DETECTING TEMPORAL TRENDS IN SPECIES ASSEMBLAGES
WITH BOOTSTRAPPING PROCEDURES AND HIERARCHICAL MODELS

Nicholas J. Gotelli¹,*

Robert M. Dorazio²

Aaron M. Ellison³

Gary D. Grossman⁴

¹Department of Biology
University of Vermont
Burlington, VT 05405 USA

²Southeast Ecological Science Center
& Department of Statistics
University of Florida
Gainesville, FL 32611 USA

³ Harvard Forest
Harvard University
Petersham, MA 01366 USA

⁴Warnell School of Forestry & Natural Resources
University of Georgia
Athens, GA, USA 30602

*Correspondence author: ngotelli@uvm.edu
ABSTRACT

Quantifying patterns of temporal trends in species assemblages is an important analytical challenge in community ecology. We describe methods of analysis that can be applied to a matrix of counts of individuals that is organized by species (rows) and time-ordered sampling periods (columns). We first developed a bootstrapping procedure to test the null hypothesis of random sampling from a stationary species abundance distribution with temporally varying sampling probabilities. This procedure can be modified to account for undetected species. We next developed a hierarchical model to estimate species-specific trends in abundance while accounting for species-specific probabilities of detection. We analyzed two long-term data sets on stream fishes and grassland insects to demonstrate these methods. For both assemblages, the bootstrap test indicated that temporal trends in abundance were more heterogeneous than expected under the null model. We used the hierarchical model to estimate trends in abundance and identified sets of species in each assemblage that were steadily increasing, decreasing, or remaining constant in abundance over more than a decade of standardized annual surveys. Our methods of analysis are broadly applicable to other ecological data sets, and they represent an advance over most existing procedures, which do not incorporate effects of incomplete sampling and imperfect detection.

KEY WORDS: Temporal trends, species abundance, null model, hierarchical model, stream fishes, grassland insects
Quantifying change in the structure of plant and animal communities is an important challenge for ecology in the 21st century (Walther et al. 2002). Species composition and abundance can respond directly to long-term changes in abiotic factors (Dunson & Travis 1991) and indirectly to changes in the occurrence or abundance of other species (White et al. 2006). Dramatic and rapid changes in community structure may result from the addition or loss of keystone species (Mills et al. 1993), foundation species (Ellison et al. 2005), ecosystem engineers (Jones et al. 1994), and some (but not all) non-native species (Manchester & Bullock 2000). Other changes may be more subtle, because the abundance of individual species can gradually increase or decrease over long periods of time, as in scenarios of a “shifting baseline” (Pauly 1995). Long-term trends may be difficult to detect because of substantial short-term noise and variability in abundances between consecutive samples.

However, not all observed changes in community structure through time are biologically relevant. Most measures of community structure and diversity are sensitive to sampling effort and to the number of individuals counted (Gotelli & Colwell 2001). These quantities are rarely constant through time, even with standardized monitoring programs. Rare species, in particular, are expected to occur more often when sampling is more thorough (Chao et al. 2009). Even the appearance of a previously unrecorded species need not signal a true change in community structure. Biodiversity sampling is labor intensive and notoriously incomplete (Lawton et al. 1998), and “new” species occurrence records—especially of plants and invertebrates—are routinely made, even in sites that have been well-sampled for many years (e.g. Longino et al. 2002).

A variety of univariate and multivariate methods have been proposed to quantify the degree of community change through time (Collins et al. 2000, Fujiwara & Mohr 2009), and to detect temporal trends in community structure (Clarke 1993, Solow 1994). However, with few exceptions (e.g., Dorazio et al. 2010), existing methods do not account for incomplete sampling and imperfect detection. Instead,
most methods assume that the absence of a species from a sampling period represents a “true” zero, and not a detection error (Royle & Dorazio 2008). Most procedures also ignore species that may have been present in a region, but were never detected in any of the samples (Colwell & Coddington 1994).

In this study, we develop new methods for quantifying temporal trends in species abundances that account for errors in detection of individuals. Our methods are appropriate for analyzing species-specific counts of individuals recorded from repeated surveys of a single site. We first develop a bootstrap procedure for testing a null hypothesis in which the counts are assumed to have arisen from sampling a stationary distribution of relative species abundances with temporally varying sampling probabilities. We then develop a hierarchical model of the counts to estimate species-specific trends in abundance while accounting for species-specific probabilities of detection. Both methods of analysis are illustrated for two long-term data sets on stream fishes and grassland insects.

