of the cyclic phase difference \( \Delta\Psi \); in this simple example, without large jumps in the phase difference, \( \Delta\phi = \Delta\Psi \); the points of (c) are used to compute the histogram of the distribution of cyclic phase difference. Here the maxima of the time series have been used to define the phase of the signal.
\[ \phi(t) = 2\pi \left[ \frac{t - t_n}{t_{n+1} - t_n} + (n - 1) \right], \]

\[ \text{eqn 2} \]

with \( t_n \leq t < t_{n+1}. \)

INDEX OF PHASE SYNCHRONIZATION

Given two signals with phases \( \phi_1(t) \) and \( \phi_2(t) \), synchronization can be detected by seeking a relation between these phases that is fixed over time (Rosenblum et al. 1996; Pikovsky et al. 1997). If the phases are locked in a 1:1 manner, the phase difference between the two signals should satisfy the relationship:

\[ \Delta \phi = \phi_1(t) - \phi_2(t) = \text{constant}. \]

\[ \text{eqn 3} \]

The constant indicates a phase shift between the two otherwise locked signals. Thus sometimes periods of synchronization can be detected simply by plotting the phase difference \( \Delta \phi(t) \) against time \( t \) and looking for horizontal plateaux (Blasius et al. 1999; Rosenblum et al. 2001).

Non-stationarity and noise are ubiquitous in ecology and noise-contaminated dynamics are typical in population biology. These features may mask phase synchronization, sometimes making direct inspection of the phase differences an inconclusive test even though the signals may be phase locked. Theoretically, noise can induce phase slips, rapid changes in the difference of the phases of two signals as they jump in and out of the synchronized state. In these circumstances, it is useful to seek for statistical signs of phase locking. As an aid, consider the cyclic phase difference (Rosenblum et al. 2001):

\[ \Delta\Psi(t) = \Delta \phi(t) \mod 2\pi. \]

\[ \text{eqn 4} \]

This is just the phase difference \( \Delta \phi(t) \) taken modulus \( 2\pi \) (Fig. 1c). The distribution of the cyclic phase differences gives information regarding the presence of synchronization. A peak in this distribution of the cyclic phase indicates there is a preferred value of \( \Delta \Psi \) and thus a statistical tendency for the two signals to be phase locked with a constant phase difference. The width of the histogram’s distribution is a measure of the synchrony. Obviously the ‘thinner’ the histogram, the tighter the synchrony, while a non-synchronous state would have a broad and uniform distribution.

The statistical test that we propose requires some measure of how peaked is the distribution of \( \Delta \Psi(t) \). A useful index is the Shannon entropy (Tass et al. 1998; Rosenblum et al. 2001):

\[ S = \sum_{i=1}^{N_p} p_i \ln p_i, \]

\[ \text{eqn 5} \]

where \( p_i \) is the proportion of counts in the \( i \)th bin of the frequency histogram of the \( \Delta \Psi \) and \( N_p \) is the number of bins of the frequency histogram. For our purposes, we make use of the normalized Shannon entropy:

\[ Q = (S_{\text{max}} - S)/S_{\text{max}}, \]

\[ \text{eqn 6} \]

where \( S \) is the Shannon entropy and \( S_{\text{max}} = \ln N_p \). Note that for a uniform distribution, \( Q = 0 \) and for a dirac distribution, \( Q = 1 \).

TESTING FOR PHASE SYNCHRONY: A SURROGATE DATA APPROACH

The detection of a peak in the distribution of the cyclic phase differences \( \Delta \Psi \) is crucial to our test for identifying synchrony. However, what precise criteria should be used to define a peak? In some situations it might not be obvious whether there is a true peak in the histogram or whether it is an artefact arising from the noise and variability of the data, or perhaps the shortness of the time series. The following approach helps detect whether the peak in the histogram is statistically significant. More precisely we test the null hypothesis:

\( H_0 \): that the observed peak is any different to that expected by chance alone;

against the alternative hypothesis

\( H_A \): that the time series are synchronized.

