Oecologia (1993) 93:435–444

Oecologia

New insights into testing for density dependence

M. Holyoak

Dept. of Biology, Imperial College at Silwood Park, Ascot, Berks SL5 7PY, UK

Received: 20 July 1992 / Accepted: 23 November 1992

Abstract. The reasons why tests for density dependence often differ in their results for a particular time-series were investigated using modelled time-series of 20 generations in length. The test of Pollard et al. (1987) is the most reliable; it had the greatest power with the three forms of density dependent data investigated (mean detection rates of 50.8–61.1%) and was least influenced by the form of the density dependence in time-series. Bulmer's first test (Bulmer 1975) had slightly lower power (mean detection rates of 27.4–56.8%) and was more affected by the form of density dependence present in the data. The mean power of the other tests was lower and detection rates were more variable. Rates were 24.6–46.2% for regression of *k*-value on abundance, 6.4–32.6% for regression of *k*-value on logarithmic abundance and 0.2–13.7% for Bulmer's second test (Bulmer 1975). Bulmer's second test is not useful because of low power. For one method, regression of *k*-value on abundance, density dependence was detected in 19.9% of time-series generated using a random-walk model. For regression of *k*-value on logarithmically-transformed abundance the equivalent figure was 18.3% of series. These rates of spurious detection were significantly ($P<0.001$) greater than the generally accepted 5% level of type I errors and so these methods are not suitable for the analysis of time-series data for density dependence. Levels of spurious detection (from random-walk data) were around the 5% level and hence were acceptable for Bulmer's first test, Bulmer's second test, and the tests of Pollard et al. (1987), Reddingius and den Boer (1989) and Crowley (1992). For all tests, except Bulmer's second test, the rate of detection and the amount of autocorrelation in time-series were negatively correlated. The degree of autocorrelation accounted for as much as 59.5–77.9% of the deviance in logit proportion detection for regression of *k*-value on abundance, Bulmer's first test, and the tests of Pollard et al. and Reddingius and den Boer. For regression of *k*-value on abundance this relationship accounted for less of the deviance (29.4%). Independent effects of density dependence were largely absent. It is concluded that these are tests of autocorrelation, not density dependence (or limitation). Autocorrelation was found to become positive (which is similar to values from random-walk data) as the intrinsic growth rate became either small or large. As the strength of density dependence (in the discrete exponential logistic equation) is dependent on the product of the intrinsic growth rate and the density dependent parameter α it is unclear whether this is because of variation in the strength of density dependent mortality or reproduction *per se*. However, small values of the intrinsic growth rate cause the amount of variation in the data to become small, which might hinder detection of density dependence, and large values of the intrinsic growth rate are coincident with deterministic chaos which hinders detection. The user of these tests for density dependence should be aware of their potential weakness when variation within time-series is small (which itself is difficult to judge) or if the intrinsic growth rate is large so that chaotic dynamics might result. Power and levels of variability in rates of detection using Reddingius and den Boer's test were intermediate between those of the test of Pollard et al. and Bulmer's first test. This, combined with the strong relationship between rates of detection of limitation and the value of the autocorrelation coefficient, make testing for limitation similar to testing for density dependence. Crowley's test of attraction gave the widest range of mean detection rates from density dependent data of all the tests (20.4–60.6%). The relative rates of detection for the three forms of density dependent data were opposite to those found for Bulmer's first test and the test of Pollard et al. I conclude that testing for attraction is a complementary concept to testing for density dependence. As dynamics represented in time-series generated using a stochastic form of the exponential logistic equation became chaotic, Bulmer's first test, the test of Pollard et al. and regression of *k* on abundance failed to detect density dependence reliably. Conversely, Crowley's test was capable of detecting attraction with a power between 96 and 100% with time-series containing both stochastically and deterministically chaotic dynamics. This difference from other tests is in agreement with the lower influence of autocorrelation.

Key words: Attraction – Autocorrelation – Chaos – Limitation – Power

---

Tests for density dependence using data that consist of abundances measured over time frequently fail to detect density dependence in natural populations. This is best illustrated by attempts to uniformly analyse data from groups of species for the presence of density dependence (Vickery and Nudds 1984, 1991; Gaston and Lawton 1987; Crowley 1992; Crowley and Johnson 1992; Woiwod and Hanski 1992; Holyoak and Lawton 1992) and delayed density dependence (Turchin 1990). These studies have enabled those tests which frequently detect density dependence to be distinguished from those that rarely identify it. Methods that consistently failed to detect density dependence (in > 80% of time-series) include double regression (test: Varley and Gradwell 1960; used by Vickery and Nudds 1984), principal and standard major axes (tests: Slade 1976, 1977; used by Vickery and Nudds 1984, 1991; Gaston and Lawton 1987) and Bulmer's second test (test: Bulmer 1975; used by Vickery and Nudds 1984; Holyoak and Lawton 1992).

The strong dependence of the rate of detecting density dependence on the length of time-series (Hassell et al. 1989; Woiwod and Hanski 1992) is best illustrated by the analysis of 5715 time-series from aphids and moths by Woiwod and Hanski (1992). Including time-series of all lengths led to rates of detection of 47%, 56% and 29% in moths for Bulmer's (1975) first test, regression of k-value against abundance (Varley and Gradwell 1960) and the test of Pollard et al. (1987) respectively. If time-series of less than 21 generations were excluded from the analysis rates of detection were raised to 80%, 79% and 57% respectively for the three methods. Analysis of data from aphids led to rates of detection of 81%, 77% and 57% for all lengths of time-series and 87%, 88% and 84% for only the longer (> 20 generations) time-series. Simulations have shown that the presence of temporal trends can cause tests to fail to detect density dependence (Slade 1977; Vickery and Nudds 1984; Pollard et al. 1987); however, this was not illustrated with field data until the work of Woiwod and Hanski (1992).