**MATERIALS & METHODS**

**DATA STRUCTURE**

The data for our analyses may be organized as a matrix of counts of individuals of $S$ species (rows) recorded during $T$ successive sampling periods (columns). The matrix entry $y_{ij}$ is the number of individuals of species $i$ that were observed at sampling time $j$ ($i = 1, \ldots, S; j = 1, \ldots, T$). This simple data structure arises in many ecological studies in which species assemblages are repeatedly sampled at a site. Although the samples do not have to be evenly spaced in time, our methods are intended for analysis of long-term trends in abundance, not for short-term or periodic changes in abundance (e.g., seasonality). We illustrate our methods with two data sets: a 13-year record of annual counts of 55 fish species seined from a mid-western U.S.A. stream (Grossman 1982), and a 14-year record of annual counts of 9 insect species collected from sticky traps in a successional grassland plot at the U.S.A. Kellogg Biological Station (KBS 1995). Table 1 summarizes sampling details for each of these
The sampling design and collecting methods for the stream fish study have been described previously (Whitaker 1976; Grossman et al. 1982, 1985) and are only summarized here. A 120 m long × 23 m wide section of Otter Creek, Vigo County, Indiana, USA, was surveyed annually between 1962 and 1974. The site contained a diversity of substrata and depths and can be considered representative of many streams in the midwestern United States. During the study period, the site retained a relatively stable physical structure. Fishes were sampled using a seine, and all collections were supervised by a single investigator, so effort was relatively consistent. There was some minor variation present in sampling efficiency produced by differences in stream depth among years, although the investigators always attempted to keep the area sampled constant. All fishes captured were identified to species and counted, except in a few cases in which a species was extremely abundant. In those cases, numerical estimates were derived from subsamples of the total catch. All fishes were immediately returned live to the site, except for voucher specimens, which were preserved for later identification (Grossman et al. 1982, 1985).

The insect data were collected as part of the Long Term Ecological Research (LTER) network (LTER 2007) at the Kellogg Biological Station (KBS) in northern Michigan, U.S.A. At KBS, a set of 7 1 ha crop rotational treatments have been replicated in 6 blocks on a single 60 ha plot (KBS 1995). We used data from Treatment 7, a native successional treatment that was abandoned after spring plowing in 1989. The plots are surveyed with yellow sticky traps that are replaced weekly from May to August. These traps collect insect predators, many of which are identified only to family (Chrysopidae and Lampyridae). We used data for 9 taxa that were identified to species and were sampled in all years of the study. Within each sampling season, we pooled data from all sticky traps and all plots to generate annual counts for each species.
In this section we describe a null model for detecting temporal trends in species’ abundances. We use the term “null model” to represent a model wherein an $S \times T$ matrix of counts of individuals is assumed to arise by randomly selecting individuals from the species assemblage according to each species’ relative abundance and a set of temporally varying sampling probabilities. The key issue is that no particular ecological process or mechanism is assumed to have generated the matrix of counts; thus the model incorporates simple sampling effects, but is “null” with respect to processes that might induce trends in species abundance (Gotelli & Graves 1996).

The total number of individuals of species $i$ in all sampling periods is:

$$m_i = \sum_{j=1}^{T} y_{ij}$$

(1)

The total number of individuals of all species observed during sampling period $j$ is:

$$n_j = \sum_{i=1}^{S} y_{ij}$$

(2)

Let $N$ equal the total number of individuals summed across all species and samples:

$$N = \sum_{i=1}^{S} \sum_{j=1}^{T} y_{ij} = \sum_{i=1}^{S} m_i = \sum_{j=1}^{T} n_j$$

(3)

We define the relative abundance of species $i$ in the source pool of $N$ individuals as:

$$s_i = \frac{m_i}{N}$$

(4)

Similarly, we define the relative sampling intensity during the $j$th survey as:

$$q_j = \frac{n_j}{N}$$

(5)
Under the assumptions of the null model, $s_i$ is regarded as the probability that an individual drawn from the source pool of $N$ individuals belongs to species $i$, and $q_j$ is regarded as the probability that an individual is observed in the $j$th sampling period, regardless of species.

To conduct a bootstrap test of the null model, we first randomly assign each of the $N$ individuals in the total collection to a particular sample, with probability $q_j$. Once all the individuals are assigned, we then assign them species identities by sampling randomly with replacement from the distribution of $s_i$ values. This two-step process does not depend on the order of conditioning; the same distribution would be obtained by first assigning individuals to species using the $s_i$ values, and then assigning these individuals to particular samples using the $q_j$ values.

This null model describes a multinomial sampling process that is conditional on $N$, the total number of individuals observed. The simulated number of individuals $y_{ij}$ of species $i$ in sample $j$ depends on $s_i$, the proportional representation of species $i$ in the source pool, $q_j$, the proportion of individuals sampled at time $j$, and $N$. The null hypothesis is that variability among species in temporal trends is no greater than would be expected from this simple model of sampling with replacement from an underlying stationary distribution of relative abundance. The alternative hypothesis is that at least some species in the assemblage are systematically increasing or decreasing, leading to changes in relative abundance that cannot be accounted for entirely by sampling effects.

The null hypothesis is defined by the following hierarchical model:

\begin{align}
    n_1, n_2, \ldots, n_T &\sim \text{multinomial}(N, q_1, q_2, \ldots, q_T) \quad \text{(6)} \\
    y_{1j}, y_{2j}, \ldots, y_{5j} &\sim \text{multinomial}(n_j, p_1, p_2, \ldots, p_5) \quad \text{(7)}
\end{align}

and our randomization procedure is entirely consistent with this model. Based on this model, the marginal distribution of the counts is multinomial:
Therefore, under the null hypothesis the expected value of each count is proportional to the product of species relative abundance and year-specific sampling probability.