In conducting the statistical test we make use of surrogate data sets which are known in advance to satisfy the null hypothesis \( H_0 \). The surrogate data sets, by construction, preserve all important features of the observed time series and mimic them well, except for the fact that they are not in any way synchronized. The first type of surrogate (type 1) shuffles or randomizes the locations of the maxima of all cycles/quasi-cycles in the observed time series. (Equivalently, the analysis may be based around the locations of the minima if it is more practical.) Let \( s_i \) and \( t_i \) be the times at which these maxima occur in each of the two observed time series that we suspect may be synchronized. Suppose that there are \( N_p \) maxima that occur at the times \( t_i \) which may or may not be synchronized to the maxima \( s_i \). Construct a surrogate set of times \( t_i' \) as follows. The first point \( t_i' \) is obtained by randomly selecting (with replacement) one of the \( N_p \) values of \( t_i \). Similarly \( t_i' \) is obtained by randomly selecting one of the \( N_p \) values of \( t_i \). Thus for example, we might find:

\[ t_1' = t_3, t_2' = t_5, t_3' = t_1, t_4' = t_3, \ldots \]

\[ \text{eqn 7} \]

Note that because random selection is performed with replacement, as is typical for Efron’s bootstrap (Efron & Tibshirani 1993), repetitions can occur as in the above case where \( t_i' = t_i' = t_i \). The surrogate time series of maxima \( t_i' \) thus preserves the distribution of the original \( t_i \) but their actual order in the time series is randomized.

We then seek to check whether the observed set of maxima \( t_i \) is any more synchronized to the \( s_i \) than the randomized set of maxima \( t_i' \). To achieve this we need some discriminating statistic, say \( Q \), which provides some measure of synchrony when applied to the resulting histogram of \( \Delta \Psi \).
Firstly, as a reference measure, the value of the discriminating statistic $Q_{\text{obs}}$ should be determined to quantify the synchrony of the quasi-cycles in the two observed time series. But it is impossible to say whether or not $Q_{\text{obs}}$ is unusual, or greater than expected, unless something is known about the distribution of $Q$ under the random null hypothesis. This is how the surrogate data sets help. Calculating $Q$ base on a large number of surrogate data sets and plotting the values as frequency histogram gives the distribution of $Q$.

Knowledge of the distribution of $Q$ now makes it possible to decide whether the observed statistic $Q_{\text{obs}}$ is exceptional or not. From the distribution, one can calculate the proportion of surrogate data sets with a value of $Q$ greater than $Q_{\text{obs}}$. If this proportion, $P$, is large (e.g. $> 5\%$) then the ‘no synchrony’ null hypothesis cannot be rejected. On the other hand, if $P$ is small (e.g. $< 5\%$), the null hypothesis $H_0$ should be rejected and the alternative hypothesis $H_A$ of synchrony accepted.

For periodic series, as the time distance between maxima (minima) is constant, the surrogate series generated by this simple method is identical to the raw series. Although this is an extreme case, it shows how the surrogate method might sometimes lead to a lack of power when testing for synchronization. Therefore a second type of surrogate (type 2) is also suggested. Type-2 surrogates mimic the observed time series in their entirety and are not just restricted to the maxima (minima). The surrogates are obtained by resampling the observed data based on a Markov process scheme that preserves the short temporal correlations in the time series. Technically these surrogates are produced in the following way:

1. The time series $\{x_n\}$ is binned to form a frequency histogram of $N_{\text{e}}$ equal sized bins that estimates the distribution of the $\{x_n\}$.
2. A transition matrix $M$ that describes the time evolution from bin-$i$ to bin-$j$ is then estimated based on the actual relative frequencies of the data contained within bin-$i$. Namely:

$$M_0 = \Pr(x_{n+1} \in b_j \mid x_n \in b_i).$$  

3. To construct a surrogate time series $\{v_n\}$, begin by choosing an initial value $v_0$ by randomly sampling from the raw series $\{v_n \in \{x_n\}\}$.
4. Suppose now that we need to determine $v_{n+1}$ having obtained already $v_n$ ($n = 0$ if beginning at the first step). We also known that $v_n$ is associated with a particular histogram bin, say $b_i$. The probability that $v_n$ from bin $b_i$ ends up in bin $b_j$ is given by the probability $M_{ij}$. Using the probabilities set by matrix $M$, randomly select the bin $b_j$ to be associated with $v_{n+1}$. Choose $v_{n+1}$ by making a random selection from the elements in bin $b_j$.
5. Iterate the last step to obtain a surrogate series of the same length as the raw series.

This surrogate definition preserves both the temporal pattern and the short time correlation of the observed values. Figure 2 illustrates the process with a time series generated based on a AR(2) model. The type-2 surrogate data has a time evolution that is very similar to the time series of the AR(2) model (e.g. the same distribution and transition probabilities).