There have been several attempts to review the results of the application of tests for density dependence to time-series data. These studies differ from those already discussed because the reviewers have accepted the findings of the original authors, so that the data are not uniformly analysed. These studies have found density dependence from 48–67% of time-series (Dempster 1983; Stiling 1987, 1988; Hassell et al. 1989). They are biased in the species that they consider; many species are of pest status, and have outbreak dynamics (Latto 1989). The majority of time-series considered in these studies were short, less than 20 generations in length, with over half being of fewer than 10 generations.

Despite the high rates of detection of density dependence with long time-series (over 20 generations) that lack temporal trends (Woiwod and Hanski 1992), tests of density dependence often differ in the results for a particular time-series (e.g. Holyoak and Lawton 1992), making their interpretation questionable. This is particularly the case with time-series of less than 20 generations. This study aims to discover why tests for density dependence differ in their ability to detect density dependence from a given time-series and hence which test results are most reliable. Simulated time-series are used to demonstrate how frequently the various methods correctly or falsely detect density dependence. The density dependent models chosen were chosen for their relevance to insect population dynamics (Bellows 1981). The following were investigated:

1. Levels of spurious detection of density dependence with random-walk data and the power (rates of detection) of tests with density dependent data
2. The specificity of tests to the form of density dependence present; where the *form* of density dependence is the nature of the link between mortality (or fecundity) and abundance. The power and specificity of tests of density dependence were compared using time-series from three different forms of density dependent model. This was carried out with the aim of showing which tests had the greatest power with a range of forms of density dependence and hence should be most appropriate for analysing field data.
3. The nature of the link between autocorrelation and the presence of density dependence was explored using four different models to generate both density dependent and density independent data. The influence of autocorrelation on tests for density dependence was identified by Maelzer (1970) and Itô (1972); however, it has not been explored in any detail. In this analysis I use density-independent random-walk data and three forms of density dependent data as sources of simulated time-series with differing levels autocorrelation and density dependence.
4. Statistics which bias detection of density dependence by tests were searched for by looking for correlations between rates of detection of density dependence and the mean, variance, variance-mean ratio, range, skewness and kurtosis of logarithmically-transformed abundances. Additionally, patterns in detection rates of density dependence and the presence of trends were searched for.
5. The effect of the presence of deterministic chaos (see May 1974 for a definition) on levels of detection of density dependence was investigated. Recently Mountford and Rothery (1992) showed that Bulmer's first test can detect density dependence when the dynamics are chaotic; this claim was investigated. In the logistic models of May (1974) the amount of chaos is directly related to the intrinsic growth rate and the amount of chaos is thought to influence detection of density dependence (May 1989). Hence the exponential logistic model was used to generate time-series for analysis.

Methods

Types of data analysed

Time-series simulated from four different forms of model were analysed for density dependence (see below). All data had a range of logarithmic variance-mean ratios that encompass the range of all natural variation (ranging from less than 0.1 to over 3). For each parameter combination of each model, 25 time-series of 20 generations in length were analysed for density dependence. A wide range of parameter values were chosen to prevent placing too much emphasis on individual combinations. In preliminary work the same analyses were carried out on time-series of 10 and 15 generations. The results of these were qualitatively similar, although the power of all tests was lower with shorter runs of data. All simulations were carried out in Turbo Pascal (version 5.5, Borland International Inc.), using the built-in random number generator as a source of evenly distributed deviates. To avoid problems of numerical overflow and underflow logarithmic values of abundances were analysed for density dependence. The mathematical properties of the models for density dependent populations are discussed by Bellows (1981) and the model authors (see below). In all of the density dependent models a density dependent parameter was made into a random variable. This was done in order to produce time-series that were more qualitatively similar to series from field populations of insects than can be achieved with deterministic versions of the models used. Alternatively, random variation could have been introduced by an additional parameter or through the intrinsic (or finite) growth rate. Preliminary investigations showed that if the amount of variation in the resulting time-series is kept constant then rates of detection of density dependence by the various tests were fairly similar for models with different types of stochastic variation. Hence, the choice of a random density-dependent parameter (as opposed to a separate stochastic parameter or a stochastic intrinsic growth rate) is apparently unimportant for rates of detection of density dependence (and limitation/attraction). However, we do not know what is the most appropriate way to introduce variation into these models to make them as realistic as possible. The data were simulated using the following methods:

1. Random-walk data. A logarithmic random-walk was used to generate density independent time-series. The random-walk used was of the form:

$$X_{t+1} = X_t + z_t$$

where X_t is the natural logarithm of abundance at time t and z_t values were independent normal deviates with zero mean and a standard deviation of s. Values of X_1 were arbitrarily set to 5, 10 or 15 and s was set to 0.2, 0.4, 0.6, 1.0, 1.4, 1.8, 2.2, 2.6, 3.0, 3.4, 3.8, 4.2 for each of these starting values. The resulting time-series of logarithmic abundances had variances in the range 0.4–28.6 and mean values between 4.6 and 16.1.

2. Density dependent data generated using the exponential logistic model. A modified form of the exponential model of May (1974) was used:

$$N_{t+1} = N_t \exp[r(1 - \alpha_t N_t)]$$

where N_t is the abundance at time t, r is the intrinsic growth rate and α_t is the inverse of the carrying capacity. The density dependent parameter α_t was made into a normally distributed random variate. To avoid introducing temporal trends into the time-series the initial abundance, N_1, was set close to the carrying capacity; actual values of the logarithm of initial abundance were generated as evenly distributed deviates with a mean of the carrying capacity and a standard deviation of 0.01. Mean values of α_t were set at 0.001, 0.01, and 0.1, with the standard deviation of α_t set at 0, 1×10^{-10}, 5×10^{-5} and 1×10^{-4} for each of these strengths of density dependence. For all of these combinations, r values of 0.5, 1.0, 1.5, 1.7, 2.0, 2.5, 2.7, 2.9, 3.1, 3.3, 3.5 and 4.0 were used. Means value of logarithmic abundances were between −9.2 and 18.7 (however, very few abundances were negative) and variances of logarithmic abundance were in the range 0.4–21.7.

