Linking environmental variability to population and community dynamics

Jelena H. Pantel\textsuperscript{1*}, Daniel E. Pendleton\textsuperscript{2}, Annika Walters\textsuperscript{3}, and Lauren A. Rogers\textsuperscript{4}

\textsuperscript{1}Centre d’Ecologie fonctionnelle et Evolutive, UMR 5175 CNRS-Université de Montpellier-EPHE, Campus CNRS, 1919 route de Mende 34293 Montpellier Cedex 05 France
\textsuperscript{2}New England Aquarium, 1 Central Wharf, Boston, MA 02110, USA
\textsuperscript{3}U.S. Geological Survey, Wyoming Cooperative Fish and Wildlife Research Unit, Department 3166, 1000 E. University Avenue, University of Wyoming, Laramie, WY 82071, USA
\textsuperscript{4}Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway

Abstract

Linking population and community responses to environmental variability lies at the heart of ecology, yet methodological approaches vary and existence of broad patterns spanning taxonomic groups remains unclear. We review the characteristics of environmental and biological variability. Classic approaches to link environmental variability to population and community variability are discussed as are the importance of biotic factors such as life history and community interactions. In addition to classic approaches, newer techniques such as information theory and artificial neural networks are reviewed. The establishment and expansion of observing networks will provide new long-term ecological time-series data, and with it, opportunities to incorporate environmental variability into research. This review can help guide future research in the field of ecological and environmental variability.

Section 1. Introduction

A species’ ecology arises from the complex interplay between intrinsic life history properties and interactions with their environment and with other species (Andrewartha and Birch 1954; Cole 1954; Hutchinson 1957). Disentangling the influence of so many possible agents on observed population dynamics, a core objective of modern ecology, is an increasingly attainable achievement in the current era of massive data collection. As the length of time spent observing natural populations increases and technology to recover historical records of environmental properties or abundance and composition of species improves, ecologists can more readily obtain potentially powerful time series datasets. Along with new opportunities, however, come new challenges. Large data records can better capture the signature of ecological processes than small data records. They can also better capture the signature of nonecological processes such as measurement error, autocorrelation, and stochasticity.

Analysis of time series data may determine whether theorized ecological mechanisms apply to observed population and community dynamics and may facilitate predictions for how these dynamics will be altered in response to changing climate. But this is only possible if the unique challenges of analyzing large data series are overcome. To realize the potential of time series data, we provide guidance from the findings of classical ecological study as well as from conceptual and analytical practices derived from other fields, such as signal processing and information theory. The goal of this review is to identify fundamental properties of environmental and biological variability, explore models that can characterize environmental and biological variability, and examine methods for determining the relationship between environmental and biological variability.
variability. We first provide an overview of features inherent to environmental variability in data series. We then compare and contrast those features with observations of population variability and describe methods to appropriately model and statistically infer properties of variability from observed time series. We further consider how intrinsic life history characteristics and interactions with other species influence populations and their abundance over time. Finally, we review an array of analytical techniques derived from non-ecological fields that can identify the signature of ecological processes embedded in complex data series and can causally link multiple data series to infer their impact on one another.

Section 2. Overview

Linking population or community dynamics to the influence of environmental properties that fluctuate over time requires three critical elements. A researcher must select a biological property of interest $y$ (e.g., population density), an external environmental feature $x$ thought to influence the biological property, and a function $f$ that relates the biological property to the environment at a given time $t$ (Laakso et al. 2001):

$$y_t = f(x_t)$$

(1)

This formulation appears simple, but defining the form and nature of each term requires knowledge of not only population and community ecology, but also potentially of geological, chemical, and physical systems as well as statistical tools that accommodate the temporal ordering, stochastic and deterministic elements, and time and frequency properties embedded in biologically relevant time series. In addition, more than one environmental feature likely influences the focal biological property and the relationship between $x$ and $y$ may also be nonlinear. We highlight past and present information from the variety of fields relevant to ecologists who use time series to ask how an organism’s endogenous properties and exogenous environment influence its ecological dynamics. We have highlighted specialized terms with italics and provide their definitions throughout the text.