To estimate temporal trends from the observed data, we first fit a simple linear model to the count data for each species $i$:

$$y_{ij} = \beta_{0i} + \beta_{1i}t_j + \varepsilon_{ij}$$

where $t_j$ is time (in arbitrary units of years, months, or time-steps), $\beta_{0i}$ is the intercept, $\beta_{1i}$ is the slope of the regression for species $i$, and the error term $\varepsilon_{ij}$ has a normal distribution $\left(\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)ight)$.

We are interested in $\beta_{1i}$ because it measures the simple temporal trend in abundance for species $i$.

Temporal change ($TC$) in the entire assemblage can be then be quantified as the sample variance of the estimated $\beta_{1i}$ values:

$$TC = \frac{1}{S-1} \sum_{i=1}^{S} (\bar{\beta}_{1i} - \beta_{1i})^2$$

The larger $TC$, the more heterogeneity there is in the temporal trends of the component species, and the more change in composition of the assemblage that will be seen at future sampling dates. As described below, the number of species generated in the null assemblages was not constant. However, for both the real and the simulated matrices, $TC$ was calculated only for species that were present at least once in the matrix. Following standard procedures for resampling tests (Manly 2009), we generated 1000 null assemblages, and calculated $TC$ for each. We estimate the probability of obtaining $TC$ if the null
hypothesis were true by comparing the observed \( T_C \) to the histogram of simulated \( T_C \) values.

Because the results are potentially sensitive to the assumption of simple linear trends in \( y_j \), we fit two alternative regression models based on log-log and log-linear transformations of \((y_j + 1)\) and \( t_j \). The same transformations were applied to the real and the simulated data. Although these alternative models incorporated non-linear trends in species temporal trajectories, the transformations had no qualitative effect on the outcome of the null model tests. Therefore, we present results only from analyses of the untransformed data fit with a linear trend line.

**Undetected species**

The construction of the null matrix is similar to a simulation of rarefaction (Sanders 1969, Hurlbert 1971), in which a small assemblage is simulated by random draws of subsamples of \( n_j \) individuals from the larger sample of \( N \). However, in rarefaction, sampling is done without replacement (Simberloff 1978). Because our null model treats the source pool as a permanent stationary distribution, we sampled from it with replacement. In practice, the results will not differ unless the sample sizes are so small that \( n_j \) is a relatively large fraction of \( N \), which is not the case for these data sets. Rarefaction also conditions on \( n_j \), the observed count in a particular sample, whereas our multinomial model conditions on \( N \), the total number of individuals.

This procedure implicitly addresses detection error because species (especially rare ones) that are present in the aggregated collection \( N \) may not be represented in any particular sample \( n_j \). In some null assemblages, species that were very rare in the original data set may be missing from all \( n_j \) samples. Because biodiversity sampling is notoriously incomplete, there are also likely to be rare species in the assemblage that were never encountered in the original samples (Colwell and Coddington 1994). We expanded our null model to incorporate these undetected species. We first estimated the minimum number of undetected species, \( \widehat{S} \), using a bias-corrected version of the familiar Chao2 estimator (Chao
1984; Equation (4) in Colwell 2009):

\[ \hat{S} = \left( \frac{T - 1}{T} \right) \left( \frac{Q_1(Q_1 - 1)}{2(Q_2 + 1)} \right) \]  

(11)

where \( Q_1 \) is the number of species represented in exactly 1 time period (“uniques”), \( Q_2 \) is the number of species represented in exactly two time periods (“duplicates”), and \( T \) is the number of samples. The Chao2 index estimates the number of undetected species in the entire assemblage, not the number that may be undetected in any single sample. For the stream fish matrix, the estimated number of undetected species (rounded to the nearest whole integer) was 16. For the insect matrix, sampling was restricted to nine common species, and the estimated number of undetected species was 0.

Once the number of undetected species was estimated, it was necessary to assign them each a relative abundance \( s_i \), so they could be included in the simulation. Estimating these \( s_i \) values would require knowledge of the precise form of the species abundance distribution, a long-standing unsolved problem in ecology (McGill et al. 2007). As a simplifying first approximation, we assumed that \( s_i \) for each undetected species was equal to 0.5* \( s_i \) for the least abundant species observed in the assemblage. The reasoning is that if any of these undetected species occurred at a frequency greater than this, they would likely have been detected at least once in the original sample. For the stream fish data, \( s_i \) for each of the 16 undetected species was set at \( 3.414135 \times 10^{-5} \). Because many of the undetected species are probably much more rare than this, our procedure allows for the greatest possible influence of undetected species. Nevertheless, the results for the stream fish matrix were identical with and without the inclusion of undetected species. However, because the observed number of species is always a biased under-estimator of true species richness, we present the full analyses here with the undetected species included in the null model.

If the observed value of \( TC \) is larger than those of 950 of the 1000 simulated \( TC \) values (\( p < 0.05 \), one-
tailed test), then the temporal trends in the set of observed species are more heterogeneous than can be accounted for by the null model: at least some species are either increasing or decreasing more rapidly than would be expected from sampling effects and undetected species. The null model was programmed and implemented in the statistical language R (R Core Development Team, 2008; Appendix A).