3. Density dependent data generated using the multiplicative logistic model. The multiplicative model of May (1974) was used in a modified form:

$$N_{t+1} = N_t[1 + r(1 - \alpha_t N_t)].$$

As in the previous model, α_t was made into a normally distributed random variate. Mean values of α_t were set at 0.001, 0.01 and 0.1, and the standard deviation of α_t was set at 0, 0.00005 and 0.0007 for each of these values. For each of these combinations 12 values of r were used (as in Model 2). Mean values of logarithmic abundance were in the range 1.23–16.2 and variances of logarithmic abundance were between 0.7 and 24.3.

4. Power-form density dependent data were generated using the model of Maynard Smith and Slatkin (1973). This model uses two parameters to give greater control of density dependent mortality; the parameter a defines the abundance at which a given proportionate mortality occurs, while b defines the strength of density dependence, which acts on abundances in a power relationship:

$$N_{t+1} = \lambda N_t (1 + (aN_t)^b)^{-1}$$

where λ is the finite growth rate. For values of b close to unity, abundances rise monotonically to an equilibrium abundance, giving contest-like competition. When $b >> 1$ scramble competition is implied and the k-value (a logarithmic measure of proportionate mortality) rises exponentially with abundance. To explore the ability of the tests to detect density dependence when the process was not sharply (deterministically) defined the parameter b was made into a normally distributed random variate with standard deviation of 0, 0.2 and 1.0. Values of λ were 1, 2, 3.5, 5, 15 and 30, mean b-values were set to 1, 5 or 10, and a was set to 0.001, 0.01 and 0.05 for each value of b. Mean values of logarithmic abundance were in the range −2.2 to 17.9 and variances of logarithmic abundance were between 0.9 and 19.4.

Tests of density dependence used

Five different tests of density dependence were carried out on each data set. Further analyses tested for limitation (*sensu* Reddingius and den Boer 1989), attraction (*sensu* Crowley 1992), significant temporal trends and finally the amount of serial autocorrelation. The tests were selected either because they are widely used (Hassell et al. 1989) or because they are relatively new and have not been investigated.

1. Least-squares regression of k-value Ln (N_t/N_{t-1}) against Ln(N_{t-1}) (Varley and Gradwell 1960). F-tests were used to judge significance $P \leq 0.05$, the slope was then tested against a slope of zero using a Student's t-test. Errors in the ordinate were assumed to be normally distributed. Both this test and Test 2 are used despite criticisms of them dating from as early as 1970 (Maelzer 1970; St Amant 1970); this is because they have continued to be widely used (see Hassell et al. 1989 for a review of their usage). These tests are used here solely to provide further evidence of their inappropriateness for analysing time-series data. The degree of bias in regression tests of density dependence was neatly shown by Vickery (1991); however, the power of these tests relative to other techniques has not been previously investigated. This subject is developed further in *How useful are tests?* in the Discussion section.

2. Least-squares regression of k-value against N_{t-1} (Varley and Gradwell 1960). Apart from the choice of abscissa, the method was identical to the first test.

3. Bulmer's first test (Bulmer 1975) was carried out using the following formulae:

$R = V/U$

where $V = \sum_{t=1}^{n} (X_t - \bar{X})^2$

and $U = \sum_{t=1}^{n-1} (X_{t+1} - X_t)^2$.

Critical values of R were calculated using $R_{0.05} = 0.2500 + (n-2)0.0366$ (Bulmer 1975); the null hypothesis of density independence was rejected for R values $\leq R_{0.05}$.

4. *Bulmer's second test* Bulmer (1975) was investigated, despite criticisms that the test is conservative (Slade 1977; Vickery and Nudds 1984; Reddingius and den Boer 1989). The following formulae were used:

$R^* = W/V$

where $W = \sum_{t=1}^{n-2} [(X_{t+2} - X_{t+1})(X_t - \bar{X})]$ and $V = \sum_{t=1}^{n} (X_t - \bar{X})^2$.

Critical values of R^* were calculated using $R^*_{0.05} = (-13.7/n) + (139/n^2) - (613/n^3)$ (Bulmer 1975). The null hypothesis of density independence was rejected for $R^* \leq R^*_{0.05}$.

5. *The test of Pollard et al. (1987)* was carried out in the following way:

a. The test statistic, the correlation coefficient between observed k value and the natural logarithm of population size, was calculated for the observed census data.
b. $\chi_1 = X_1$ was calculated, where $\chi_i[1 < i < (n-1)]$ is used to denote the natural logarithm of abundance in a randomized data set at time i.
c. $d_t = X_{t+1} - X_t$ was calculated. These values were then randomly shuffled using the random number generator in GLIM (Nelder and Wedderburn 1972) to generate uniformly distributed variates which were used as positions in the series, such that the new positions in the sequence were called i for the d_i values. Following Manly (1991) a total of 25 000 simulated time series were used.
d. Randomized data sets were constructed using $\chi_{i+1} = \chi_i + d_i$ and the test statistic was calculated for each randomized data set.
e. The proportion of values of the test statistic which were larger or equal to the observed value was calculated. This proportion is the conditional probability of density independence.

This test compares the time-series under consideration with density dependent time-series that can have a temporal trend if such a trend exists in the time-series. This is in contrast with the other tests, which make no allowances for trends.

6. *The test of Reddingius and den Boer (1989)* tests for limitation (*sensu* Reddingius and den Boer 1989) rather than density dependence. Despite this difference, it is remarkably similar to the test of Pollard et al. (1987), the only difference being the use of the logarithmic range $(X_{max} - X_{min})$, as the test statistic and not the correlation coefficient as chosen by Pollard et al. Unlike that of Pollard et al. the test is probably sensitive to temporal trends (Reddingius and den Boer 1989).