Section 3. The nature of environmental variability

There are numerous observations to inform how $x$, an environmental variable thought to influence the ecology of living organisms, changes over time. A ubiquitous feature of variability in natural systems is its tendency to increase with the length of time considered (Steele 1985). This means that measures of environmental fluctuation, such as standard deviation or coefficient of variation, increase as the time window increases, e.g., from 10 to 1000 years (Fig. 1). This pattern of variance growth with time has been observed in numerous geophysical (sea level, stream discharge, ice cover) and climatic (temperature, precipitation, pressure, humidity, CO$_2$ concentration) variables, measured in aquatic (streams, rivers, lakes, oceans) and terrestrial habitats over a broad range of geographies and time spans (Steele 1985; Weber and Talkner 2001; Cyr and Cyr 2003; Wunsch 2003; Vasseur and Yodzis 2004; Sabo and Post 2008).

Another signature of variability in environmental properties is that low-frequency events influence time series variability more than high-frequency events (Steele 1985; Vasseur and Yodzis 2004). This dominance of low-frequency processes is best illustrated when the time series is decomposed into constituent periodic functions (such as $\sin(t)$; Platt and Denman 1975). Placing the distribution of periodic functions that describe a time series into a histogram, sorted by increasing frequency of the functions, creates a spectrum or spectral density distribution (Halley 1996). Despite the range of possible frequency distributions—including ‘white noise’ where the contribution of all frequencies is uniform and ‘blue noise’ with an increased high frequency component—environmental variables relevant to ecological processes reliably display ‘reddened’ noise, characterized by a dominance of cycles with long periods (low frequencies) (Fig. 2; Steele 1985; Vasseur and Yodzis 2004).

On geologic time scales, the overrepresentation of environmental variability at low frequencies is often attributed to Milankovitch cycles (the effects of Earth’s movement on its climate) with 100,000-year periods (McManus et al. 1999; Jouzel et al. 2007; but see Wunsch 2003 for a purely stochastic explanation). At shorter time scales, other physical properties may be involved, such as water’s high heat capacity increasing the ‘memory’ of low-frequency processes in aquatic systems by buffering against short-term environmental changes (Halley 2005; Sabo and Post 2008). However, a third feature of environmental variability contributes greatly to this reddened pattern, positive temporal autocorrelation, where measures at nearby time points are more similar to one another than to measures observed at distant time intervals (Fig. 3). If these autocorrelations decay slowly with time, the system has a ‘long memory’—fluctuations in $x$ at time $t$ influence subsequent $x_{t+k}$ values for long time lags $k$ and thus increase the contribution of low-frequency events to series variability (Granger and Ding 1996). Given the observed relationships between variance and time, accurate models of environmental properties must include some form of temporal autocorrelation, or association between observations gathered over successive units of time.

Though variance growth with time, dominance of low frequencies, and positive autocorrelation are common properties of environmental time series, these properties can differ among habitats. Variability is much higher in terrestrial systems than in marine at all time scales (Steele 1985). Terrestrial habitats also display constant variance (white noise) over short time intervals and increasing variance (red noise) from about 50 to 100 years onwards, whereas marine environmental properties tend to show variance growth at all time scales (Steele 1985; Cyr and Cyr 2003). This terrestrial versus marine pattern is consistent with the hypothesis that water’s
heat capacity buffers aquatic systems against more frequent fluctuations (Halley 2005; Sabo and Post 2008). Cyr and Cyr (2003) also found increased dominance of low-frequency cycles in spectral distributions as the size of aquatic systems increased from river to lake, Great Lake, and ocean temperature recordings.