**Results**

**PHASE ANALYSIS OF THEORETICAL COMMUNITIES**

To illustrate phase analysis we first test the method on short synthetic time series generated by non-linear foodweb models. Blasius *et al.* (1999) described a model in which vegetation, herbivores and predators oscillate with cycles that recur regularly (because of the model’s *Uniform Phase* evolution) but the maxima of the cycles are chaotic (*Chaotic Amplitudes*). The model’s so-called UPCA dynamics generates subtle forms of synchronization when two or more such oscillators are coupled. We used the two-patch modified version suggested by Cazelles & Boudjemâa (2001) without dispersion between patches and with external environmental forcing. The model reads:

$$\frac{dx_i}{dt} = f_i(x_i, y_i) = a_i f_1(u_i, v_i) + \xi_{w_i}$$

$$\frac{dy_i}{dt} = -h_i v_i + \alpha_1 f_1(u_i, v_i) - \alpha_2 f_2(v_i, w_i) + \xi_{w_i}$$

$$\frac{dw_i}{dt} = -c(w_i - w^*) + \alpha_2 f_2(v_i, w_i) + \xi_{w_i}$$

where $i = 1, 2$ is the patch index, $u$ the primary producers, $v$ the herbivores and $w$ the predators. The parameter $a$ represents the primary producer growth rate, $k_0$ the primary producer carrying capacity, $h$ the natural mortality of the herbivores, $c$ the mortality rate of the predators, and $\alpha_1, \alpha_2$ the interaction strengths. The growth of the primary producers is modelled by a logistic function and the interactions by a Holling type II function: $f_i(x, y) = \frac{xy}{1 + k_i x}$, with the $k_i$ being parameters...
that scale according to the half saturation rates of the interactions. Except for the herbivore mortality rates $b_1$ and $b_2$, both patch populations are given identical parameter values. In addition, environmental forcing is added at each trophic level $f_i = \xi_i = \alpha \cos(\Omega t) + \eta$, with $A$ and $\Omega$, the amplitude and the frequency of the forcing disturbance, respectively, and $\eta$ an additive gaussian noise component with zero mean and variance $\sigma^2$.

Cazelles & Boudjema (2001) showed that for an appropriate external environmental forcing, populations of the otherwise uncoupled patches, become phase-locked to the same collective rhythm although their abundances remained chaotic and uncorrelated (see Fig. 3). This result generalizes the Moran effect to unlinked oscillating populations having non-identical and non-linear complex dynamics.

We use the above statistical methodology to test whether there is detectable synchronization between two predator time series generated from each patch of Cazelles & Boudjema’s (2001) model. The analysis is based on short time series of approximately 10 quasi-cycles containing 100 points in total. The time series were deliberately left ‘short’ to reflect the length of natural time series.

Figure 3a shows the evolution of the predator dynamics in the two patches. A visual inspection indicates a well-defined association between these two series; to each oscillation in one patch there often (but not always) corresponds a simultaneous oscillation in the second patch. The two time series appear to be synchronized and in-phase. However the classical cross-correlation fails to indicate a significant in-phase (lag 0) correlation (see Fig. 3b). This is because even though the two time series are reasonably locked in phase, their amplitude dynamics are chaotic and uncorrelated. Nevertheless the cross-correlation coefficient indicates a significant link between the one oscillation of one patch and the previous (or the next) oscillation of the other (i.e. lag $-7$).

Figure 3c shows the distribution of the cyclic phase difference of the two time series has a sharp peak at $\Delta \Psi = 0$ (i.e. at lag 0), testifying the visual impression of synchronization in Fig. 3a. The cyclic phase difference fluctuates around the preferred values $\Delta \Psi = 0$. As further confirmation, the surrogate null test was applied. Recall that this checks whether the ‘peakedness’ of the histogram of the cyclic phase differences might be expected by chance, against the alternative hypothesis that it is a significant indicator of synchronization between the series. The test statistic, the normalized Shannon entropy of the histogram, is $Q_{obs} = 0.27$.

Figure 3d plots the reverse cumulative density function of the statistic $Q$, estimated from $N = 500$ surrogate (type-2) data sets. For any value of the statistic $Q$, one can directly read off this graph the proportion $P$ (vertical axis) of all surrogates that had a statistic $Q$ greater or equal to $Q$. In particular it is easy to look up the proportion of surrogates that had a statistic $Q$, greater or equal to the observed value $Q_{obs}$, i.e. $Pr(Q \geq Q_{obs})$. In this case $Pr(Q_5 \geq Q_{obs}) < 0.002$. That is, not one of the surrogates had a value $Q$ as high as the observed data $Q_{obs}$. Hence the null hypothesis can be rejected in favour of the alternative hypothesis that the time series are significantly synchronized. These results are confirmed by the analyses of longer time series generated by the same model (Cazelles & Boudjema 2001).