7. *Crowley's test* of attraction (Crowley 1992) was carried out using the following sequence of calculations:

a. The n observed abundances are sorted into ascending sequence.
b. These abundances were used as $n+1$ density intervals. The $n-1$ density increments between density observations were each calculated as $\ln(N_{t+1}+1) - \ln(N_t+1)$ for growth between generations t and $t+1$.
c. The test statistic was the *minimum violation number*. Each of the $n+1$ density intervals each represent trial attractors. The violation number is obtained by counting the number of observed abundances in the time-series that were outside the range of abundances encompassed by the attractor for each of the $n+1$ density intervals. The trial attractor that gave the smallest violation number is then the minimum violation number.
d. The randomization procedure and method of calculation of the conditional probability are identical to those used by Pollard et al. (1987) (Test 5, above) and Reddingius and den Boer (1989) (Test 6, above).

8. *Temporal trends* in abundances were assessed; the presence of a trend was determined for each of the time-series by regressing the natural logarithm of abundance against time. The significance of the regression was judged using an F-test. The use of a regression test to quantify temporal trends might be questioned because successive values of the dependent variable are not independent. The test is used here because of its simplicity and so that results are comparable to those of Woiwod and Hanski (1992) who carried out an analysis of 5715 time-series from insects and moths. The potential inadequacy of this regression is a minor problem because the influence of autocorrelation was also assessed (test 9) and autocorrelation and temporal trends cannot normally be separated (Diggle 1990). This topic is discussed further in *What influences rates of detection?* in the Discussion.

9. *Autocorrelation* was quantified using the one-step autocorrelation coefficient (ACF). The coefficient lies in the range -1 to $+1$, -1 representing perfect negative correlation and $+1$ representing perfect positive correlation; a value of zero indicates the absence of correlation. Values of the ACF were used regardless of their significance.

Methods of analysis

The results of testing for density dependence were obtained for each of 25 replicate time-series for each parameter combination of each model. The results consisted of the proportion of cases where the test rejected density independence at $\leq 5\%$ probability. For each group of 25 time-series from a given parameter combination several basic statistics were calculated from the natural logarithm of abundance: (1) mean, (2) variance, (3) variance: mean ratio, (4) logarithmic range of abundances, (5) coefficient of skewness ($\gamma_3 = 1/N\sum\{[(X_t - \bar{X}_t)/s]^3\}$ where N is the number of generations), (6) coefficient of kurtosis ($\gamma_4 = 1/N\sum\{[(X_t - \bar{X}_t)/s]^4\} - 3$), (7) number of replicates with significant temporal trends at $P \leq 0.05$, and (8) the one-step autocorrelation coefficient, ACF.

The relationship between the frequency of detection of density dependence and these explanatory variables was analysed using multivariate analysis of covariance (MANCOVA), which was used to search for patterns in a way that allowed for differences between time-series generated using different models. The analyses were carried out in GLIM (Nelder and Wedderburn 1972; McCullagh and Nelder 1989).

The MANCOVA consisted of stepwise linear multiple regression analysis into which a factor coding for the model which generated the data was added. This was used to relate the proportion of replicates in which density dependence was detected with the statistics listed above. The regression was weighted for sample size, using a logit-link function, and assuming that errors in the ordinate followed a binomial distribution (Cox 1970). The potential explanatory variables (listed above) were all added into the model, followed by the factor coding for data-type, and the interactions between explanatory variables and data-type. Those variables that caused the smallest change in deviance were removed first. Changes in deviance were assumed to follow a chi-square distribution, allowing significance levels to be assigned to them. The treatment of overdispersion of McCullagh and Nelder (1989) was followed; that is, the model was rescaled in GLIM using a dispersion factor which was calculated by dividing Pearson's χ^2 by the residual degrees of freedom.

Table 1. The frequency of detection of density dependence (and limitation or attraction) using the various tests on the seven types of modelled data

Data type	Mean proportion of replicates in which density dependence (or limitation or attraction) was detected						
	regression of k on $\ln(N_t)$	regression of k on N_t	Bulmer's first test	Bulmer's second test	Pollard *et al.*'s test	Reddingius and den Boer's test	Crowley's test
Random walk (900)	19.9%	18.3%	6.0%	5.9%	5.0%	3.8%	2.6%
Exponential logistic (3600)	32.6%	46.2%	56.8%	0.2%	61.1%	45.9%	20.4%
Multiplicative logistic (3600)	16.3%	26.5%	27.4%	13.7%	52.2%	15.1%	60.6%
Power-form density dependent (3600)	6.4%	24.6%	41.5%	5.1%	50.8%	37.9%	41.5%

All time series were 20 generations long. Details of data used are given in the materials and methods section. The total number of time-series of each type analysed are given in parentheses in the first column

Results

Frequency of detection of density dependence

The frequency with which tests detected density dependence is shown in Table 1. Levels of spurious detection of density dependence from density-independent random-walk data varied with the amount of stochastic variation placed into time-series. No single test detected density dependence in less than 5% of time-series for all parameter combinations of the model; whether levels of spurious detection are acceptable depends on the choice of parameters. As expected, the two regression tests showed levels of detection of density dependence that had mean levels significantly greater than 5% (in a G-test) for random-walk data (Table 1). Other tests had mean rates of spurious detection that were similar to the conventional 5% probability level, and so were acceptable.