Hierarchical Model of Trend in Abundances

The null model provides a simple test for heterogeneity in species trends. If this null hypothesis is rejected, the next step is to estimate the rate of change in abundance of each species. As before, \( y_{ij} \) is the count of species \( i \) in sample \( j \). We assume that the count \( y_{ij} \) depends on the abundance \( N_{ij} \) present during the \( j \)th survey and on each individual’s probability of capture \( p_{ij} \), as follows:

\[
y_{ij} \mid N_{ij}, p_{ij} \sim \text{binomial}(N_{ij}, p_{ij})
\]  

(12)

To estimate trend, we assume population abundances can be described as:

\[
N_{ij} \mid \lambda_{ij} \sim \text{Poisson}(\lambda_{ij})
\]  

(13)

\[
\lambda_{ij} = \lambda_{i0} \exp(r_i t_j)
\]  

(14)

where \( \lambda_{ij} \) denotes mean abundance of species \( i \) during survey \( j \) and where \( t_j \) denotes the year of the \( j \)th survey. Trend in \( \lambda_{ij} \) values is specified using the familiar exponential growth model (Eq. 14), which includes a species-specific intercept parameter \( \lambda_{i0} \) and a net population growth rate parameter \( r_i \).

Note that \( N_{ij} \) is not actually observed. \( N_{ij} \) is a parameter of the model that represents the number of individuals of species \( i \) which are present and available to be captured during the \( j \)th survey. The observation \( y_{ij} \) can be interpreted as a negatively biased estimator of \( N_{ij} \), with the level of bias depending on the magnitude of \( p_{ij} \), the unknown probability of capture for individuals of species \( i \).

In the absence of replicated observations, we cannot estimate temporal changes in both \( N_{ij} \) and \( p_{ij} \).
Therefore, we assume that capture probabilities vary among species but not among surveys (i.e., we assume $p_{ij} = p_i$). Even with this simplifying assumption, the hierarchical model composed of Eqs. 12–14 contains more parameters than can be estimated from the data. To solve this problem, $N_{ij}$ may be eliminated from the model by integrating the joint distribution of $y_{ij}$ and $N_{ij}$. This integration can be done analytically to obtain the following marginalized version of the hierarchical model:

$$y_{ij} | p_i, \lambda_{i0}, r_i \sim \text{Poisson} \left( p_i \lambda_{i0} \exp(r_i t_j) \right)$$  

(15)

Note that this model may be viewed conceptually as a Poisson regression model. For example, let $\mu_{ij}$ denote the Poisson mean for $y_{ij}$. The logarithm of $\mu_{ij}$ is a linear combination of the marginal model’s parameters:

$$\log(\mu_{ij}) = \log(p_i) + \log(\lambda_{i0}) + r_i t_j$$  

(16)

However, $p_i$ and $\lambda_{i0}$ are not identifiable parameters in Eq. 16 (i.e., both parameters cannot be estimated); therefore, we combine these parameters into a common regression intercept parameter (say, $a_i = \log(p_i \lambda_{i0})$) to obtain

$$\log(\mu_{ij}) = a_i + r_i t_j$$  

(17)

From this equation, the $T$ observations, $y_{i1}, y_{i2}, \ldots, y_{iT}$, can be used to estimate the parameters $a_i$ and $r_i$.

We are interested primarily in the latter parameter $r_i$, which denotes the trend in abundance of species $i$; however, our formulation of the intercept parameter $a_i$ reveals explicitly the combined roles of mean abundance and capture probability in the model.

The model specified by Eqs. 15 and 17 can be fitted to each species separately. However, doing so may produce estimates of trend that are unstable or highly imprecise for species whose abundance appears to be low (as indicated by counts that contain several zeros and ones). Therefore, we extend the model as
follows:

\[ r_i | \beta, \sigma \sim \text{normal} (\beta, \sigma^2) \]  

(18)

where \( \beta \) denotes the average trend in abundance among species in this assemblage and \( \sigma \) denotes the level of variation in \( r_i \) values among species. This distributional assumption allows the counts of all species to be analyzed jointly so that information associated with species of moderate or high abundance can be used to stabilize the estimates of trend for species of low apparent abundance. Nevertheless, even with this assumption, there were not enough data to reliably estimate trends for very rare species that were represented by less than 10 individuals in the entire survey (25 of 55 stream fish species, and 2 of 9 insect species).

Equation 18 implies an exchangeability of trend parameters among species. This exchangeability formalizes the notion that although abundances may be increasing, decreasing, or constant for any particular species, each is also a member of a common assemblage. We expect that the temporal trends of the species in the stream fish assemblage are more similar to one another than they are to, say, the temporal trends of the species in the grassland insect assemblage. A restricted version of this model that corresponds to the null model assumes an identical growth rate \( r_i = \beta \) for all species, so that \( \sigma = 0 \). We can fit this restricted model and compare it to the unrestricted model to assess whether the data support the null hypothesis that all species abundances have an identical trend.