Incorporating the fundamental properties of environmental variability—variance growth with time, dominance of low
frequencies, and positive autocorrelation—will increase the accuracy of models meant to represent these processes. One model to characterize the variation accompanying environmental signals is ‘\(1/f\)-noise’ (Keshner 1982). \(1/f\)-noise was first observed in signals processed from current-carrying vacuum tubes that contained unexpected variability at low frequencies (Johnson 1925; Schottky 1926) and has since proved ubiquitous in physical, biological, astronomical, economic, and even musical and linguistic systems (Press 1978; Voss and Clarke 1978; West and Shlesinger 1990; Gisiger 2001; Li and Holste 2005). \(1/f\)-noise is a random process with a spectral density \(S(f)\) inversely related to frequency (\(f\)) raised to a power \(\beta\) (\(S(f) \propto 1/f^\beta\)), defined over \(0 < \beta < 2\) (Keshner 1982; Halley 1996). When \(\beta = 0\), this process produces white noise, where each point in the time series is drawn independently from the distribution of possible values, the signal is not autocorrelated, and so the spectral density has no relationship with frequency. \(\beta = 2\) produces a random or Brownian walk (brown noise) where each point in the time series equals the previous point’s value plus a constant representing the average change between points. Because of the importance of recent values, brown noise has a shorter ‘memory’ than reddened noise (\(\beta = 1\)). Values of the variable under consideration from the distant past reach their maximal importance when \(\beta = 1\) (pink or reddened noise), leading to a dominance of low frequencies in the spectral density distribution commonly observed in time-series of environmental variables (Keshner 1982; Halley and Inchausti 2004).

Following broad introduction to ecologists and evolutionary biologists in a seminal 1985 paper by John H. Steele, \(1/f\)-noise has persisted as a reliable and flexible choice for modeling abiotic environmental quantities (Halley 1996; Halley and Inchausti 2004; Vasseur and Yodzis 2004). This is largely due to properties of \(1/f\)-noise ‘memory.’ Autocorrelation gives a system ‘memory’ because current values are dependent

Fig. 2. Reddened spectral density of NGRIP \(^{18}\)O time series. Spectral density was estimated in R (version 3.0.0, R Core Team 2013) using a fast Fourier transform and a modified Daniell smoothing kernel with two dimensions. This plot shows estimated spectrum values increase at low frequencies (the log of both spectrum and frequency values are plotted).

Fig. 3. Variance increases with length of time considered for autocorrelated series. Time series generated using (A) uniform (where \(x\) values at each time point \(t\) are chosen at random uniformly over a fixed range) and (B) normal (where \(x\) values at each time point \(t\) are chosen at random from a normal distribution) models vary around a mean value over time, but this variance is fixed at one true value and thus estimates of standard deviation do not vary with time. Time series generated using (C) a Brownian random walk (where \(x_{t+1}\) is generated by successively adding a random normally distributed value \(\alpha\) to \(x\)) also vary around a mean value but the variance measure increases with time (D). Figure after Ariño and Pimm (1995), Fig. 1.
time considered (Keshner 1982; Halley and Inchausti 2004). Distant past events are most influential for current values when $\beta = 1$ and autocorrelation decay is logarithmic, slower even than a power of time (Keshner 1982). Environmental quantities whose observed power spectra do not adhere to a $1/f$ process may fit better to an AR model or values of $\beta = 0$. This is more likely for observations drawn from short time intervals, especially in terrestrial ecosystems (Cyr and Cyr 2003; Vasseur and Yodzis 2004).

**Section 4. The nature of population variability**

If a species’ population dynamics are strongly influenced by their external environment, some reflection of that environment’s variance properties are expected in the population time series. Though variance growth over time and reddened spectral densities are observed in population surveys, the difference in variance among habitats (marine, terrestrial, or otherwise) is not observed (Pimm and Redfearn 1988; Gaston and McArdle 1994; Ariño and Pimm 1995; Inchausti and Halley 2002). A survey of 544 populations from 123 species produced a near-symmetric distribution of spectral exponents (an estimation of $-\beta$, of the closest $1/f^\beta$-process) with a median value of $-1.02$ (reddened spectra) and a decelerating rate of variance growth. However, no influence of taxonomic group, trophic level, latitude, or habitat type was detected (Inchausti and Halley 2001, 2002).