The power of the test of Pollard et al. ranged from 50.8% to 61.1% detection of density dependence when the data incorporated some form of density dependence (power form and exponential logistic density dependence, respectively). For each form of density dependent data these levels of detection were greater than those from all other tests of density dependence (Table 1). Bulmer's first test had the next greatest power: from 27.4% detection with multiplicative logistic data to 56.8% with exponential logistic data. Regression of k-value against abundance had low power with both the multiplicative logistic data (mean 26.5%) and the power-form density dependent data (24.6%). The power was greatest with exponential logistic data (46.2%), although this is perhaps surprisingly low for data which contain the same form of density dependence that is explicitly being tested for. Regression of k-value against logarithmic abundance had a power of 32.6% with exponential logistic data, only 16.3% with multiplicative logistic data (the same form of density dependence as is being tested for) and 6.4% with power-form density dependent data. The mean power of Bulmer's second test did not exceed 15% for density dependent data.

The power of Reddingius and den Boer's test was intermediate between that of regression of k-value against logarithmic abundance and Bulmer's first test. The levels of detection showed the same pattern of detection as Bulmer's first test; greatest power (45.9%) was with exponential logistic data and least (15.1%) with multiplicative logistic data. The power of Crowley's test was the exact opposite of this, greatest power (60.6%) occurred with the multiplicative logistic data, intermediate levels (41.5%) with power-form density dependent data and lowest power (20.4%) was found with exponential logistic data. The range of variation in levels of detection of density dependence (specificity) of Crowley's test is greater than all other tests.

What influences rates of detection?

It was shown by the MANCOVA's (Table 2) that tests of density dependence differed widely in the statistics from time-series which influenced the rate of detection. There is, however, one pattern that is common to all tests; the rate of detection of density dependence is correlated with the value of the autocorrelation coefficient. For Bulmer's second test the relationship between detection rate and autocorrelation accounted for only 0.7% of the deviance. For other tests this relationship was negative (increasingly negative autocorrelation resulted in greater frequencies of detection of density dependence) and accounted for a greater amount of the deviance in proportion detection than other variables. The pattern was strongest for Bulmer's first test and the test of Pollard et al. Where values of the slope of logit proportion detection against the autocorrelation coefficient were -4.90 ($P<0.001$) and -4.95 ($P<0.001$), respectively. In Bulmer's first test and the test of Pollard et al. this relationship accounted for as much as 77.9% and 65.3% of the deviance in logit proportion detection. The frequency of detecting limitation using Reddingius and den Boer's test was influenced less strongly by autocorrelation than Bulmer's first test and the test of Pollard et al.

Table 2. Results of analysis to test the influence of explanatory variables from time-series on the rate of detection of density dependence using multiple analysis of covariance

Test	Parameter	Value	SE	Deviance	Percentage of deviance
Regression of *k*-value against log abundance	Intercept for random-walk	0.789	0.548	232.8 ($P<0.001$)	24.6%
	Intercept for density dependent models	−1.144	0.511		
	ACF	−2.46	0.17	278.9 ($P<0.001$)	29.4%
	Kurtosis	0.016	0.006	7.1 ($P<0.01$)	0.1%
	Total				53.7%
Regression of *k*-value against abundance	Intercept for random-walk	2.359	0.422	64.3 ($P<0.001$)	4.8%
	Intercept for density dependent models	1.244	0.387		
	ACF	−3.71	0.27	309.4 ($P<0.001$)	23.1%
	Logarithmic range	−0.016	0.004	27.6 ($P<0.001$)	2.1%
	r	−0.067	0.013	27.2 ($P<0.001$)	2.0%
	Total				68.4%
Bulmer's first test	Intercept	3.24	0.14		
	ACF	−4.90	0.27	473.1 ($P<0.001$)	16.7%
	Trend	−0.118	0.009	199.0 ($P<0.001$)	7.0%
	Total				84.9%
Bulmer's second test	Intercept for random-walk data	−2.63	0.28	118.9 ($P<0.001$)	20.1%
	Intercept for density dependent models	−4.48	0.42		
	ACF	0.46	0.22	3.9 ($P<0.05$)	0.7%
	r	0.066	0.011	38.5 ($P<0.001$)	6.5%
	Total				28.0%
Pollard *et al.'s* test	Intercept	2.271	0.142		
	ACF	−4.95	0.24	836.5 ($P<0.001$)	65.3%
	Kurtosis	0.019	0.007	9.2 ($P<0.005$)	0.7%
	Total				66.1%
Reddingius and den Boer's Test	Intercept	2.17	0.12		
	ACF	−3.40	0.28	185.0 ($P<0.001$)	9.3%
	Kurtosis	−0.079	0.008	108.2 ($P<0.001$)	5.4%
	Trend	−0.126	0.013	105.6 ($P<0.001$)	5.3%
	Total				78.4%
Crowley's test	Intercept for random-walk data	−0.967	0.330	54.6 ($P<0.001$)	9.3%
	Intercept for density dependent models	0.524	0.186		
	ACF	−3.93	0.50	92.3 ($P<0.001$)	15.7%
	Trend	0.053	0.016	10.9 ($P<0.001$)	1.9%
	Kurtosis	0.024	0.011	5.5 ($P<0.05$)	0.1%
	Total				37.2%

The *y* variable is logit proportion detection of density dependence from a given test. Errors in *y* were assumed to be binomially distributed. Only variables that were significantly ($P \leq 0.05$) correlated with detection rates are given in the table. For Bulmer's first test, Pollard *et al.'s* test and Reddingius and den Boer's test the MANCOVA reduced to a multiple regression when non-significant variables were removed. For some of the tests for density dependence dissimilar intercepts of the regression lines resulted as a result of the different data types being separated in the ANCOVA. In other cases the ANCOVA reduced to a single intercept for all data types

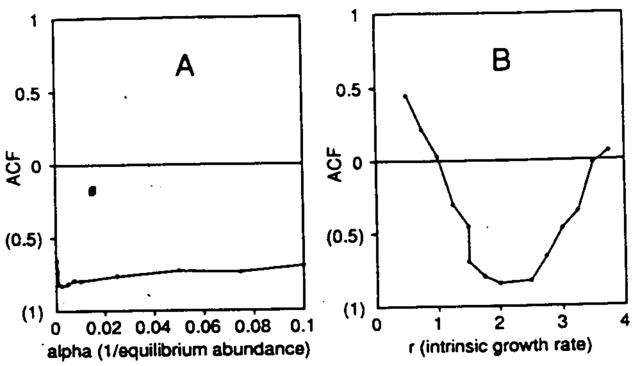

Fig. 1A, B. The effects of size of **A** the equilibrium abundance and **B** intrinsic growth rate on the strength of autocorrelation from the exponential logistic model (May 1974). Other parameter values were SD = 0.00005, $r = 2.5$ for A and $\alpha_t = 0.001$ for B. In A the size of the equilibrium abundance is given by $1/\alpha_t$. All series were 20 generations in length and the initial abundance was set close to the carrying capacity. A total of 10000 series were used to obtain a mean value of the autocorrelation coefficient for each value of intrinsic growth rate or mean value of α_t

(slope = −3.40, $P < 0.001$), but this relationship accounted for 67.7% of the deviance. Equivalent figures for regression of k-value against abundance were a slope of −3.71 ($P < 0.001$) and 59.5% of the deviance. For regression of k-value against logarithmic abundance and Crowley's test, levels of detection were less strongly influenced by autocorrelation; slopes were −3.93 and −2.46 respectively, but only 29.4% and 25.9% of the deviance was accounted for by the respective methods.

The effect of autocorrelation on detection

The size of the equilibrium abundance ($1/\alpha_t$) had little effect on the strength of autocorrelation within the normally expected range of values (Fig. 1). Note that although the data in Fig. 1 are from the exponential logistic model, the same effect applies to data from the other two models which were used, and the parameter values (the size of α_t in Fig. 1B and the value of r in Fig. 1A) are not critical to the findings. In all three density dependent models the value of the intrinsic growth rate had a much greater influence on the value of the autocorrelation (Fig. 1B). It is unclear whether the effect of r is due to a direct effect of reproduction or whether it is an effect of the strength of density dependence ($r \cdot \alpha_t$). However, at small values of r the amount of variability in the time-series is relatively small and measured values of the autocorrelation coefficient become close to zero. As r becomes large dynamics become deterministically chaotic which hinders detection of density dependence and causes the value of the autocorrelation coefficient to again approach zero or become positive.

The effect of chaos on detection

The results of applying tests to data in which the dynamics are chaotic and are from the exponential form of the logistic equation are shown in Table 3. For the deterministic version of this model detection levels remained high (100%) for Bulmer's first test, the test of Pollard et al. and Crowley's test. However detection rates declined to 82% for regression of k on abundance with the series in which the dynamics were most chaotic ($r = 3.5$), and detection levels could presumably become even lower as dynamics become increasingly chaotic. The frequency with which tests failed to detect density dependence increased as series were made more stochastic. For Bulmer's first test and the test of Pollard et al. only the most stochastic time-series considered showed levels of detection which declined as dynamics became more chaotic. Crowley's test was minimally influenced by chaos; detection levels observed were not less than 96.7%.

Discussion

How useful are tests?

Overall the test Pollard et al. (1987) emerges as the best test. When data were not density dependent it detected

Table 3. The frequency of detection of density dependence from time-series in which dynamics were chaotic

r	Mean proportion of replicates in which density dependence (or limitation or attraction) was detected											
	Regression of k on N_t			Bulmer's first test			Pollard *et al.*'s test			Crowley's test		
	$SD = 0$	$SD = 0.00005$	$SD = 0.00025$	$SD = 0$	$SD = 0.00005$	$SD = 0.00025$	$SD = 0$	$SD = 0.00005$	$SD = 0.00025$	$SD = 0$	$SD = 0.00005$	$SD = 0.00025$
2.7	99.8%	100%	100%	99.8%	100%	100%	99.8%	100%	100%	99.8%	100%	99.9%
2.8	99.9%	100%	100%	99.9%	100%	100%	99.9%	100%	100%	99.9%	100%	99.9%
2.9	99.8%	100%	100%	99.8%	100%	100%	99.8%	100%	100%	99.8%	99.9%	100%
3.0	99.9%	100%	58.6%	99.9%	100%	94.4%	99.9%	100%	89.7%	99.9%	99.9%	96.9%
3.1	99.9%	100%	53.2%	99.9%	100%	93.8%	99.9%	100%	90.5%	99.9%	100%	96.7%
3.2	99.8%	99.9%	46.3%	99.8%	100%	92.8%	99.8%	100%	87.9%	99.8%	99.9%	97.3%
3.3	100%	96.3%	34.1%	100%	100%	90.6%	100%	100%	83.9%	100%	100%	98.4%
3.4	91.2%	86.6%	32.6%	100%	100%	89.5%	100%	100%	81.3%	100%	100%	97.3%
3.5	82.1%	64.4%	21.6%	100%	100%	87.1%	100%	100%	77.6%	100%	100%	96.7%

Series were of 20 generations in length and were generated using the exponential form of the logistic equation. The table gives the % of 1000 time-series for which the process was detected at the 5% probability level. Time-series were generated using $N_{t+1} = N_t \exp[r(1 - 0.001 N_t)]$ and an initial abundance of $1000 - z$, where z is a uniform random variate with a mean of zero and a variance of 10 density dependence from approximately 5% of time-series. When data were density dependent it was the test which was most reliable at detecting this density dependence (50.8–61.1% detection, depending on the form of density dependence present). Other tests were less reliable, either because of unreasonably high (significantly >5%) rates of detection when density dependence was absent, or because of very different detection rates with different forms of density dependent data. Regressions of k-value on abundance and k-value on logarithmic abundance are rejected because they both produced high levels of detection with density independent data (18.3% and 19.9% respectively). The bias in regression tests has also been reported by St. Amant (1970), Maelzer (1970), Kuno (1971), Itô (1972), Vickery and Nudds (1984), Latto (1989) and Vickery (1991). Tests which are weak because they showed low levels of detection with certain forms of density dependent data (a high *specificity*) were Bulmer's first test, Reddingius and den Boer's test for limitation and Crowley's test for attraction. The least useful tests were Bulmer's second test and both regression tests, because of low rates of detection with all three forms of density dependent data. Additionally, both regression tests are precluded from general use because they had greater mean frequencies of detection with density independent random-walk data than density dependent data (Table 2).

The effect of autocorrelation on detection

A number of tests appear to be tests of autocorrelation, not density dependence (or limitation/attraction). This is because the tests were more influenced by autocorrelation than the presence of density dependence. However, since the concepts of density dependence and autocorrelation are inextricably related any conclusions about this must remain tentative. The MANCOVAs showed that autocorrelation could account for a high proportion of the deviance in the rate of detection; 77.9%, 65.4% and 67.7% (assuming that the variance that is shared between explanatory variables is split in proportion to the amount of variance that is specific to that variable) for Bulmer's first test, the test of Pollard et al. and Reddingius and den Boer's test, respectively. For these three tests there were no differences in detection (other than differences which could be explained by autocorrelation) for data where density dependence was present and data where it was absent. Regression of k-value on abundance was also mainly a test of autocorrelation: autocorrelation accounted for 59.5% of the deviance, where as the difference between density dependent and density independent data accounted for only 4.8% of the deviance. It is not surprising that Bulmer's first test is a test of autocorrelation, as the method was derived from von Neumann's ratio, which is a measure of autocorrelation (von Neumann 1941). Crowley's test is the only technique which is apparently not solely a test of autocorrelation and does not have low levels of detection. For Crowley's test autocorrelation accounted for only 25.9% of the deviance in logit proportion detection. Bulmer's second test and regression of k-value on logarithmic abundance were least strongly influenced by autocorrelation, but they are not useful for other reasons (see above).

The relationship between density dependence and autocorrelation is complex (see Figure 1). The value of the equilibrium abundance ($1/\alpha_i$) has relatively little effect on the value of the autocorrelation coefficient within normally observed limits; however, both small and large values of the intrinsic growth rate greatly increase the value of the autocorrelation coefficient. Positive values of the autocorrelation coefficient are similar to those expected from random walks. When r is small the amount of variability in abundances is small, which might limit the resolving power of tests for density dependence. When r is large, dynamics are chaotic and this might limit detection and break down autocorrelation. Because of the structure of the population models used in the present simulations the effects of reproduction *per se* cannot be separated from those of the strength of density dependence. With tests for density dependence that are less strongly influenced by autocorrelation, it is unclear how much of the this effect is attributable to a loss of statistical power (through autocorrelation) and how much is due to qualitative changes in the density dependence which are expressed through autocorrelation. Qualitative changes in density dependence might be caused by different intrinsic growth rates causing differences in the frequency with which density dependence acts. However, all findings of this sort must remain tentative because of the close relationship between density dependence and autocorrelation.

What influences rates of detection?

As we have seen, tests differ in their rates of detection with different forms of density dependent data and when density dependence is absent. One way forward may be to use these differences in specificity of tests to our advantage by carrying out several tests which differ in their specificity. From patterns of detection of density dependence in particular time-series, it may be possible to make some (albeit weak) inferences about the form of density dependence. Tests also differ in the extent to which detection is influenced by autocorrelation. Hence, a test which is little influenced by autocorrelation might be used with a test for which autocorrelation has a greater effect, giving a wider range of possible ways of detecting density dependence (and limitation/attraction). In particular, the relative detection levels for the three forms of density dependent data of Crowley's test were opposite to those of the test of Pollard et al., Reddingius and den Boer's test and Bulmer's first test (Table 1). Crowley's test was also less influenced by autocorrelation than the other tests which did not have low rates of detection. Hence, Crowley's test is apparently complementary to the test of Pollard et al., Reddingius and den Boer's test and Bulmer's first test. The idea of combining several tests into a single test, in a statistically acceptable way, will form the subject of a later paper (Holyoak and Crowley, in preparation).

Testing for limitation is apparently similar to testing for density dependence. Levels of detection for different forms of density dependent data with Crowley's test were intermediate between those of the test of Pollard et al. and Bulmer's first test and the effects of autocorrelation were also similar. Given the similarity of testing for limitation to testing for density dependence, it is questionable whether the concept of limitation (*sensu* Reddingius and den Boer 1989) is useful. There would seem to be little point in recognizing another concept which in practice is identical to density dependence.

Only Bulmer's first test and Reddingius and den Boer's were influenced by temporal trends in abundances. These tests gave lower rates of detection with time-series that contained temporal trends in abundances (Table 2). It is not surprising that temporal trends have this effect, since trends are obviously and strongly related to autocorrelation (Diggle 1990). However, it is interesting that trends have an effect which is independent of autocorrelation for these tests and not for the test of Pollard et al. since Pollard et al. (1987) found that their test was not influenced by temporal trends, but that Bulmer's first test was influenced by trends. Additionally, Reddingius and den Boer (1989) predicted that their test would be influenced by temporal trends. This adds to the arguments in favour of the test of Pollard et al. or a combination of tests.

The effect of chaos on detection

It has been argued that Bulmer's first test can continue to efficiently detect density dependence as dynamics become increasingly chaotic (Mountford and Rothery 1992). In analyses of time-series of 20 generations, it was found that detection levels from regression of k on abundance became low (as low as 21.6%) as the dynamics represented in series became increasingly chaotic and stochastic. Bulmer's first test also declined in power (down to 87.1%) with series of 20 generations and (54.3%) with 10000 replicates of 10 generations (and $SD = 0.00025$). The test of Pollard et al. showed detection levels which declined in a similar way to those of Bulmer's first test with increasingly chaotic and stochastic dynamics. Detection levels of attraction using Crowley's test remained high (96.7–100%). It is concluded that Mountford's results for Bulmer's first test are incorrect; that is, we cannot detect density dependence reliably using this conventional technique if dynamics are made slightly more chaotic or stochastic. This is also true of the test of Pollard et al. and regression of k on abundance. Conversely Crowley's test seems capable of continuing to detect attraction even in the most stochastic and chaotic dynamics. This difference from the other techniques concords with the low influence of autocorrelation on Crowley's test. Autocorrelation becoming more positive would be expected to decrease the detection rates of those tests which rely on it to show the presence of density dependence (and attraction or limitation).

Conclusion

The test of Pollard et al. Bulmer's first test and Reddingius and den Boer's test are the best techniques for detecting density dependence available to us. Of these techniques, the test of Pollard et al. appears to be the most able to detect density dependence when the process was made increasingly stochastic and when the form of density dependence acting was varied, and so is of most general applicability. It should be added that testing for attraction (Crowley 1992) is complementary to testing for density dependence. However, the usefulness of individual tests will depend partly on which population model is the best descriptor of population data. The inappropriateness of many tests for analysing data of this kind might partly explain why many tests of density dependence frequently fail to detect density dependence in insect populations. Testing for limitation does not appear to differ from testing for density dependence; based on rates of detection and the influence of autocorrelation.

Acknowledgements. I wish to thank Phil Crowley, John Lawton, Mike Hassell, Mick Crawley, Hefin Jones, Katriona Shea, Mark Rees and an anonymous referee for their helpful comments on the manuscript. The study was supported by a studentship from the Natural Environment Research Council.

References

Bellows TS Jr (1981) The descriptive properties of some models for density dependence. J Anim Ecol 50:139–156

Bulmer MG (1975) The statistical analysis of density dependence. Biometrika 31:901–911

Cox DR (1970) Analysis of binary data. Monographs on statistics and applied probability. Chapman and Hall, London

Crowley PH (1992) Density dependence, boundedness and attraction: detecting stability in stochastic systems. Oecologia 90:246–254

Crowley PH, Johnson DM (1992) Variability and stability of a dragonfly assemblage. Oecologia 90:260–269

Dempster JP (1983) The natural control of butterflies and moths. Biol Rev 58:461–481

Diggle PJ (1990) Time series. A biostatistical introduction. Oxford statistical science series 5. Clarendon Press, Oxford

Gaston KJ, Lawton JH (1987) A test of statistical techniques for detecting density dependence in sequential sequences of animal populations. Oecologia 74:404–410

Hassell MP, Latto J, May RM (1989) Seeing the wood for the trees: detecting density dependence from existing life-table studies. J Anim Ecol 54:323–334

Holyoak M, Lawton JH (1992) Detection of density dependence from annual censuses of bracken-feeding insects. Oecologia 91:425–430

Itô Y (1972) On the methods for determining density dependence by means of regression. Oecologia 10:347–372

Kuno E (1971) Sampling error as a misleading artefact in key-factor analysis. Res Popul Ecol 13:28–45

Latto J (1989) The analysis of density dependence in insect populations. Unpublished Ph.D. thesis, Imperial College, University of London

Maelzer DA (1970) The regression of log (N_{n+1}) on $\log(N_n)$ as a test of density-dependence: an exercise with computer constructed density-independent populations. Ecology 51:810–822

Manly BFJ (1991) Randomization and Monte Carlo methods in biology. Chapman and Hall, London

May RM (1974) Biological populations with non-overlapping generations: stable points, stable cycles and chaos. Science 186:645–647

May RM (1989) Detecting density dependence in imaginary worlds. Nature 338:16–17

Maynard Smith J, Slatkin M (1973) The stability of predator-prey systems. Ecology 54:384–391

McCullagh P, Nelder JA (1989) Generalized Linear Models, 2nd ed. Chapman and Hall, London

Mountford MD, Rothery P (1992) Density-dependent populations. Nature 356:391

Nelder JA, Wedderburn RWM (1972) Generalised linear models. J R Stat Soc A 135:370–384

Neumann JW von (1941) Distribution of the ratio of the mean square successive difference to the variance. Ann Math Stat 12:367–395

Pollard E, Lakhani KH, Rothery P (1987) The detection of density dependence from a series of annual censuses. Ecology 68:2046–2055

Reddingius J, den Boer PJ (1989) On the stabilization of animal numbers. Problems of testing. 1. Power estimates and estimation errors. Oecologia 78:1–8

St Amant JLS (1970) The detection of regulation in animal numbers. Ecology 51:823–828

Slade NA (1976) Statistical detection of density dependence in a series of sequential censuses. Bull Ecol Soc Am 57:35

Slade NA (1977) Statistical detection of density dependence from a series of sequential censuses. Ecology 58:1094–1102

Stiling PD (1987) The frequency of density dependence in insect host-parasitoid systems. Ecology 68:844–856

Stiling PD (1988) Density-dependent processes and key-factors in insect populations. J Anim Ecol 57:581–594

Turchin P (1990) Rarity of density dependence or population regulation with lags? Nature 344:660–663

Varley GC, Gradwell GR (1960) Key factors in population studies. J Anim Ecol 29:399–401

Vickery WL (1991) An evaluation of bias in *k*-factor analysis. Oecologia 85:413–418

Vickery WL, Nudds TD (1984) Detection of density dependent effects in annual duck censuses. Ecology 65:96–104

Vickery WL, Nudds TD (1991) Testing for density-dependent effects in sequential censuses. Oecologia 85:419–423

Woiwod IP, Hanski I (1992) Patterns of density dependence in moths and aphids. J Anim Ecol 61:619–630