**METHOD OF ESTIMATION**

The hierarchical model described by Eqs. 15, 16, and 18 may be fitted by maximizing the likelihood function obtained by integrating away the latent trend parameters. In our situation, however, this approach is counter-productive. In addition to the minor technical challenges of computing the integrals numerically, the trend parameters \( r_i \) are the quantities of primary scientific interest. Estimates of these
parameters and their uncertainties are actually needed to solve the inference problem. We therefore
adopt a Bayesian approach to inference, which allows every parameter in the model to be estimated
directly, including the species-specific trends in abundance.

In a Bayesian analysis, all inferences are based on the joint posterior distribution of model parameters.
In our case the unnormalized, probability density function (pdf) of this distribution is

\[
\pi(a, r, \beta, \sigma|Y) \propto \pi(\beta, \sigma, a) \prod_{i=1}^{S} g(r_i|\beta, \sigma) \prod_{j=1}^{T} f(y_{ij}|\exp(a_i + r_i \tau_j))
\]  

(19)

where \(a = (a_1, ..., a_n)'\), \(r = (r_1, ..., r_n)'\), and \(Y = (y_1, ..., y_n)'\). Here, \(g(\cdot|\beta, \sigma)\) denotes the
pdf of a normal distribution with mean \(\beta\) and variance \(\sigma^2\), \(f(\cdot|\mu_{ij})\) denotes the probability mass function
of a Poisson distribution with mean \(\mu_{ij}\), and \(\pi(\beta, \sigma, a)\) denotes the pdf of the prior distribution of the
parameters \(\beta\), \(\sigma\), and \(\bar{a}\).

The posterior pdf cannot be written in closed form owing to the analytically intractable integrals in the
normalizing constant (not shown in Eq. 19). Therefore, we estimated the model’s parameters using
Markov chain Monte Carlo algorithms (Robert and Casella, 2004) to obtain an arbitrarily large sample
of the joint posterior distribution. Specifically, we fit the model using the WinBUGS software (Lunn et
al. 2000), which is an implementation of the BUGS language for specifying models and doing Bayesian
analyses (Gilks et al., 1994).

To obtain the posterior sample, we assumed a set of mutually independent noninformative prior
distributions for \(\beta\), \(\sigma\), and \(\bar{a}\). We assumed normal(0, 100^2) priors for \(\beta\) and \(\bar{a}\), and a uniform(0, 10) prior
for \(\sigma\). Each of five Markov chains was independently initialized and computed for a total of 21,000
draws. The first 1000 draws in each chain were discarded as “burn-in”, and every fifth draw in the rest
of each chain was retained to form the posterior sample. Consequently, these calculations yielded a posterior sample of 20,000 draws, which was used to compute estimates of the model’s parameters and their 95% credible intervals (Appendix B).

RESULTS

NULL MODEL ANALYSIS

For the stream fish data, there was a non-significant decreasing trend in total abundance (Figure 1), caused primarily by extremely high abundances in the November 1966 sample ($n_g = 5344$ individuals).

For the null model analysis, this decreasing trend leads to the expectation of negative slopes for individual species, with a moderate amount of variation among species (Figure 2, left panel). However, the observed slopes were much more heterogeneous than this expectation: several species showed sharply increasing or decreasing trend lines (Figure 2, right panel), and the observed $TC$ index was larger than that of all 1000 simulated assemblages (Table 1).

For the insect data, there was a marginally non-significant increasing trend in total abundance (Figure 3), with systematically greater abundances during the final sampling years. For the null assemblages created from this matrix, most species had increasing trend lines (Figure 4, left panel). However, the observed slopes were again much more heterogeneous than expected (Figure 4, right panel). As with the stream fish data, the observed heterogeneity among slopes ($TC$) was greater than that of any of the simulated assemblages (Table 1).

TRENDS IN ABUNDANCES

For the stream fish data, the hierarchical model identified 7 species with significant increases in abundance, 17 species with significant declines in abundance, and 6 species with no significant trend (Figure 5). A negative estimate of average trend, $\hat{\beta} = -0.152$ (95% credible interval: (-0.289, -0.024)), also indicates that species with declining abundances outnumbered those with increasing abundances.
There is little doubt that trends in population abundance differed substantially among species. The posterior distribution of \( \sigma \) (Figure 6, left panel) provides virtually no support for the hypothesis that \( \sigma = 0 \).

For the grassland insect data, the hierarchical model identified 2 species with significant increases in abundance, 3 species with significant declines in abundance, and 2 species with no significant trend (Figure 7). The estimate of average trend, \( \hat{\beta} = -0.026 \) (95\% credible interval: (-0.352, 0.297)), reflects the nearly equal numbers of species with increasing and decreasing abundances. As with the stream fish data, the posterior distribution of \( \sigma \) (Figure 6, right panel) does not support the null hypothesis (\( \sigma = 0 \)) of identical trend lines for these 7 species.