Though the underlying drivers of population versus abiotic environmental autocorrelation structure may vary, time series of both types are effectively described by $1/f$-noise models. A breadth of tools to generate or analyze $1/f$-noise in ecological time series now exists. For example, Cuddington and Yodzis (1999) use a spectral synthesis approximation method to generate a distribution of amplitudes and periods with a desired spectral exponent that is then sent to an inverse fast Fourier transform to generate a time series. The spectral mimicry method (Cohen et al. 1998, 1999) was developed to generate new or modify existing time series that possess identical elements but in different orders, so that only their spectral density distributions differ. Using this method to generate regimes of environmental variability allows researchers in theoretical (e.g., Hiltunen et al. 2006; Vasseur 2007a) and empirical (Reuman et al. 2008) settings to isolate how the structure of environmental stochasticity influences population and community dynamics.

A method to generate colored time series that are also correlated with one another, phase partnering, extends the utility of colored noise models to multivariate ecological systems (Vasseur 2007a). Miramontes and Rohani (2002) devised a method known as multiple segmenting to estimate the spectral exponents of short-term (<100 points) time series, a situation that is challenging but common for ecologists. A variety of methods to statistically evaluate properties of time series to test particular hypotheses exist as well. Rouyer et al. (2008) developed a series of null models with colored noise to distinguish the signal of nonstochastic forcing from stochastic signals in wavelet analysis of time series. Lindström et al. (2012) used hierarchical Bayesian modeling to estimate posterior distributions of autocorrelation parameters and compare these spectral properties among different time series.

The lack of complete concordance between variance properties in environmental and population time series raises an essential question: what factors besides environmental forcing influence observed population dynamics? It may be that ecologists must broaden their definition of ‘the external environment’ when seeking drivers of population variability. We explored biotic properties that may explain variability in populations: the organism’s life history traits and its interactions with other species in the community.

**Section 5. Life history influences population variability**

Any attempt to determine the cause of population variability over time is incomplete without considering the relative influence of intrinsic (demographic and density-dependent) versus extrinsic (e.g., interactions with environment or with other organisms) factors. Our exploration benefits greatly from basic principles established by Robert May (1976) and colleagues (May et al. 1974) who almost forty years ago found that species-specific properties, such as intrinsic growth rate ($r$) and generation time ($\tau$) play a large role in determining how population densities vary (Box 1 provides an applied example of these principles to fishery populations). A population’s characteristic return time ($t_R$) determines how quickly that population returns to equilibrium density after a disturbance. Characteristic return time is inversely proportional to $r$ ($t_R = 1/r$) in logistic and some other simple population growth models (see Vasseur 2007b for characterization of return times in a more complex consumer-resource ecological system using bioenergetics models), as larger values of $r$ increase the rate of a population’s approach to carrying capacity. Through a combination of models and laboratory growth experiments, May (1976) and colleagues (May et al. 1974) found that populations with rapid growth rates (and thus quick return times) were more likely to reflect environmental variability in their population density than populations with slower growth rates (and thus longer return times). Slow-growing populations tend to average over environmental variability, leaving few or no environmental signals in their density record. Population densities do not respond to frequencies of environmental variability that exceed their return time ($f_T > t_R$), as equilibrium population densities change before they are approached, but do track the environmental variability that occurs at low frequencies ($f_T < t_R$) (Vasseur 2007b). Variability in population density is also influenced by the organism’s generation time ($\tau$). Populations exponentially return to their equilibrium value when $t_R > \tau$, but become unstable and overshoot or oscillate around equilibrium when $t_R < \tau$ (May et al. 1974).

In addition to generation time ($\tau$), a population’s ability to track environmental variability is also influenced by the
relationship between its growth rate \( (r) \) and body mass \( (m) \). Allometric scaling, where a broad variety of ecological and physiological properties scale exponentially with body mass, is a well-supported theory that is conserved across taxonomic groups (West et al. 1997). As population growth rate scales as \( r \sim m^{-0.25} \), so does the response time scale as \( t_r \sim m^{0.25} \), indicating that small-bodied organisms have a shorter response time than large-bodied organisms (Savage et al. 2004; Vasseur 2007b). Larger organisms with their longer response times thus tend to average over environmental variability at a wider range of frequencies than smaller-bodied organisms. This increased influence of low-frequency environmental processes for large-bodied organisms should produce a reddened spectral density, which is observed in population dynamic models and surveys of population time series (Kaitala et al. 1997; Inchausti and Halley 2002). Indeed, the survey of Inchausti and Halley (2002) found a positive association between population variability and body size and that larger species had more reddened time series.