The second example analyses two AR(2) time series simulated with two independent noise realizations:
Imperfect population synchrony

\[ x_n = a + bx_{n-1} + cx_{n-2} + \zeta_n \]  

Eqn 10

with \(a, b, c\) the model parameters and \(\zeta\) an i.i.d. gaussian noise component.

The simulations were run for \(N = 100\) time steps producing approximately 10 quasi-cycles. Figure 4 displays the results of the analysis. By chance in this example, the two time series are very similar (Fig. 4a) and the cross-correlation indicate a strong significant cross-correlation at lags \(-1\) and \(-2\) (Fig. 4b). The inspection of the distribution of the cyclic phase difference shows a peak around \(-\pi/3\) (Fig. 4c) and the normalized Shannon entropy of the histogram is \(Q_{obs} = 0.125\). Figure 4d shows that the proportion of \(N = 500\) type-2 surrogates having \(Q\) greater or equal to \(Q_{obs}\) is 30%, i.e. \(\Pr(Q \geq Q_{obs}) = 0.3\), meaning that this association between the phases of these series is not significant. Thus the null hypothesis cannot be rejected, and the apparent association between these two noisy time series may be due to chance. Of course, we know that this is how it should be since the initial test data was constructed to be unsynchronized.

These examples have shown that phase analysis can reveal weak forms of association between noisy time series and can also discriminate between false and true phase associations between short noisy series.

**PHASE ANALYSIS OF POPULATION ABUNDANCES**

Here we apply the phase analysis approach on observed time series to identify synchronization patterns between actual populations. For the computation of the phase of a considered time series, a key point of the proposed approach is the determination of the maxima (minima). Sometimes this determination is not easy in noisy observed time series. This point will be addressed in the Discussion section.

We begin by analysing the Dungeness crab populations (Higgins et al. 1997b; Higgins, Hastings, & Botsford 1997a). This data set consists of yearly catch records of male crabs from 1951 to 1992 at eight locations ranging over 1000 km of the Pacific coastline. The populations show large amplitude fluctuations with a period of about 10–11 years. Higgins et al. (1997b) showed that Dungeness crab population dynamics are governed by intertwined density-dependent mechanisms and exogenous disturbances. The populations are connected by the pelagic dispersal of larvae, but this dispersal effect was thought to be insignificant and therefore was not included in their model. On the other hand, these authors emphasized with model simulations that environmental disturbances may lead to large fluctuations with multiyear cycles that seem to be locked. Figure 5 displays the results of our analysis based on two spatially amalgamated time series (see Fig. 1 in Higgins et al. 1997a). Figure 5a shows the time evolution of the amalgamated population based on two distant groups of locations and Fig. 5b displays the evolution of the phase of these two series as determined from the minima of the data sets. Despite the good visual association between the raw data sets, it should be noted that cross-correlation coefficients (Fig. 5c) between these two spatially amalgamated time series are of borderline significance (\(P = 0.05\)). Nevertheless, the phase analysis results (Fig. 5d,e) strongly suggest that the populations fluctuate in synchrony. Figure 5e displays the reverse cumulative density
function obtained from $N = 500$ type-1 surrogate data sets, and based on the normalized Shannon entropy. The observed entropy of the Dungeness crab populations is $Q_{obs} = 0.50$. Not one of the surrogates had a $Q_s$ value this high, i.e. $\Pr(Q_s \geq Q_{obs}) < 0.002$. Figure 6a, b generalizes the previous results when taking into account all eight crab times series. The distribution of cyclic phase differences calculated for all pairs of Dungeness crab populations shows a significant peak ($\Pr(Q_s \geq Q_{obs}) < 0.002$). In all cases the cyclic phase difference fluctuates around the preferred value $\Delta \Psi = 0$, signifying that the dynamics are phase-locked for all population pairs.

We also analysed the Canadian lynx data, which is perhaps the most famous example of multiannual cycles in vertebrate populations. Phase analysis was implemented on the four longer lynx series from the 19th century in central Canada (MacKenzie River, Athabasca Basin, Winnipeg Basin and North Central; Elton & Nicholson 1942; Keith 1963) and the results are presented in Fig. 6(c and d). The results make clear that the phases of these four populations are synchronized over all the considered regions of Canada, confirming previous analyses (Blasius et al. 1999; Stenseth et al. 1999).