**DISCUSSION**

The null model and the hierarchical model provide complementary information on temporal trends, and they both point to strong temporal re-organization of stream fish and insect grassland assemblages over periods greater than a decade. Because the insects were sampled in a successional plot, it is no surprise that strong temporal trends would be detected, as vegetation structure and arthropod prey assemblages were systematically changing through time. In fact, the two most rapidly increasing species, *Harmania axiridis* and *Hippodamia glacialis*, never appeared in any of the traps until 6 and 7 years respectively after the sampling began. This is exactly the pattern that would be expected in a classic facilitation model of succession (Connell & Slatyer 1977). On the other hand, the abundance of the most common species in the samples, *Coccinella septempunctata* (\( \bar{x} = 200 \) individuals/year), did not change significantly during the 14-year successional sequence (Figure 7).

The pattern for the stream fishes is more complex. Although no obvious physical changes were
observed in the habitat during the 15-year sampling period, 17 species showed significant declines, whereas only 7 species increased in abundance. The causal mechanisms behind these patterns are unclear because both generalist and specialist species were found in all categories, as were representatives of most North American taxonomic groups. Perhaps the preponderance of declining populations suggests that some species successfully invaded the site early in the time series, but were not able to sustain populations through local reproduction and began to decline. It is likely that flow variation plays some role in these trends, perhaps facilitating establishment of species in benign periods and causing substantial mortality during periods of high water (Grossman et al. 1982; 1998). High flow events may cause substantial mortality in stream fishes, especially if they occur during the reproductive period and destroy an entire year-class (Grossman et al. 1982; 1998). However, there was no evidence during the sampling period of declining flows or increased numbers of extreme flow events that might be linked to the decreasing abundance of 17 of the 55 species (Grossman & Sabo 2010). The decreasing trends in abundance of many stream fish species (Figure 5) are consistent with a shifting baseline scenario, but the causes of these declines are still unknown.

The results of both the null model and the hierarchical model are potentially sensitive to the functional form that is used to describe temporal trends. For the null model analysis, the results for these data sets were the same when the trends were fit with linear, semi-logarithmic, or log-log transformations of the original data. The estimated heterogeneity among species in temporal trends does not seem to be sensitive to the fitting procedure, perhaps because deviations caused by extreme sample numbers (such as the high counts in the stream fish data set in 1966) are also incorporated into the pattern in the null assemblages. Both the null model and the hierarchical model assume that species are independent of one another. However, it is unclear how the violation of this assumption (from strong species interactions) would systematically affect the estimates of temporal trends in abundance.
Because the hierarchical model is being used for parameter estimates of change (rather than just a simple dichotomous null model test), it is potentially more sensitive to violation of its assumptions. As we noted, one important assumption in this model is that capture probabilities are constant through time. Although this assumption may not be true, it probably matches the perspective of most field biologists, who typically assume that long-term monotonic changes in species counts with standardized sampling methods primarily reflect changes in abundance, rather than changes in detection or capture probabilities.

If species-specific capture probabilities are not constant, the magnitude of the deviations between observed and expected counts may be inflated. As long as these deviations do not vary systematically with time, the point estimates of trend will not be affected, although the credible intervals may be too narrow. Alternatively, if the deviations between expected and observed counts vary systematically with time, say changing from positive to negative values, the trend estimates will be very sensitive to an incorrect assumption of constant capture probability. For the data sets we analyzed, there was no evidence of a systematic lack of fit (Figure 8).

In the hierarchical model, the assumption of constant sampling probabilities was necessary only because of the extremely simple and unreplicated structure of the data matrix. With replication, it may be possible to estimate parameters for temporal trends in both abundance and detection probabilities. For example, the KBS insect data actually consist of weekly sticky trap counts collected from 6 replicated plots. Rather than pooling the data as we have done in this analysis, the individual trap records could be fit to a more complex hierarchical model (Royle & Dorazio 2008, Kery et al. 2009). The hierarchical model could also be expanded to incorporate species-specific covariates \( Z \) (such as body size, geographic range size, or degree of habitat specialization) that might be of interest for conservation purposes. Species-specific covariates could be used to model either the mean structure of
the elements of $r$ in Eq. 18 or their covariances.

Both the bootstrap test and the hierarchical model assume that changes in abundance through time are monotonic. If species show more complex patterns of temporal change (such as periodic fluctuations), these could be accommodated by fitting polynomial or sine functions to the time series. However, at least for these data sets, diagnostic analysis of residuals indicated little evidence for departures from linearity over the time periods that were sampled. Moreover, a monotonic function is appropriate for very short data series such as these ($T = 15$ samples for stream fishes and $T = 14$ samples for grassland insects)

Finally, the frequent occurrence of rare species in natural assemblages continues to pose statistical challenges. In the null model, all species, no matter how rare, are included in the analysis, and the null assemblages even incorporate the possibility of species that were never detected in any of the samples. In theory, rare species can contribute to the size of the observed $TC$ index. For example, if all of the rare species occurred in only the very first or the very last sample, the sample variance in the trend lines would tend to be large compared to that found for the null assemblages. However, less extreme distributions of rare species look very similar to those generated by the null model, and therefore would not contribute substantially to the $TC$ index.