**Section 6. Community interactions influence population variability**

Species interactions may leave detectable signatures on a population’s abundance over time. As computing capability, data collection methods, and statistical techniques have improved, the presence of such signatures are increasingly detected in natural systems. For example, Taylor’s power law gives a null expectation for the relationship between a population’s variance \( (\sigma^2) \) and mean abundance \( (\mu) \) over time \( (\log \sigma^2 = log a + b \log \mu) \), with a slope of \( b = 2 \) when populations experience constant environmental variability (Taylor 1961). Though log variance and log abundance are highly correlated in surveys of natural populations, the slope \( b \) is typically between 1 and 2 (Taylor and Woiwod 1982; Kendal 2004). Kilpatrick and Ives (2003) successfully simulated populations with \( b < 2 \) using models of population growth that allowed competition among species, indicating that interspecific interactions alone can explain deviations of populations from growth dynamics expected in isolation. Another study (Bjornstad et al. 2001) seeking evidence that interspecific interactions mark a focal species’ population dynamics reared a moth population in isolation, with a strongly interacting parasitoid, and with a weakly interacting (but highly specialized) virus. The number of parameters needed to fit the moth’s population dynamics to a time series model increased for the population raised with the parasitoid relative to the population raised in isolation. There was no increase in model dimensionality for the population raised with the virus, suggesting that comparative studies of dimension number among populations may reveal species interaction links.

Because species interactions do leave detectable signals on a focal population’s time-series, it should be possible to infer the presence and strength of trophic interactions from time-series of multiple species. However, a large amount of information is captured in a population’s abundance time-series, including environmental, intra-, and interspecific influences.

**Box 1. Marine fisheries and temporal environmental variability**

Understanding how marine fish populations respond to environmental variability has been an area of active research for over a century, driven by the need to manage fisheries harvests (Hjort 1914; Lehodey et al. 2006). The abundance of fish stocks varies over interannual, decadal, and even centennial time scales (Spencer and Collie 1997; Rogers et al. 2013), and many studies have sought to determine how these modes of variation might be related to changes in the environment. The link between Pacific salmon \( (Oncorhynchus spp.) \) abundance and an interdecadal mode of climate variability operating in the North Pacific (the Pacific Decadal Oscillation) is a particularly striking example of how fish populations can respond to environmental variability, in this case at relatively low frequencies (Mantua et al. 1997).

Different species show different patterns of variability in their abundance, with some varying from year to year, and others demonstrating relative stability on interannual time scales, but strong long-term trends or cycles (Spencer and Collie 1997). These differences between species can be attributed to their life histories, with fast growing, short-lived species generally showing increased interannual variation in abundance, whereas long-lived species with slower growth rates are generally more stable (Spencer and Collie 1997). This stability is, at least, partly due to the buffering associated with having multiple age classes in a population, such that higher frequency environmental variability is filtered out (Berkeley et al. 2004).

Fisheries exploitation can alter the way that populations respond to environmental variability, in particular by altering the age-structure of a population (Ottersen et al. 2006; Hsieh et al. 2010; Planque et al. 2010). Larger, older fish are generally targeted by fisheries and high harvest rates can result in populations dominated by smaller, younger individuals (Berkeley et al. 2004). Fisheries-induced evolution may also contribute to fish reaching maturity at a younger age (Law 2000), further altering the age structure of a population as well as shortening the generation time. Cod and herring populations have been found to track environmental variability more closely once their age-structure has been truncated (Ottersen et al. 2006; Rouyer et al. 2011), and increased mortality in general has been shown to increase the high-frequency variation in fish abundance (Rouyer et al. 2012). These examples show how the demographic properties of a population (e.g., age structure or generation time) can influence the relationship between population density and environmental drivers.
Correlations among static values of population densities are likely unreliable for inferring the direction and magnitude of interactions among species. For example, strong competitors should have a negative interaction, but a weak or nonexistent interaction may be observed if the two species compete strongly and respond similarly to environmental fluctuations. This situation grows more complex when, as is often the case, environmental variables are also temporally autocorrelated (Ripa and Ives 2007).