Finally, we analysed another known and impressive example of population synchrony noted in Finnish tetraonid populations (Lindström et al. 1995; Ranta, Lindström & Lindén 1995b; Lindström, Ranta & Lindén 1996). The three observed species (black grouse, hazel grouse and capercaillie) come from route censuses performed in 11 provinces over 20 years (1964–83) and display cyclic dynamics with a periodicity of 4–6 years. We applied phase analysis on the three species. Figure 6(e,f) shows the distribution of cyclic phase differences computed over all pairs of black grouse populations in the 11 provinces (similar results were obtained with the hazel grouse and capercaillie). Cyclic phase difference fluctuated around the preferred value $0$ ($\Pr(Q_s \geq Q_{obs}) < 0.002$). This confirms previous speculations of synchrony obtained by cross-correlation analysis (Lindström et al. 1995; Ranta et al. 1995b).

These observations of synchrony, at least in the sense of phase, over quasi-continental scale may be explained by invoking the Moran effect (Moran 1953; Royama 1992). Assuming that the different local crab, lynx or tetraonid populations have the same structure of density dependence, any density-independent environmental factor can synchronize them. Most likely, dispersal and the Moran effect may act together to enhance the synchronization of these populations (Ranta, Kaitala & Lindström 1999). Thus, in the next section we will check the potential influence of climatic oscillations.

---

**Fig. 5.** Synchrony between populations, the example of Dungeness crab populations. (a) Evolution of the two amalgamated time series. (b) Time evolution of the phases of these two series. In this example the minima of population time series were used to define the phases (see Fig. 1). (c) Cross-correlation $\rho$ between time series; bold lines indicate significant correlation coefficient levels with $\alpha = 5\%$ based on a white noise hypothesis (see Box & Jenkins 1976: Chapter 11); dashed line these same 5% levels computed with $N = 500$ surrogate time series (type 2). (d) Frequency distribution of the cyclic phase difference between the two time series. (e) Reverse cumulative density function of the discriminating statistic, the normalized Shannon entropy, estimated with $N = 500$ surrogate time series (type 2). Arrow indicates the observed value of the normalized Shannon entropy characterizing the observed cyclic phase distribution (d).
Environmental variations and environmental forcings can play an important role in population biology, not only in generating cyclic patterns of fluctuation as suggested, for example, by studies of rodent populations (Saitoh, Stenseth & Bjørnstad 1997, 1999; Bjørnstad et al. 1998; Stenseth, Bjørnstad & Saitoh 1998a; Stenseth et al. 1998b; Hansen, Stenseth & Henttonen 1999a,b), but also in causing spatial synchrony over large geographical areas (Ranta et al. 1995a, 1997a,b; Lindström et al. 1996; Sutcliffe et al. 1996; Grenfell et al. 1998; Stenseth et al. 1999). We therefore applied the phase analysis to identify relations between population abundances and environmental signals as quantified by standard climatic indices. Fluctuations in climate are strongly correlated with interannual variation in atmospheric circulation. Indices of these oscillations have been constructed based on the difference in normalized sea level pressures between different locations: the North Pacific Index (NPI) for the northern Pacific, and the North Atlantic Oscillation (NAO) for the northern Atlantic (www.cgd.ucar.edu/cas/climind). For instance, the NAO affects the pattern of temperature, precipitation and wind over the Northern Hemisphere. High positive NAO values are associated with strong wind circulation in the North Atlantic, moisture and warm temperatures in western Europe and low temperatures on the east coast of Canada (Hurrell 1995). Conversely low NAO values are associated with cold and dry climate in western Europe.

First we tested the relationship between the NAO oscillations and the irregular fluctuations of the sheep population in the St Kilda archipelago (Grenfell et al. 1998). Milner et al. (1999), studying the influence of climatic fluctuations on the survival of Soay sheep, showed that despite a visual impression of a strong association between population abundance and the NAO climatic index, the two variables were in fact not significantly correlated. We obtained identical results with the sheep population from Hirta (see Grenfell et al. 1998), a good visual impression of link between the two time series (Fig. 7a), but the correlation between the two variables was not significant (Fig. 7c). We rechecked this result by applying phase analysis. Figure 7b displays the time evolution of the phase derived from the two series of the sheep abundance and the NAO winter index. The distribution of cyclic phase difference is plotted in Fig. 7d. This distribution shows a clear peak for $\Delta \Psi = 2\pi/3$ indicative of a synchronization of the phases in these time series. The observed cyclic phase difference distribution is significantly different from the distribution under the null hypothesis with $\Pr(Q \geq Q_{\text{obs}}) \approx 0.008$, as based on $N = 500$ type-1 surrogates.

Because phase analysis uses quasi-cycles with a constant duration $2\pi$, the results presented in the Fig. 7d can be interpreted as showing the existence of a delay of two-thirds of a quasi-cycle between the peak of the NAO winter index and the peak of the sheep population. This merely reflects the fact that several years are required for the sheep population to recover from the occurrence of harsh winters (corresponding to peaks in the NAO winter index). The results are in agreement...
with the recent work of Forchhammer et al. (2001) who showed that high NAO winters depressed juvenile survival of Soay sheep population in the island of Hirta (St Kilda). This decrease of juvenile survival may then induce a delay in the population abundance peak.

Recently a large number of studies have shown that environmental variation can play an important role in generating cyclic patterns of vole fluctuations (Saitoh et al. 1997; Bjørnstad et al. 1998; Stenseth et al. 1998a,b; Hansen et al. 1999a,b; Saitoh, Bjørnstad & Stenseth 1999). We analysed the relation between the abundance of vole populations and a regional winter climatic index: between grey-sided voles in northern Finland (Hansen et al. 1999a,b) and NAO (Fig. 8); between total voles in northern Finland (Hansen et al. 1999a,b) and NAO (Fig. 9a, b); between grey-sided voles in Japan (Saitoh et al. 1997) and NPI (Fig. 9c,d). The Finnish data set includes a 48-year time series (1949–96) of autumn abundance of the vole community and these populations have a 4 – 5 year cycle (Hansen et al. 1999a,b). The Japanese data set comes from northern Hokkaido (Japan) and the autumn time series used span 31 years (1962–92) (series #15 in Saitoh et al. 1997), similar results not shown are obtained with the series #39).

The phase analysis demonstrates a strong and significant phase relation and thus synchronization between population abundances and climatic fluctuations in all cases (Figs 8d, e, 9a – d). In contrast, cross-correlation analysis always failed to detect any significant signs of synchronization (Fig. 8c). Instead of quasi-cycles with an irregular period length, phase analysis uses quasi-cycles with a constant duration, $2\pi$. Thus, Figs 8d and 9a, with an average phase difference around $\pi$, can be interpreted as showing the existence of a delay of a half quasi-cycle between the trough of NAO winter index and the trough of the vole population, independent of the duration of each quasi-cycle. For example, if the duration of the given winter NAO quasi-cycle is 4 years, the value of this is 2 years. These results agree with the conclusions of Hansen et al. (1999a,b), who pointed out that winter regulation appears to be an important element of vole dynamics in this region, and suggested that long severe winters, coupled with delayed density dependence, may be a direct cause of the observed population fluctuations.

Our last example concerns the tetraonid populations in Finland. The phase analysis approach was used to analyse the links between the population of black grouse and the NAO winter index (Fig. 9e,f) (similar results were obtained with hazel grouse and capercaille). Again, the results indicate a significant relation between climatic winter oscillations and population abundance (Fig. 9e). These results, together with those in Fig. 6e, may also underline that climatic winter conditions drive and maintain the spatial synchrony of tetraonid populations over the entire region.
Fig. 8. Synchrony between climatic signal and population abundance, the example of the grey-sided vole population in northern Finnish Lapland (Hansen et al. 1999a,b). (a) Evolution of the grey-sided vole population (solid line) and the winter NAO index time series (dashed line). (b) Time evolution of the phases of these two series. In this example the minima of population time series were used to define the phases (see Fig. 1). (c) Cross-correlation $\rho$ between time series; bold lines indicate significant correlation coefficient levels with $\alpha = 5\%$ based on a white noise hypothesis (see Box & Jenkins 1976: Chapter 11); dashed line these same 5% levels computed with $N = 500$ surrogate time series (type 1). (d) Frequency distribution of the cyclic phase difference between the two time series. (e) Reverse cumulative density function of the normalized Shannon entropy, estimated with $N = 500$ surrogate time series (type 1). Arrow indicates the observed value of the normalized Shannon entropy characterizing the observed cyclic phase distribution (d).

Fig. 9. Synchrony between climatic signal and population abundance. (a), (b) Total voles population in northern Finland and NAO. (c), (d) Grey-sided voles population in Japan and NPI. (e), (f) Black grouse population in Finland and NAO. (a), (c), (e) Frequency distributions of the cyclic phase difference between the population and climatic time series. (b), (d), (f) Reverse cumulative density function of the normalized Shannon entropy of the phase difference distribution estimated with $N = 500$ surrogate time series (type 1). Arrows indicate the observed values of the normalized Shannon entropy characterizing the observed cyclic phase distributions. In these examples, for the voles time series, the minima of the population and the climatic index series were used to define the phases, for the tetraonid population, the maxima of the population and the climatic index series were employed (see Fig. 1).
Discussion

Using the phase synchrony concept (Rosenblum et al. 1996; Pikovsky et al. 1997, 2001), we have shown that it is possible to study weak and imperfect synchrony (or interactions) between irregular, non-stationary and noisy time series, which are very common in population biology. The approach has a major advantage over the more classical linear techniques in that it does not share their particularly restrictive requirement of stationarity, where all moments of the time series must be constant in time. The phase analysis method transforms a time series with quasi-cycles of irregular duration (period) to a time series with regular cycles of 2π duration, thereby allowing analysis of the phase of non-stationary time series.