In the hierarchical model, the assumption of exchangeability of $r_i$ values allowed us to use information from common species to estimate trends for less common species. Nevertheless, when abundance is so low that there are fewer than 10 individuals counted in 14 or more consecutive annual samples, estimating temporal trends with any statistical model is a dubious exercise. For these cases, auxiliary information, stratified sampling, and additional data may be necessary (Dixon et al. 2005).
In summary, quantifying temporal trends in species abundances is an important forecasting problem. Given the accelerating rates of habitat alteration and global climate change, the strong heterogeneity that we detected in the stream fish and grassland insect data sets (Figures 5 and 7) may be typical; it seems unlikely to us that most long-term temporal trends will be accounted for entirely by the sampling effects in our null model. In these cases, the hierarchical models provide a sensible framework for predicting what the future may hold.
ACKNOWLEDGMENTS

We thank Anne Chao and 2 anonymous reviewers for comments that improved the manuscript. NJG was supported by the U.S. National Sciences Foundation (NSF DEB-0541936) and the U.S. Department of Energy (022821). GDG was supported by the Warnell School of Forestry and Natural Resources and U.S. Department of Agriculture (USDA) Forest Service McIntire-Stennis program grant GEO-00144-MS. Portions of this work represent a contribution from the Harvard Forest Long-Term Ecological Research Site, supported by NSF 06-20443. This work was also facilitated by meetings at the Binary Matrices Working Group at the National Institute for Mathematical and Biological Synthesis, sponsored by NSF, the U.S. Department of Homeland Security, and the USDA through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville.
Appendix A. R-script file for null model analysis.

# This program tests for heterogeneity in linear trend lines
# of species abundances.

# GLOBAL VARIABLES

ObservedMatrix <- as.matrix(read.csv("Grossman Streamfish Data.csv", header=TRUE, row.names=1))
# input has species as rows, columns as times, with labels in the first row and column

Time <- c(11, 21, 34, 35, 45, 47, 57, 59, 71, 81, 82, 93, 117, 129, 141, 153)
# these are months since start of 1962 in Grossman streamfish datafile

NRep <- 1000
# this is the number of random replicates to use in the null model analysis

SimVar <- numeric(NRep)
# this is the vector of simulated variances in growth rates

ObsVar <- 0
# this is the single observed variance in growth rates for comparison with null

# FUNCTION SourcePool
# This function takes an input matrix and a calculated number of missing species
# It returns a probability vector for creating null matrices
# For the missing species we assume their frequency = 0.5*rarest observed species

SourcePool <- function(FMatrix, FMissingSpecies) {
    SourcePool <- rowSums(FMatrix)
    SourcePool <- SourcePool/sum(SourcePool)
    SourcePool <- c(SourcePool, rep(0.5*min(SourcePool), FMissingSpecies))
    SourcePool <- SourcePool/sum(SourcePool)
}

# FUNCTION RandomSampler
# This function takes as input the SourcePool vector and the original data matrix
# It returns a single null matrix created by sampling individuals one at a time
RandomSampler <- function(FSourcePool, FMatrix) {

TimeProbs <- colSums(FMatrix)/sum(FMatrix)
PoissonN <- numeric(ncol(FMatrix))

for (j in 1:sum(FMatrix)) {
  i <- sample(seq(1:length(PoissonN)),1,replace = TRUE,TimeProbs)
  PoissonN[i] <- PoissonN[i] + 1
}

SampleMatrix <- matrix(0,length(FSourcePool),ncol(FMatrix))

for (j in 1:ncol(FMatrix)) {
  for (n in 1:PoissonN[j]) {
    i <- sample(seq(1:length(FSourcePool)),1,replace = TRUE,FSourcePool)
    SampleMatrix[i,j] <- SampleMatrix[i,j] + 1
  }
}

RandomSampler <- SampleMatrix

# FUNCTION GrowthVar
# This function takes as input a data matrix and a vector of times
# It fits a linear model to temporal trends for each species and
# calculates the variance among species in the measured slopes
# The calculation is only made for species that occurred at least once in the matrix
# This function can be modified to fit other kinds of temporal models
GrowthVar <- function(FMatrix, FTime) {

Growth <- NULL #creates empty vector of unspecified length

for (i in 1:nrow(FMatrix)) {
  if (sum(FMatrix[i,]) > 0) {
    model <- lm(FMatrix[i,]~FTime) #fits model for species that are present
  }
}

var(Growth, use = "complete") #species with 0 generate missing values that are ignored

#########################################
#  FUNCTION Chao2
#  This function takes as input the observed matrix
#  It calculated the bias corrected version of Chao2 (Equation 4 in EstimateS manual)
#  Note that is this function is needed only once to get missing species for the simulation
#  Results are rounded to the nearest whole integer

Chao2 <- function(FMatrix) {
  Occurrences <- NULL
  for (i in 1:nrow(FMatrix)) {
    Occurrences[i] <- sum(FMatrix[i,] > 0)
  }
  Uniques <- sum(Occurrences == 1)
  Duplicates <- sum(Occurrences == 2)
  m <- ncol(FMatrix) # m = number of samples