Regression models that consider time lags are a promising method for improving the quality of information extracted from multi-species time series data. An approach increasingly used to estimate the strength and direction of species interactions is first order multivariate autoregressive, MAR(1) or MAR, models. A MAR model applied to multi-species time series data estimates the strength and direction of interactions among species by regressing the abundance or biomass of each species against the abundance or biomass of all other species at the previous time step. Exogenous variables, such as sea surface temperature, can be included as covariates (Ives 1995; Ives et al. 1999; Hampton et al. 2013). The form of a MAR model can be written as

\[ X_t = A + BX_{t-1} + CU_{t-1} + E \]  

(2)

where \( X_t \) is a vector of abundances for each species at time \( t \), \( A \) is a vector of intrinsic growth rates for each species, \( B \) is the interaction matrix whose elements describe the influence of species \( j \) on the per capita growth rate of species \( i \) (the diagonal of \( B \) represents density dependence), the elements of \( C \) describe the effect of covariate \( j \) on species \( i \), \( U_{t-1} \) is the vector of covariate values at time \( t-1 \), and \( E \) is the vector of process errors drawn from a multivariate normal distribution with mean 0 and covariance matrix \( Q \). Once the elements of the interaction matrix \( B \) have been estimated, eigenvalue analysis can also be performed to quantify stability of the system with respect to the environment, asymptotic return time of the system to a stable solution, and reactivity of the system to external perturbations (Ives et al. 2003).

Accurate models of population time series that appropriately consider autocorrelation and community interactions allow researchers to ask intriguing questions in their systems and about their sampling strategy (see Box 2 for examples in aquatic communities). For example, Telfer et al. (2010) found that a field vole population’s susceptibility to infection by most individual microparasite species was best explained not by host or environmental properties, but by that population’s infection status for other microparasites. This result gives one example of the exciting potential for community-wide time series analyses to address not only responses to climate change at the community and ecosystem level, but also the mechanisms driving these responses.

### Section 7. New approaches for reducing complexity

Our initial formulation of the link between population...
dynamics and their environmental drivers, as some function \( f \) relating population density \( y \) to an external environmental property \( x \) at a given time \( t \), may be accurate but fails to address the complexity needed to accurately and fully describe ecological systems. Environmental variables have complex temporal dynamics and species vary in life history parameters, which influence population dynamics. The shape of the functional relationship between \( x \) and \( y \) dictates how an environmental signal is transformed into a biological outcome and remains empirically underexplored. The influence of this relationship may be quite large, as demonstrated by Laakso et al. (2001). In their study, time series of \( x \) generated with the same autoregressive process but filtered through an asymptotic, symmetric, and peaked, or sigmoidal function produced time series of \( y \) with very different frequency characteristics. Correlations between \( x \) and \( y \) changed signs or disappeared altogether, and in some instances, the filter \( f \) altered the ‘blue’ environmental signal \( x \) to produce a reddened biological signal \( y \). The properties inherent to environmental and population variability and the nature of the relationship between them are at once complex and necessary to understand for accurate modeling and interpretation of ecological time series. Newer methods or methods from other fields may provide approaches for reducing complexity. Here we review three promising methods: information theory, convergent cross mapping, and artificial neural networks.

**Information theory** provides a promising alternative formulation to our question of how to link population and community dynamics to one another and to environmental properties that fluctuate over time. Information theory is the study of communication—the ability to reproduce at one point a message specified at another point (Shannon 1948; MacKay 2003). The inevitability of noise or error as a message is transmitted along a channel led to general theory and an entire field of study of how to encode, transmit, and decode messages with the smallest probability of being wrong (MacKay 2003). The conceptual similarity between audio, visual, and other signal processing and ecological time series, where data are extracted from natural systems and information is attributed to sources such as measurement error, external abiotic and biotic influences, and the biological process of interest is increasingly catching the attention of ecologists (Oppenheim et al. 1999; Wiegand et al. 2003).