Another advantage of the phase analysis approach is that it is capable of detecting relatively weak signs of synchrony. Indeed, the notion of phase synchronization implies only some interdependence between phases, whereas the irregular amplitudes may remain uncorrelated. The irregularity of amplitudes can mask phase locking so that traditional techniques (e.g. based on the correlation coefficient), that focus on the signals themselves rather than on the phase, may be less sensitive in the detection of the signal dependencies. This approach is in full agreement with Myers (1998), who remarked that when two complicated dynamics are compared the amplitudes are often uncorrelated, but if the phases are studied a well-defined relationship may appear. It is noteworthy that in many of the examples given here a conventional cross-correlation method failed to detect any significant synchrony (dependence) between population abundance (Fig. 5) or between abundance and environmental signal (Figs 7 and 8), whereas calculating the phases clearly showed phase locking, at least in a statistical sense.

Recently two other methods for the phase computation have been employed in ecology. The first is based on a wavelet analysis (Grenfell, Bjørnstad & Kappey 2001) and the second based on a pseudo-Poincaré reconstruction (Haydon & Greenwood 2000). For wavelet phase definition one needs to choose a fixed frequency band for the computation of the phase of a time series (Torrence & Compo 1998; Lachaux et al. 2000; Le Van Quyen et al. 2001). However in numerous cases, it is difficult to define an average cyclic pattern with only one significant frequency band, and, in this case, the wavelet approach appears restrictive (see for example Le Van Quyen et al. 2001). Moreover the wavelet approach has never been tested on short time series as those commonly encountered in ecology (i.e. less than 128 data points).

In the approach developed by Haydon & Greenwood (2000), a pseudo-Poincaré section with an embedding dimension $E = 2$ was reconstructed by plotting the population abundance at time $t + \tau$, $N_{t+\tau}$ vs. the population abundance at time $t$, $N_t$ with $\tau$ the lag of the reconstruction. For cycling population this plot, in the reconstructed space $N_t - N_{t+1}$ defined quasi-cycle. First, they identified the beginning and end points of each quasi-cycle, assigning phase values 0 and $2\pi$, respectively. They then computed the phase associated with each point as the proportion of the quasi-cycle transversed since the starting point. The difficulty with this approach lies in defining the beginning and the ending points of the each quasi-cycle especially if the time series is noisy.

Another difficulty with this approach is the choice of the embedding dimension $E$ and the lag $\tau$ for the reconstruction of the Poincaré section by plotting $N_{t+\tau}$ vs. $N_t$ (Kantz & Schreiber 1997). We have also applied this type of phase space reconstruction method using the transformation proposed by Pikovsky et al. (1997):

$$\Psi(t) = \arctan \left( \frac{N_{t+1} - C_y}{N_t - C_y} \right) + \pi, \text{ with } \{C_x, C_y\} \text{ the coordinates of the centre of the quasi-cycles.}$$

Figure 10 summaries the results for theoretical and observed time series based on this phase reconstruction. Figure 10a,b concerns the time series generated by the tri-trophic food web (cf. Fig. 3) while Fig. 10c,d concerns the Dungeness crab populations (cf. Fig. 5). The results with this phase computation are almost identical with those obtained from the approach suggested in this paper. Note though that the approach is very sensitive to the centre definition, i.e. $\{C_x, C_y\}$.

The simple definition of phase used here is based on clear detection of the time series maxima (minima). It is sometimes the case that strong random fluctuations make it difficult to define clearly maxima (minima). Although a subjective approach is possible, sometimes a more objective is preferred. We have found that smoothing the data often overcomes the problem and succeeds in highlighting the locations of the maxima (minima). The simple smoothing methods (e.g. moving averages) often introduce autocorrelations in the reconstructed series that are artefacts, and in some case the phases of the series are modified. A powerful method that overcomes these problems is the singular spectrum analysis (Broomhead & King 1986; Vautard & Ghil 1989; Vautard, Yiou & Ghil 1992). This method decomposes the time series based on the eigenvectors of a projection of the series in an orthogonal space. The main part of the dynamics of the time series is associated with the leading eigenvalues of this projection. Hence with the main oscillating components, determined based on the larger eigenvalues, a smooth series can be reconstructed. This method has been used by Pascual et al. (2000) to extract the main oscillating components of cholora and El Niño time series before spectral analysis. We have also employed the singular spectrum analysis to smooth the sheep population and winter NAO index time series (Fig. 7). We have then applied our simple phase method to the reconstructed series with their main oscillating components. Figure 11 reports the results of this analysis and displays similar significant results (Fig. 7d,e). The simple approach advocated here appears to be robust to different peak definitions.
Our simulations also reveal that the phase analysis based on detection of maxima or minima is often very robust, and often insensitive even to the deletion of some of them. (The robustness increases with the length of the series.) To illustrate this last point, Fig. 12 displays the results after the addition or removal of one maximum in the time series of the sheep population in the St Kilda archipelago (Fig. 7). For the graphs of Fig. 12a,b, we have added a maximum in the sheep population at the year 1977 (see Fig. 7a), and for the graphs in

Fig. 10. Detection of association between two time series by the phase analysis approach; influence of the method used to define the phase of each time series. The phase of each time series is computed based on a pseudo-Poincaré section reconstruction (see the main text). (a), (b) Example of the two predator populations of two non-linked tri-trophic food web (see Fig. 3). (c), (d) Example of Dungeness crab populations (Higgins et al. 1997a) (see Fig. 5). (a), (c) Distributions of the cyclic phase difference between the series. (b), (d) Reverse cumulative density functions of the discriminating statistic, the normalized Shannon entropy, estimated with $N = 500$ surrogate time series (type 2). Arrows indicate the normalized Shannon entropy characterizing the observed cyclic phase distributions (a) or (c).

Fig. 11. Detection of association between two time series by the phase analysis approach, influence of the use of a smoothing technique. Phase analysis is applied to smoothed time series of the sheep population from Hirta island and the NAO winter index (see Fig. 7) reconstructed by singular spectrum analysis (Vautard et al. 1992). (a) Frequency distributions of the cyclic phase difference between the series. (b) Reverse cumulative density functions of the discriminating statistic, the normalized Shannon entropy, estimated with $N = 500$ surrogate time series (type 1). Arrow indicates the normalized Shannon entropy characterizing the observed cyclic phase distributions (a).

Fig. 12. Detection of association between two time series by the phase analysis approach, influence of the number of maxima used. The time series tested are the sheep population from Hirta island and the NAO winter index (see Fig. 7) reconstructed by singular spectrum analysis (Vautard et al. 1992). (a), (b) A maxima is added for the year 1977. (c), (d) A maxima is randomly removed. (a), (c) Frequency distributions of the cyclic phase difference between the series. (b), (d) Reverse cumulative density functions of the discriminating statistic, the normalized Shannon entropy, estimated with $N = 500$ surrogate time series (type 1). Arrows indicate the normalized Shannon entropy characterizing the observed cyclic phase distributions (a) or (c).
Fig. 12c,d we have randomly removed a maximum. In both cases the cyclic phase difference distribution and its statistical significance are similar to those originally found in Fig. 7d,e (Pr \( Q \geq Q_{\text{null}} \) < 0.01).

Quantifying and explaining synchronization of fluctuating populations over large geographical regions present important analytical challenges. We have developed a method based on phase analysis (Rosenblum et al. 1996; Pikovsky et al. 2001) to quantify synchrony between two time series that is appropriate for studying irregularly fluctuating population and/or environmental signals. Our approach analyses the similarity in rhythm within the time series. More recently Haydon et al. (2003) proposed a somewhat different methodology from the one described here that permits the quantification of the degree of dynamic cohesion between population fluctuations. Their approach requires the availability of multiple and/or replicate time series (e.g. spatially explicit ecological data sets) while here we focus on detecting synchronization within two time series.

With the phase analysis approach, patterns of weak synchrony that were previously undetectable are found to be resolvable, or at least statistically likely. We believe further such analyses will contribute substantially towards resolving controversies over population regulation and the causes of spatial synchrony within and between populations. As Lloyd & May (1999) emphasized, the concept of phase synchronization ‘might provide much-needed new tools for population biologists to study their well-worn data sets’.

Acknowledgements

We thank the support of the James S. McDonnell Foundation and of the French Ministry of Education and Research, ACI ‘Jeunes Chercheurs 2000’. We are grateful to Anders Møller, M. Justin O’Riain and Andy Gonzalez for discussions and comments.

References


Received 16 October 2002; revision received 6 June 2003