  Chao2 <- ((m - 1)/m)*((Uniques*(Uniques - 1))/(2*(Duplicates + 1)))
  Chao2 <- round(Chao2)
}
MissingSpecies <- Chao2(ObservedMatrix)
SourcePoolVector <- SourcePool(ObservedMatrix, MissingSpecies)
ObsVar <- GrowthVar(ObservedMatrix, Time)
for (i in 1:NRep) {
  RandomMatrix <- RandomSampler(SourcePoolVector, ObservedMatrix)
  SimVar[i] <- GrowthVar(RandomMatrix, Time)
}
ConfidenceInterval <- quantile(SimVar, c(0.025, 0.975))
ConfidenceInterval
ObsVar
SES <- (ObsVar - mean(SimVar))/sqrt(var(SimVar))
SES
sum(ObsVar > SimVar)/NRep
summary(SimVar)
# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop

# PROGRAM Basic Simulation Loop
Appendix B. WinBUGS script file for simulation of prior distributions and fit of marginal trend model.

model {
    # priors
    beta ~ dnorm(0, 0.0001)
    sigma ~ dunif(0, 10)
    tau <- pow(sigma, -2)
    for (i in 1:S) {
        a[i] ~ dnorm(0, 0.0001)
    }

    # marginal model of trend
    for (i in 1:S) {
        r[i] ~ dnorm(beta, tau)
        for (j in 1:T) {
            y[i,j] ~ dpois(mu[i,j])
            log(mu[i,j]) <- a[i] + r[i]*t[j]
        }
    }


Grossman, G. D. & Sabo, J. L. 2010 Incorporating environmental variation into models of community


Kellogg Biological Station Sampling Protocols. 1995 http://lter.kbs.msu.edu/protocols/36


McGill, B. J. et al. 2007. Species abundance distributions: moving beyond single prediction theories to


invasion biology. *Diversity and Distributions* 12, 443-455.

Table 1. Empirical details of two data matrices and null model results for temporal change analysis. See text for definitions of variables.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxon</td>
<td>Stream fishes</td>
<td>Grassland insects</td>
</tr>
<tr>
<td>Study site</td>
<td>Central Illinois stream reach</td>
<td>Successional grassland plot</td>
</tr>
<tr>
<td>Sampling method</td>
<td>Seining</td>
<td>Sticky Traps</td>
</tr>
<tr>
<td>Observed number of species</td>
<td>55</td>
<td>9</td>
</tr>
<tr>
<td>Estimated number of undetected species (Chao 1984)</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Sampling interval</td>
<td>~ annual</td>
<td>annual</td>
</tr>
<tr>
<td>Number of sampling dates ($T$)</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Average abundance per sample (minimum, maximum)</td>
<td>914 (87, 5344)</td>
<td>401 (71, 1152)</td>
</tr>
<tr>
<td>Average abundance per species (minimum, maximum)</td>
<td>266 (1,4304)</td>
<td>624 (2, 2793)</td>
</tr>
<tr>
<td>Total abundance ($N$)</td>
<td>14142</td>
<td>5614</td>
</tr>
<tr>
<td>Observed Temporal Change index ($TC$)</td>
<td>0.256</td>
<td>156.90</td>
</tr>
<tr>
<td>Average of 1000 simulated values of $TC$ (95% confidence interval)</td>
<td>0.095 (0.082, 0.111)</td>
<td>39.16 (31.81, 47.22)</td>
</tr>
<tr>
<td>P(observed $T_C$</td>
<td>null model)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
Figure 1. Temporal trends in total abundance for the stream fish samples of Grossman et al. (1982). The dashed line indicates the regression line for a simple linear model ($n_t = 1371 - 6.212 t; r^2 = 0.04; P = 0.446$)
Figure 2. Observed and simulated trend lines for the stream fish data of Grossman et al. (1982). Left panel: results of a single replicate of the null model simulation, in which total abundances for each species are sampled randomly from the abundance distribution of all species pooled through time. See text for details of the simulation model. Each line is the least-squares regression for one of the simulated species. Right panel: same graph for the observed data.
Figure 3. Temporal trends in total abundance for the insect samples of the successional plot. The dashed line indicates the regression line for a simple linear model ($n_t = -68253.0 + 34.4t; r^2 = 0.27; P = 0.057$).
Figure 4. Observed and simulated trend lines for the insect data from the successional plot. Details as in Figure 2.
Figure 5. Hierarchical model estimates of $r_i$, the intrinsic rate of increase ($= \ln(\lambda)$) for 30 species of stream fishes. Each circle represents the estimated $r_i$ and the error bar is the asymmetrical 95% credible interval.
Figure 6. Posterior distribution of $\sigma$, the variation among species in temporal trends, estimated from the hierarchical model (Eq. 19) by Markov chain Monte Carlo simulation. Left panel: stream fishes. Right panel: grassland insects.
Figure 7. Hierarchical model estimates of $r$, the intrinsic rate of increase ($= \ln(\lambda)$) for 7 species of grassland insects. Details as in Figure 5.
Figure 8. Estimated trends for representative stream fish and insect species. Upper panel: Estimated trends in captures of stream fish (shown by lines) of two species superimposed on the observed counts of these species. Ordinate is logarithmically scaled. Three zero-valued counts (for years 1964, 1966, and 1968) of Noturus miurus are not shown. Lower panel: Trends for insect species. Five zero-valued counts (for years 1989-1993) of Harmonia axyridis are not shown.