One information theoretic approach to compare time series uses the mutual information in two or more sequences of data (Cazelles 2004). The information contained in a data series can be estimated by entropy—the certainty of an outcome given the probability that a set of possible events occurs. If we refer to these probabilities for \( n \) possible events as \( p_1, p_2, \ldots, p_n \), entropy is represented using the Shannon index:

\[
H = -\sum_{i=1}^{n} p_i \log p_i
\]  
(3)

We are most certain of an outcome when \( H = 0 \) and least certain at the maximal value of \( H \) for a given \( n \) (\( H_{\text{max}} = \log n \)) when all the probabilities \( p_i \) are equal (\( p_i = 1/n \), for \( i = 1, 2, \ldots, n \)). Given two events, \( x \) and \( y \), with \( m \) possibilities for event \( x \) and \( n \) possibilities for event \( y \), we can calculate the probability of a joint occurrence of possibility \( i \) and \( j \) (\( p(i,j) \)) for event \( x \) and \( y \), respectively (Shannon 1948):

\[
H(x,y) = -\sum_{i,j} p(i,j) \log p(i,j)
\]  
(4)

Cazelles (2004) described the relationship between two ecological data series by discretizing each time series into successions of four possible states—increasing, a peak, a trough, or decreasing with respect to temporal neighbors—and calculating the joint entropy of state value pairs observed at each time point and at various time lags (see also Zunino et al. 2010 for more on time lags). The statistical significance of entropy values was assessed by comparing observed joint entropy to a null distribution of values calculated after resampling the observed time series in a way that preserved important features such as temporal autocorrelation (Cazelles and Stone 2003). The cross-mutual information approach successfully detected a two-year delay between peaks of the North Atlantic Oscillation winter index and of St. Kilda archipelago sheep population abundance, confirming the findings of previous detailed studies in this system.

Estimates of information content can help researchers evaluate the association between time series and whether their signals oscillate similarly regardless of the shape of the function connecting these series. They can also be used to determine whether time series structure responds to some treatment or change. The ‘normalized spectral entropy’ (\( H_s \)) of Zaccarelli et al. (2013) quantifies the predictability of a time series’ power spectrum. Possible states are the relative contribution of a frequency \( \lambda_x \) to the overall power. The measure is normalized to lie between 0 and 1, so spectral densities are more predictable for small values of \( H_s \), where a small number of frequencies dominate the spectrum and are less predictable for large values of \( H_s \), where the spectra has the more uniform density characteristic of white noise. This metric and associated confidence intervals were used to compare climate measures from multiple meteorological stations, to detect changes over time in a series’ spectral density pattern, and for comparison of multiple time series along a spatial gradient (Zaccarelli et al. 2013).

**Convergent cross mapping (CCM)** is a potentially groundbreaking technique developed by Sugihara and colleagues (2012). It utilizes shared information content between time series to address transient or spurious associations that can arise in the absence of causality—the statistical bugaboo that is correlation without causation. CCM tests for causation in three ways: (i) by measuring the extent to which one time series can be used to reliably estimate states of a second time series, (ii)
by testing for convergence in the skill of this reconstruction with increasing time series length, and (iii) by assessing the direction of causality ($X \rightarrow Y$, $X \leftrightarrow Y$, $X \leftarrow Y$) in separate tests. They illustrated that if a causal relationship between two variables exists—for example, the one-directional influence of temperature on population abundance—coordinates of time series variable $Y$ reliably estimate lagged points in $X$ but not vice-versa when using CCM. The accuracy of this estimation, measured by the correlation coefficient $\rho$, increased with the length of the time series used for the reconstruction. CCM also diagnosed causality in three situations that are common in ecological systems and that conflict with prevailing tests of causality: when time series are weakly to moderately coupled, are non-separable (the signal of $X$ embedded in $Y$ is not completely erased when $X$ is removed from the system), and when shared forcing variables cause correlations or apparent synchrony between noninteracting species. A technique well suited for complex and nonlinear systems that eliminates the signal of noncausal correlations is greatly useful to ecologists.

We encourage increased application of CCM to disentangle relationships among time series variables and look forward to the determination of this method’s applicability to a broad array of ecological interaction networks.

Artificial neural networks also model information transmission and may help ecologists understand complex signals embedded in time series. An artificial neuron receives various inputs of information, modifies these according to weights that represent transmission efficiency, and produces a single output (Fyfe 2000; MacKay 2003). Multiple neurons acting in concert create a powerful processing unit. For ecology, artificial neural networks can model relationships between inputs, such as a time series of biotic or abiotic properties, and outputs, such as an individual species’ population density at various time points. The networks are robust enough that numerous forms of information can be included in any model. For example, Thrush et al. (2008) grouped macrobenthic intertidal species into functional groups and used a neural network model to weight the variables that best explained temporal changes in each group’s abundance. The inputs were the other group’s abundance at the same time point, the value of environmental variables at the same time point, and the other group’s abundance at previous time points. The neural network model produced an interaction network among ecosystem components, and network structure was subsequently compared among different collection sites and related to other ecological descriptors such as species diversity and network stability.

Neural network models can estimate the most likely connections among elements given the input data, but an important potential shortcoming for ecologists is the unknown attributes of the ‘hidden’ layers of weights that explain how information passing to the output layer is processed. Weights modeling interactions among neurons and their inputs often remain uninvestigated ‘black boxes.’ The study of Millie et al. (2012) addressed this concern with a ‘Gray-Box’ method that produced the most likely network linking environmental variables as inputs and algal abundance as the output and quantified the influence of specific environmental predictors on the variance and magnitude of algal abundance. Water temperature and salinity had the greatest influence on algal abundance estimates. By varying these two predictors along their observed value ranges and holding the other predictors constant in the neural network model, a response surface linking temperature and salinity to algal abundance was generated and fit to a particular curvilinear mathematical expression. This study is a remarkable example of what can be done with ecological time series — successfully identifying the presence and functional form of significant relationships between environmental variables ($x$) and population dynamics ($y$).

**Section 8. Conclusions**

Understanding how populations and communities respond to their environment is a fundamental objective for ecology and the role of variability has long been acknowledged. Examples where temporal variability or severe disturbance in environmental properties affects individual species as well as the communities and ecosystems they inhabit are numerous (Poff and Allan 1995; Helmuth et al. 2006; Fraterrigo and Rusak 2008; Sabo and Post 2008; Shurin et al. 2010). Time series of both environmental and biological systems are increasingly needed to quantify the form of their relationship and the signature of fluctuating environmental impacts on dynamic ecological systems. This need will be met by continuing expansion and establishment of ecological observing networks (Keller et al. 2008; Duffy et al. 2013), but must be matched with improved methods to process the complex signals embedded in these large datasets.

Our review of past ecological findings and some modern analytical tools is intended to assist ecologists with this challenge. Identifying fundamental properties of environmental and ecological time series, such as variance growth with time or the relationship of population variability to life history traits, allows more accurate models of these series and thus improved predictions of future dynamics. Theoretical and empirical studies up to this point have done a remarkable job isolating the influence of external environment, biotic interactions, and intrinsic properties of species for driving population dynamics. However, these studies also indicate a staggering number of factors interact to leave complex signatures on population variability. The influence of temporally fluctuating environmental properties on population and community dynamics is well informed by past research, but its future will be increasingly driven by the next generation of economatic techniques (Hale and Hollister 2009; Michener and Jones 2012). Ecologists will benefit greatly if methods to take advantage of ‘big data’ are in place to greet accumulating time series datasets.
References


128


Wiegand, T., F. Jeltsch, I. Hanski, and V. Grimm. 2003. Using pattern-oriented modeling for revealing hidden information:


