Abstract. – An analysis of alternative methods for detecting trends in a series of abundance indices is carried out through simulation. The alternative procedures explored have been applied to analysis of relative abundance indices of dolphins in the eastern Pacific Ocean. They include a linear test over a moving time-period, and a nonparametric procedure based on smoothing of the time-series of abundance indices. Results indicate that the nonparametric procedure outperforms the linear tests in most of the situations tested.

A comparison of tests for detecting trends in abundance indices of dolphins

Alejandro A. Anganuzzi
Inter-American Tropical Tuna Commission
8604 La Jolla Shores Drive, La Jolla, California 92037-1508

Manuscript accepted 26 January 1993.
Fishery Bulletin, U.S. 91:183-194 (1993).

An important part of the analysis of any set of abundance indices is the application of an objective procedure or test to determine whether changes in the estimates are due to random fluctuations in conditions of the sampling procedure or to actual changes in the population size. Such a procedure must exhibit certain properties in order to be effective. Among these properties, perhaps most important is the power of the test for a given significance level.

In deriving conclusions about changes in the size of a population, we can fall into two types of error. First, we can erroneously conclude that population size has changed when, in fact, differences in estimates are due to random errors. This is usually known as a Type-I error. A Type-II error occurs when we conclude that the estimates reflect random fluctuations when, in fact, there has been a change in population size. The probability of falling into a Type-I error is usually referred to as the significance level of the test. The power of a test is defined as 1 minus the probability of a Type-II error. An ideal test will minimize the trade-offs between both types of error. Another desirable property of a test is robustness to underlying assumptions about the populations. For example, tests commonly carried out to detect changes in population size are based on specific assumptions about the error structure of the estimates (e.g., normality) and the model that would best describe the population size as a function of time (e.g., linear, exponential; see, for example, Gerrodette 1987).

In the case of dolphin stocks involved in the tuna fishery in the eastern Pacific Ocean (EPO), it has been recommended that their management should include both estimates of absolute abundance, derived from research-vessel data (RVD), and analysis of trends in relative abundance, derived from tuna-vessel observer data (TVOD) (IWC 1992:218). In the case of EPO dolphin stocks, the use of TVOD seems the natural choice for analyzing trends, given the vast amount of low-cost information available from the observer programs. However, for this analysis to be effective, it is necessary to obtain abundance estimates with a constant bias over the years, or, at least, a bias that shows no trend over the years. Procedures developed by the Inter-American Tropical Tuna Commission (IATTC) to analyze the TVOD, described in Buckland & Anganuzzi (1988) and Anganuzzi & Buckland (1989), were specifically aimed at reducing the magnitude of year-to-year fluctuations in the estimates due to changing biases. These procedures were complemented with more specific analyses when there were reasons to suspect that biases might be changing, for example, due to widespread use of high-resolution radar for the detection of birds (Anganuzzi et al. 1991). However, in spite of the robustness of the methods, randomly fluctuating biases (an extra source of

variability) may still affect estimates from year to year. This problem may not be exclusive to the TVOD estimates; interannual variability also seems to affect estimates of relative abundance derived from research-vessel data (Wade & Gerrodette 1992). These biases and imprecise estimates will affect the performance of statistical tests designed to detect trends and, ultimately, our ability to draw conclusions about the status of populations.

For the analysis of trends in the EPO dolphin stocks, Buckland & Anganuzzi (1988) applied a linear test for trends over a moving period of 5 yr, although they expressed concern about the low power of such a test. Edwards & Perkins (1992) extended the moving time-frame to 10 yr to increase the power. However, such a test still shows some undesirable properties. Given the inadequacy of the tests based on linear regressions, Buckland et al. (1992) proposed a different procedure, based on a nonparametric regression, which addresses some of the problems exhibited by the linear test.

In this paper, the characteristics of these tests are discussed and compared by analyzing their performance in a number of simulated scenarios.

Current tests for trends

Linear tests

Buckland & Anganuzzi (1988) tested for linear trends over successive 5 yr periods by carrying out a weighted linear regression of abundance index vs. time. Each individual estimate was weighted by the inverse of its variance, calculated by applying a bootstrap procedure. The null hypothesis for the test is that no change has occurred in the population, i.e., that the slope of the regression is equal to zero. As the authors noted, the test has low power since it estimates precision from the deviations of only five estimates from a straight line. Power can be increased by extending the moving time-period to incorporate more years in the test. Unfortunately, this also increases the probability of violating the assumption of a constant rate of change implicit in the linear model being fitted to the estimates (Edwards & Perkins 1992).

The linear test also fails to consider the precision of estimates adequately. Variances of the estimates are not taken into account except as weights in the regression. As a consequence, only the ratios of the variances between estimates are relevant, and not their absolute values. For example, if for any given series of estimates we double the variance of each individual estimate, the results of the test will remain unchanged.

Weighting by the inverse of the variance can also present other problems. Suppose, for example, that the variance of the estimate is not independent of the estimate itself, but that the variance is correlated with the estimate, i.e., the coefficient of variation (CV=ratio of standard error to point estimate) is constant. In this case, a very low estimate (and especially in the case of an outlier) with a correspondingly small variance will become an influential observation. A linear test for trends will then indicate that there was a decline in the population if that estimate is at the end of the moving period, or a significant increase if it is at the beginning. An example from the EPO dolphin abundance estimation is the case of the 1983 index of relative abundance for the northern stock of offshore spotted dolphin *Stenella attenuata*, which was an anomalous index as a result of a very strong El Niño event (Fig. 1). In such cases, where the error distribution of the estimates seems to be better approximated by a lognormal distribution, it would be more appropriate to apply weights (wt) defined as

$$wt = \ln(1+CV^2)^{-1}. \quad (1)$$

For the comparisons in this paper, two versions of the linear test are applied: the original 5 yr linear test with inverse variance weighting applied by Buckland & Anganuzzi (1988), and a 10 yr linear test with weights as described by Eq. 1.

Smoothed trends

The approach taken by Buckland et al. (1992) differs considerably from the method just described. First, they replaced the assumption of a parametric model for the underlying change in population with a nonparametric model. Among the many possible choices for such a model, they selected the smoothing algorithm known as '4253H, twice' (Velleman & Hoaglin 1981) on the basis of a comparison described by K. L. Cattanach and S. T. Buckland (SASS Environ. Model. Unit, MLURI, Craigiebuckler, Aberdeen, Scotland, unpubl. manuscr.). The adoption of a nonparametric model increases the robustness of the test to model misspecification, a problem that affects the linear test. Furthermore, the procedure, which involves the use of a compound running median, incorporates the information from nearby years into calculation of the smoothed estimate for a particular year, therefore reducing the influence of possible outliers and increasing the precision of each smoothed estimate. The smoothed test also provides a different way of looking at the trend. Instead of the trend being described by a single parameter (the slope of a linear regression), the sequence of smoothed estimates constitutes the best estimate of the underlying trend.

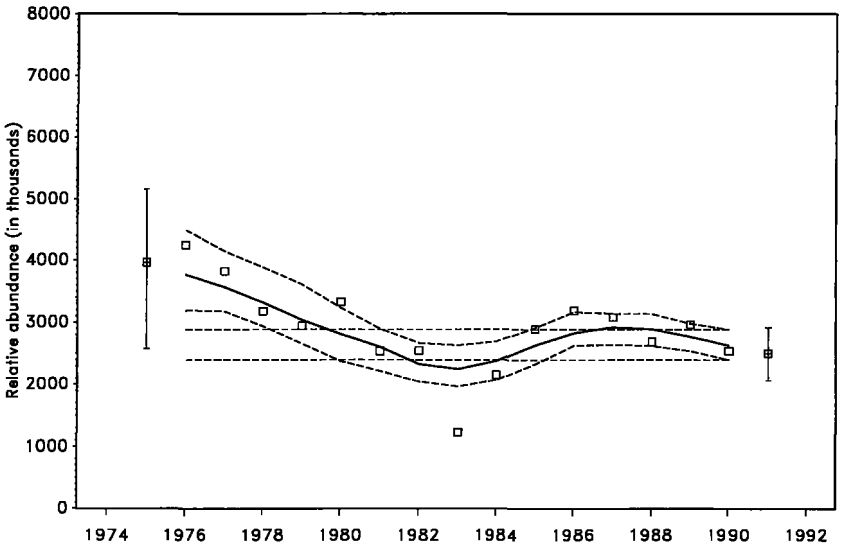

Figure 1

Smoothed trends in abundance of the northern offshore stock of spotted dolphin *Stenella attenuata* in the eastern tropical Pacific Ocean. Broken lines indicate ~85% confidence limits. Horizontal lines correspond to 85% confidence limits for the 1990 estimate. If both the 1990 confidence limits lie above the upper limit for an earlier year, abundance has increased significantly between that year and 1990 ($p<0.05$); if both limits lie below the lower limit for an earlier year, abundance has decreased significantly.

To obtain confidence intervals of the smoothed estimates, Buckland et al. (1992) combined smoothed estimates and bootstrap replication using the following procedure. First, they obtained 79 bootstrap estimates of the abundance index for each year. Next, they built bootstrap replicates of the series of estimates by taking, for example, the first bootstrap estimate for each year to obtain the first replicated series. They smoothed each replicated series, thereby obtaining 79 smoothed estimates for each year. Finally, they sorted the smoothed estimates within each year and obtained 85% confidence intervals based on the percentile method (Buckland 1984). The median of the smoothed bootstrap replicas is considered to be the best smoothed estimate.

The confidence intervals thus calculated allow a direct comparison between estimates. If the confidence intervals for two estimates do not overlap, then they are significantly different at a level of ~5%. An example based on estimates of relative abundance for the northern stock of offshore spotted dolphin is shown in Fig. 1. Since the last and first smoothed estimates are too variable, Buckland et al. (1992) recommend excluding them from the comparisons. The significance level is approximate, since it depends on normality of the estimates, homogeneity of the variances of estimates, and their independence. The second condition is often not met, but it can be shown that results are robust-to-moderate departures from it. The third condition is not met for estimates that are close in time, due to the correlation introduced by smoothing, and the relative importance of this effect is discussed further below.

Simulation comparison between tests

To illustrate the difference in performance between methods, a simulation study was carried out. Series of estimates were simulated by assuming different scenarios of underlying trends in the population over a period of 25 yr. The following scenarios were chosen.

Stable population Population exhibits no trend over the simulated period. This scenario provides us with an estimate of the probability of a Type-I error detecting a trend when, in fact, there is none, or obtaining a "false positive." Under this scenario, a percentage of detected trends close to 5% would be expected for a test based on a significance level of 5%.

Rapid decline Population remains at a constant level for a period of time, and then declines sharply over a period of 3 yr to 50% of its previous level. After the decline, the population recovers at a rate of 5% per yr.

Steady decline Population declines exponentially at an annual rate of 10%.

Steady cycle Population follows a sinusoidal change over the simulated period, completing one cycle over the 25 yr. Amplitude of the cycle is 30% of the original population size.

These scenarios are intended as a means of highlighting properties of the different tests and not as an exhaustive list of possible situations. Each scenario was replicated 100 times for different combinations of sources of variation. Variability in the estimates was assumed as coming from two different sources: (1) **Interannual variability**, resulting in a point estimate for each year t of

$$I_{0t} = N_t e^{z},$$

where z is a random variable distributed as N (0, σ^2), and (2) **precision of the estimate**, represented in the distribution of the simulated bootstrap replicas for year t, I_{bt}, as

$$I_{bt} = I_t e^v, \text{ for } b=1, \ldots, 79,$$

where v is a random variable distributed as N (0, ε^2).

The rationale for a setup with independent control of two sources of variation and a lognormal error structure lies in the properties of the estimates and in the fact that the linear and smoothed tests deal differently with both components of variation. The choice of a lognormal error structure can be justified by considering that the abundance estimates are naturally constrained to be positive. The choice of error structure for the simulation can be further justified by an analysis of the available TVOD estimates (Fig. 2), which shows that the main target stocks tend to have constant CV's, particularly in recent years when observer coverage of the tuna fleet increased and more information was available for abundance estimation. There is considerable variability in some stocks, due to changing levels of targeting from the purse-seine fleet that result in unequal sample sizes from year to year.

The argument for including two sources of variation in the estimates is based on the possibility that actual relative abundance indices are affected by random biases from year to year. Under standard assumptions, σ^2 and ε^2 should be equal. However, estimates of dolphin abundance may exhibit an additional variability, represented in this setup as $\sigma^2 > \varepsilon^2$. This can be attributed to randomly-fluctuating biases, such as those introduced by changing environmental conditions or differences in the way the purse-seine fishery operates. It is important to separate these two components since, for example, in the case of the linear test, the results are affected only by the variability represented by σ^2.

Therefore, it was assumed in the simulation that estimates are lognormally distributed around the underlying trend with a constant coefficient of variation. Figure 3 illustrates one simulated series for each of the different scenarios. The simulations were repeated for different combinations of CV_σ and CV_ε, the coefficients of variation for both sources of variation.

To compare the test for trends, two diagnostics were used.

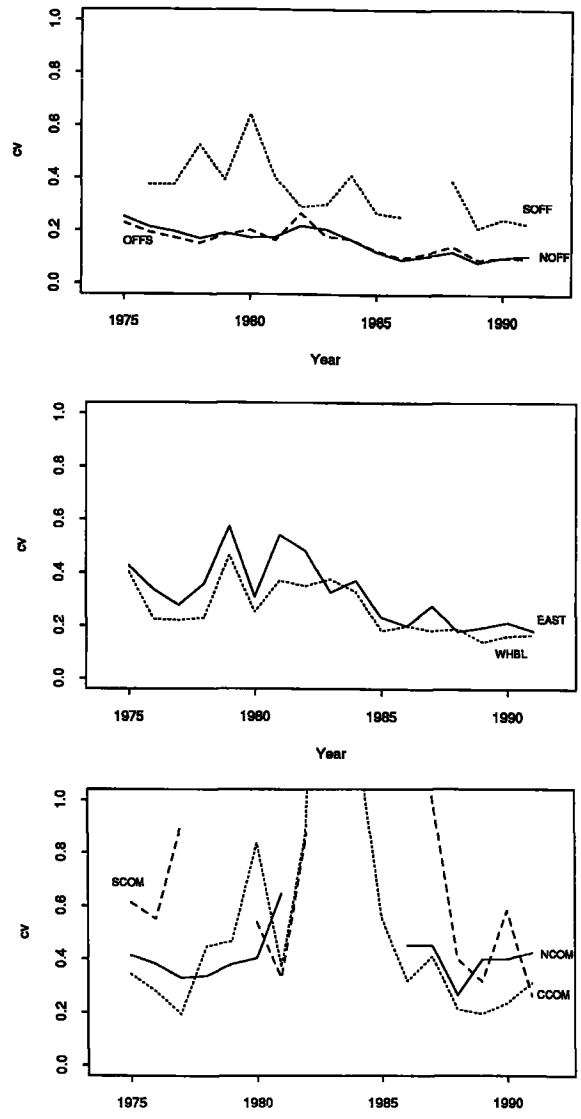

Figure 2

Coefficients of variation in relative abundance estimates of dolphin stocks in the eastern Pacific Ocean as a function of time. **NOFF**=northern stock of the offshore spotted dolphin *Stenella attenuata*; **SOFF**=southern stock of the offshore spotted dolphin; **EAST**=eastern stock of the spinner dolphin *Stenella longirostris*; **WHBL**=whitebelly stock of the spinner dolphin; **NCOM**=northern stock of the common dolphin *Delphinus delphis*; **CCOM**=central stock of the common dolphin; **SCOM**=southern stock of the common dolphin.

Number of detected trends For the linear trends, this is the number of 10 yr periods with slopes significantly different from zero at the 5% significance level. For the smoothed test, it is the number of significant differences between the next-to-last estimate and the estimate 10 yr earlier. In this way, the comparison is based on the same number of tests for each method. Since there are 25 yr simulated in each replica, a total of 1500 tests were carried out in the simulation of each scenario.

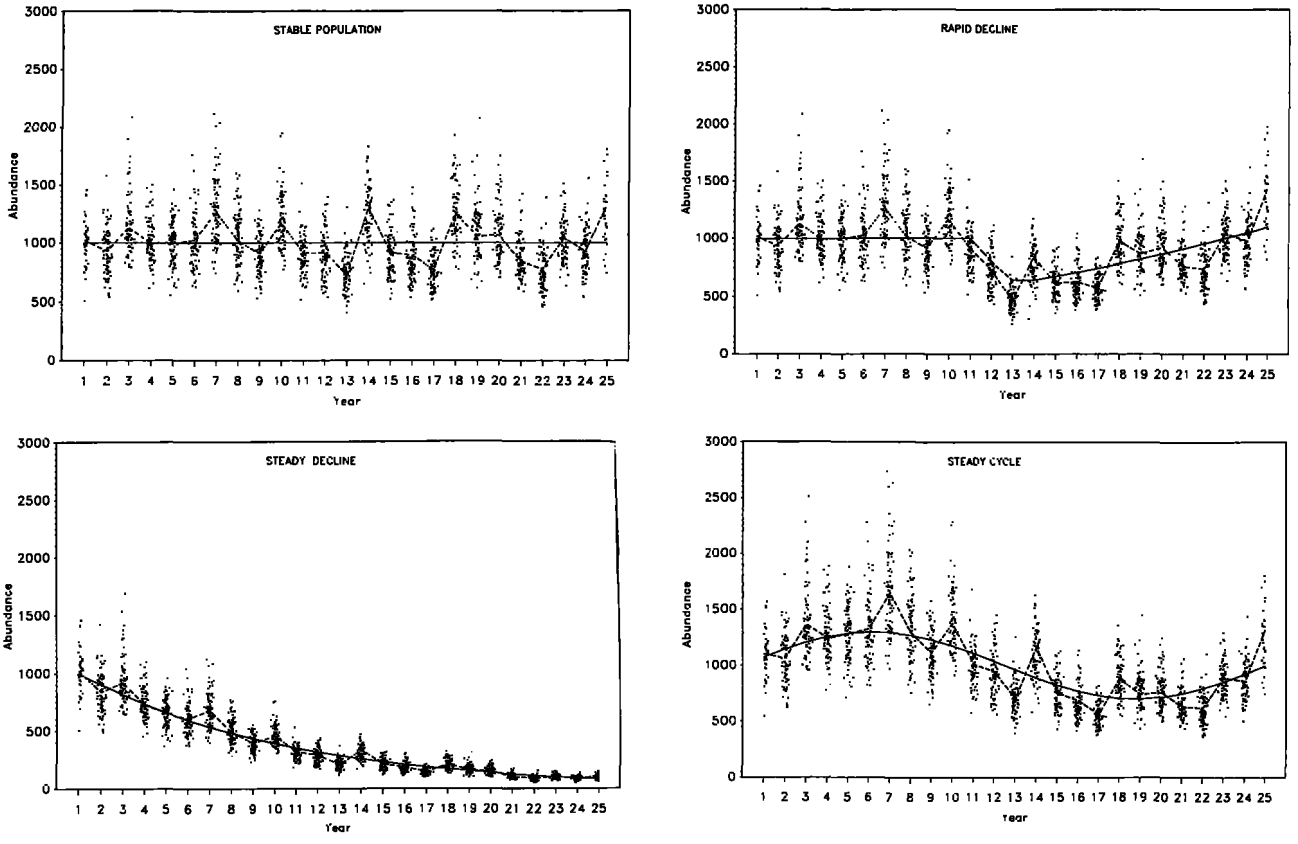

Figure 3

Example of a simulated series of relative abundance estimates for each of four scenarios. Solid line represents true underlying trend in the population. Dashed line connects simulated point estimates. Distribution of simulated bootstrap estimates is represented by the distribution of dots.

Ratio between estimates As a way of assessing how well each method describes the underlying trend in the population, an estimated rate of change was obtained. For the smoothed test, this is the ratio of two smoothed estimates separated by 10 yr. For the linear test, it is the ratio of the corresponding estimates calculated from the linear regression. These estimated rates of change were then compared with the true rates of change and the discrepancies summarized as average absolute error.

Correlation in the smoothed estimates

One of the problems of the smoothed test is that the smoothing procedure induces a correlation between estimates. This lack of independence affects the results of the comparison between estimates close in time, and it is therefore important to assess the magnitude of this correlation and how it is reduced as the separation in time between estimates increases. To investigate this, the following Monte Carlo procedure was carried out on the series of relative abundance estimates for dolphin stocks in the EPO reported by Anganuzzi et al. (1992).

1 For each year, 79 estimates were sampled with replacement from the distribution of bootstrap estimates of relative abundance. The 79 estimates were available from the standard bootstrap procedure used to estimate confidence bounds in the relative abundance estimation (Anganuzzi & Buckland 1989).

2 Each of the 79 trajectories obtained in the previous step were smoothed, and 85% confidence limits for the resulting smoothed estimates were obtained based on the percentile method. This step is essentially the application of the smoothed test.

3 Steps 1 and 2 were repeated 100 times, therefore obtaining 100 estimates of the lower and upper confidence bounds for the smoothed estimates for each year.

4 A correlation matrix between years for both upper and lower limits of the confidence bounds was estimated on the basis of results of the previous step.

5 Estimates of correlation as a function of the distance in time between estimates were obtained. This was done by averaging over all correlation coefficients between estimates separated by a given number of years, that is, by taking the average of elements in the subdiagonals of the correlation matrix obtained in step 4.

Results

Number of detected trends

The results of this analysis are shown in Figs. 4–7, as the number of detected trends each year in 100 simulations for the different scenarios. The underlying trends are shown on an arbitrary scale to relate changes in performance of the tests to changes in population trajectory.

Stable population This scenario can be used to assess the actual level of significance of the tests. An ideal procedure for detecting trends would indicate significant trends in ~5% of the tests under this scenario, given that the significance level is set at 5%. Results from the simulations are shown in Fig. 4. Both linear tests show an actual significance level close to the expected value. These results were relatively robust to the different values of CVs tested. The smoothed test also performs well in all cases, except when interannual variation exceeds the precision of the estimate. For example, for CV_σ=0.2 and CV_ε=0.3, the percentage of detected trends was ~10%.

Rapid decline In this scenario, different trade-offs of the tests are illustrated by their performance along the simulated period (Fig. 5). For CV_σ=CV_ε, at the beginning of the period both the linear and smoothed tests have similar proportions of detected trends, close to the nominal significance level. As the underlying trend in the period included in the tests departs from linearity, the smoothed test tends to outperform both linear tests. For CV_σ=CV_ε=0.2, the smoothed test indicates a maximum of almost 80% significant trends in

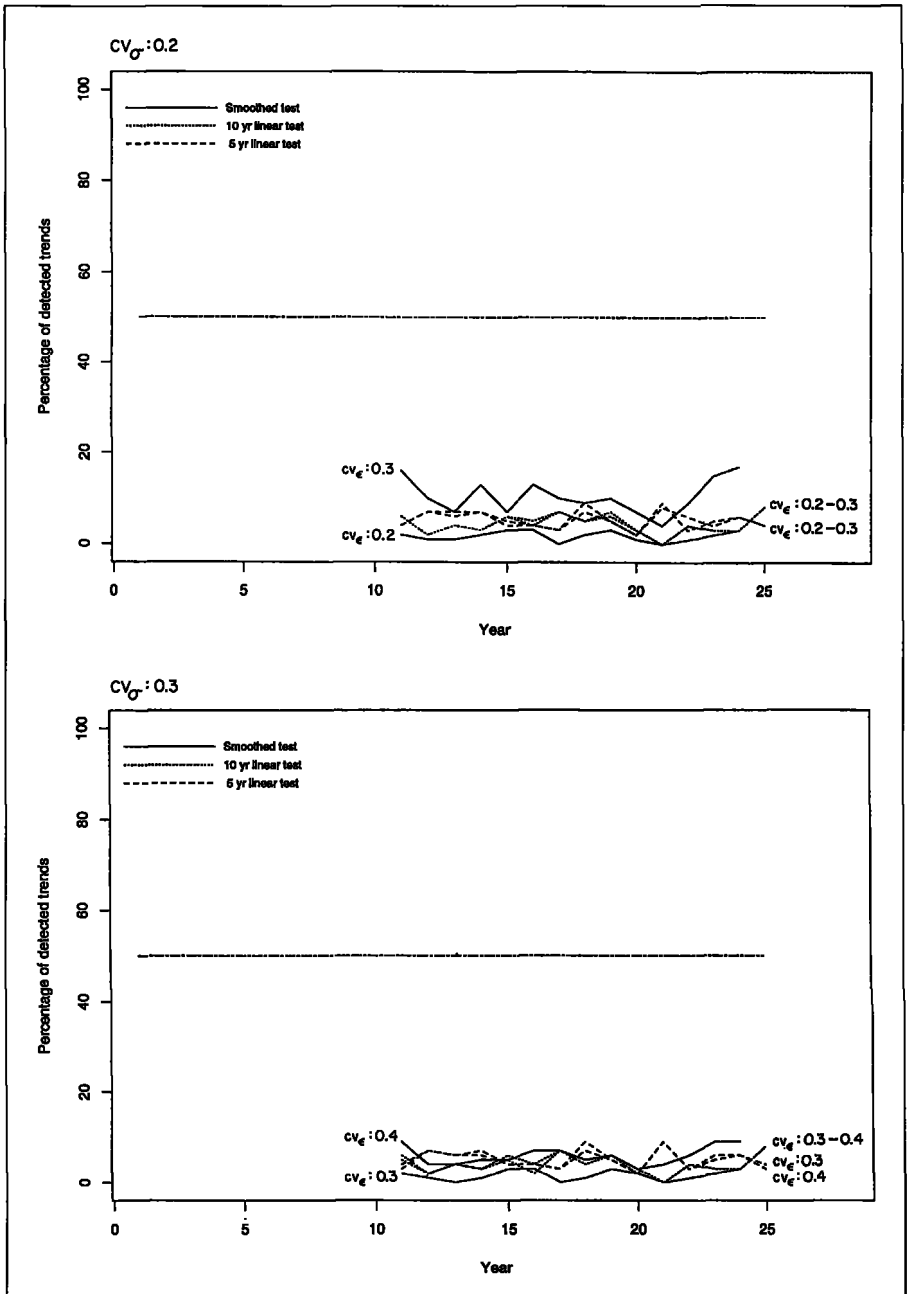

Figure 4

Percentage of trends detected in 100 simulations of the **'stable population'** scenario for two levels of precision: **(top)** CV_σ=0.2, and **(bottom)** CV_σ=0.3. The two lines for each smoothed test represent different levels of interannual variation, CV_ε. Broken line (– – –) represents underlying trend in the population on an arbitrary scale.

comparison with <60% for the linear test. In the recovery phase of the trajectory (starting when the tests cover the period between years 12 and 22), the linear test improves its performance relative to the smoothed test. In absolute terms, the performance of both tests seems to be poor between years 20 and 23, but this is the result of the small difference between the first and last year of the period included in the moving timeframe. After that, the 10 yr linear test performs almost as well as the smoothed test, as the result of an underlying trend that can be approximated well by a linear model. The 5 yr linear test does not perform well, although it detects the decline more frequently than the 10 yr test. In spite of the improvement in the last part of the trajectory, both linear and smoothed tests reach a maximum of only ~70% of significant trends. The performance of all tests degrades quickly for higher amounts of variability; the maximum percentage of detections for the smoothed test falls to around 40% for $CV_\sigma=CV_\varepsilon=0.3$.

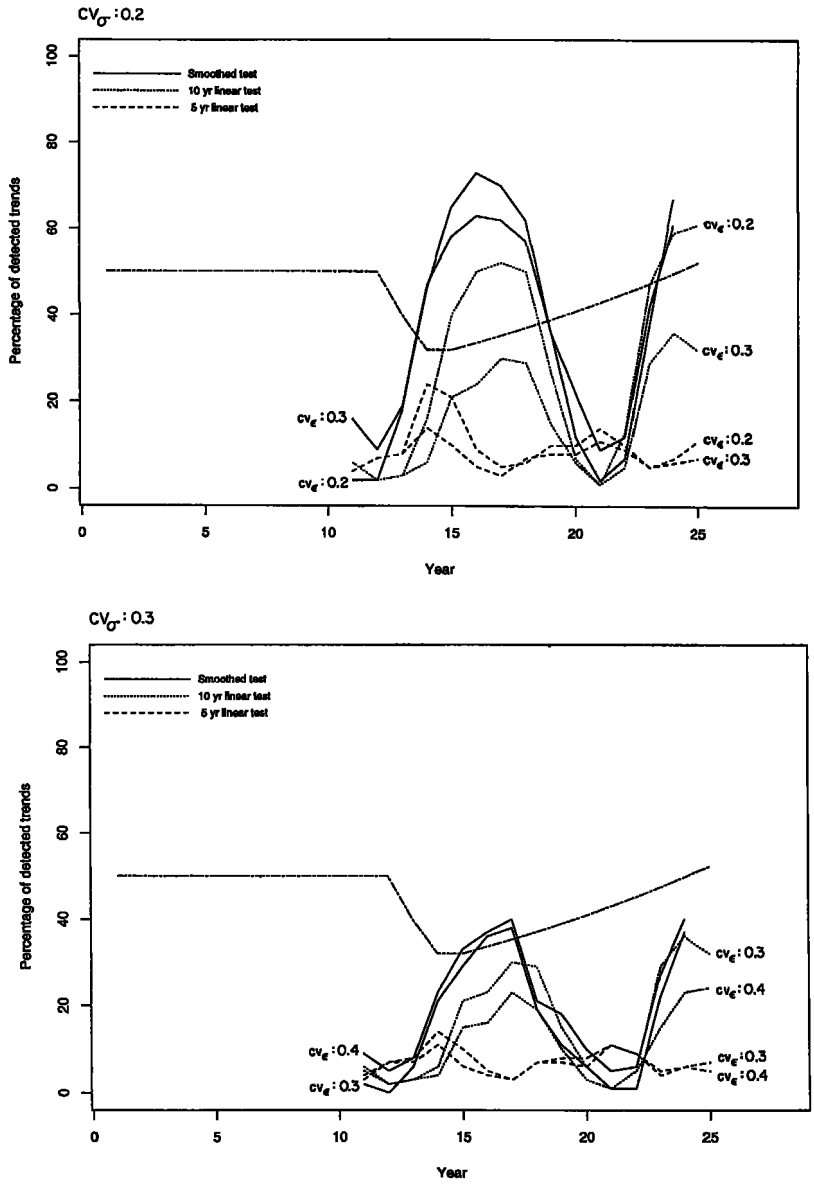

Figure 5

Percentage of trends detected in 100 simulations of the **'rapid decline'** scenario for two levels of precision: **(top)** $CV_\sigma=0.2$, and **(bottom)** $CV_\sigma=0.3$. The two solid lines for each smoothed test represent different levels of interannual variation, CV_ε. Broken line (- - -) represents underlying trend in the population on an arbitrary scale.

In the case of the smoothed test, the number of detections increases when $CV_\sigma>CV_\varepsilon$. This result seems to be a consequence of higher Type-I error probabilities, as suggested by the higher number of detections in years 11 and 12.

Steady decline The power of both types of tests improves under this scenario relative to the previous one, due to the smoother nature of the underlying trend (Fig. 6). For $CV_\sigma=CV_\varepsilon$ the smoothed test outperforms the linear tests for all levels of variability. The percentage of significant trends detected by the smoothed test ranges from over 95% for CV's of 0.20 to ~80% for a CV of 0.30. The power of the 10 yr linear test seems to be more affected by increasing variability in the estimates, falling to ~50% detections for CV=0.40. For $CV_\sigma>CV_\varepsilon$, the power of the smoothed test seems to increase although, as before, this is probably the result of greater Type-I error rates. The 5 yr linear tests show low power for all levels of variability.

Steady cycle The results of this set of simulations are very similar to those from the 'rapid decline' scenario (Fig. 7). Overall performance for both tests im-

proves relative to that scenario, due to the smoother underlying trend, reaching a maximum of 90% for the smoothed test for CV=0.2. This performance falls rapidly, as the amount of variability increases, to a maximum of slightly over 20% when CV=0.40. For $CV_\sigma > CV_\varepsilon$, the smoothed test again shows an apparent increase in power related to high Type-I error probabilities. Once more, the 5 yr linear test shows very low power throughout the series.

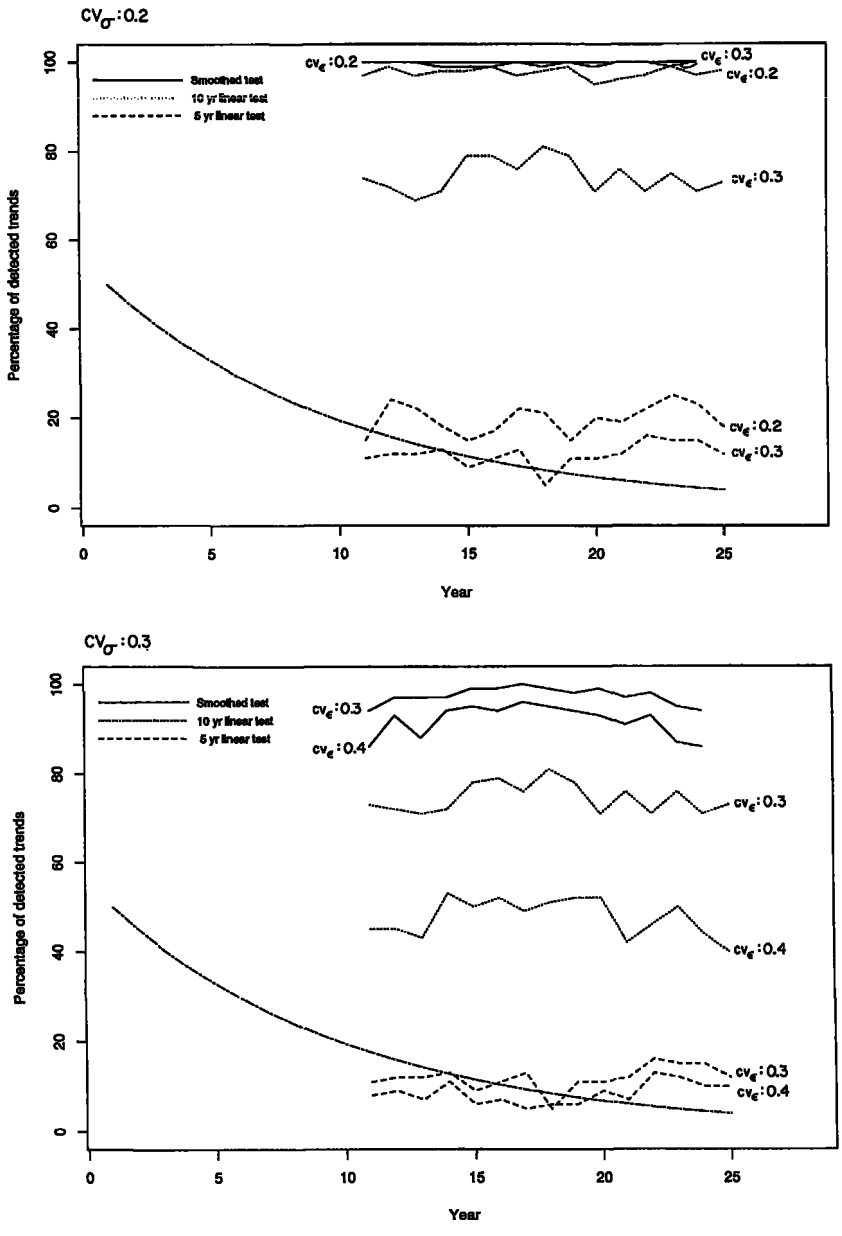

Figure 6

Percentage of trends detected in 100 simulations of the **'steady decline'** scenario for two levels of precision: **(top)** CV_σ=0.2, and **(bottom)** CV_σ=0.3. The two solid lines for each smoothed test represent different levels of interannual variation, CV_ε. Broken line (- - -) represents underlying trend in the population on an arbitrary scale.

Ratio between estimates

The purpose of this analysis is to assess sensitivity of the estimated rates of change, obtained by smoothing or linear procedures, to departures from linearity of the underlying trend. It also measures the ability of the procedures to reconstruct the true underlying population trajectory. The results of this analysis are shown for each of the simulated trajectories in Fig. 8. Results indicate that the estimated rates of change obtained through the smoothing procedures are better, in terms of the average absolute error, than estimates obtained by any of the linear methods, even when the simulated trend is linear ('stable' scenario, Fig. 8, page 192). Estimates from the 5 yr linear regressions are consistently worse than estimates from the 10 yr regressions, as expected due to the greater number of points on which the latter is based. The only exception is for the scenario with cyclic fluctuations, where the 10 yr linear estimates are poorer than the 5 yr estimates. This is a consequence of the period of the sinusoidal cycle in the underlying population that can be better approximated by a linear model over a short period of time. For longer periods in the cycle, 10 yr linear estimate should improve, as suggested by its performance in the scenario with an exponential decline. Results for the smoothed procedure are consistent over the sets of scenarios, indicating its robustness to departures from linearity in the population trajectory.

Correlation between smoothed estimates

Results from this analysis are shown in Figs. 9 and 10 (pages 193, 194), which indicate that the correlation

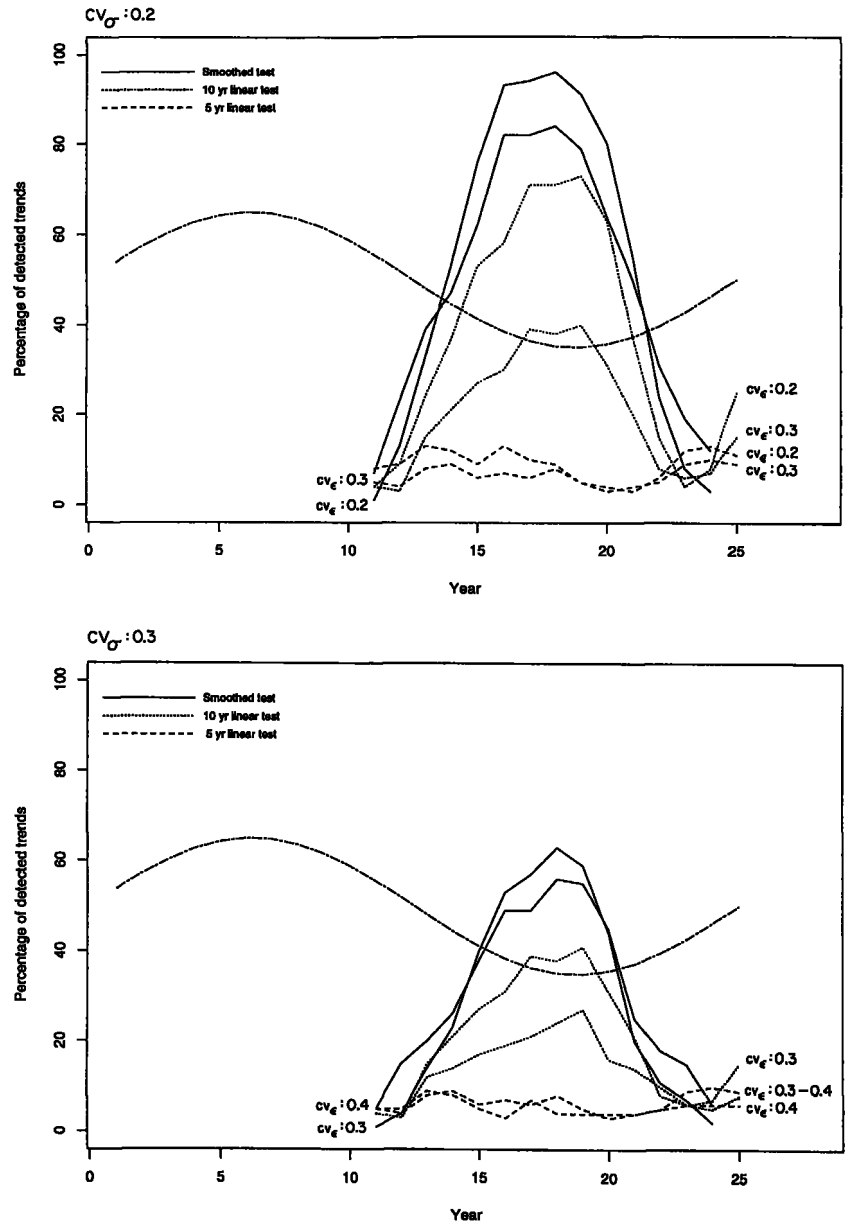

Figure 7

Percentage of trends detected in 100 simulations of the **'steady cycle'** scenario for two levels of precision: **(top)** CV_σ=0.2, and **(bottom)** CV_σ=0.3. The two solid lines for each smoothed test represent different levels of interannual variation, CV_ϵ. Broken line (– – –) represents underlying trend in the population on an arbitrary scale.

declines rapidly as distance between the estimates increases, approaching very low values as the separation between estimates exceeds 4 yr. This suggests that the validity of the test will not be seriously compromised by the correlation induced from the smoothing procedure, when the comparison is carried out on estimates that are separated by at least 4 yr.

A characteristic of the smoothed test is that, while the smoothing procedure induces a correlation between estimates, the correlation across years between fixed percentiles of the distribution of smoothed estimates is lower. This is a result of the (implicit) sorting of estimates that removes some of the dependency. To illustrate this, the average correlation between smoothed estimates obtained before sorting is also shown in Figs. 9 and 10, suggesting that the reduction in correlation due to sorting is ~30% for estimates close in time.

Discussion

In summary, the two types of tests represent different ways of looking at the data, and a comparison between them based on the same criteria only partially reflects these differences. The smoothed test provides more information since it allows a graphic comparison between individual estimates. It is based in a more robust assumption about the underlying trend of the population, since it assumes only that the change has been smooth over a short time-period. It also incorporates the lack of precision of the estimates in a much more effective way than the linear test.

In some cases, however, the linear test might be more suitable. For example, if the amount of interannual variation is low relative to the precision of the estimates, or if changes are closely approximated by a linear function, the linear test should perform better. Also, if large changes in population size occur over a very short time-period, smoothing the series will tend to underestimate the rate of change.

Despite limitations of the comparison, the results from simulations indicated that the smoothed test outperforms the linear test in most situations. The only exception is the tendency to detect spurious trends

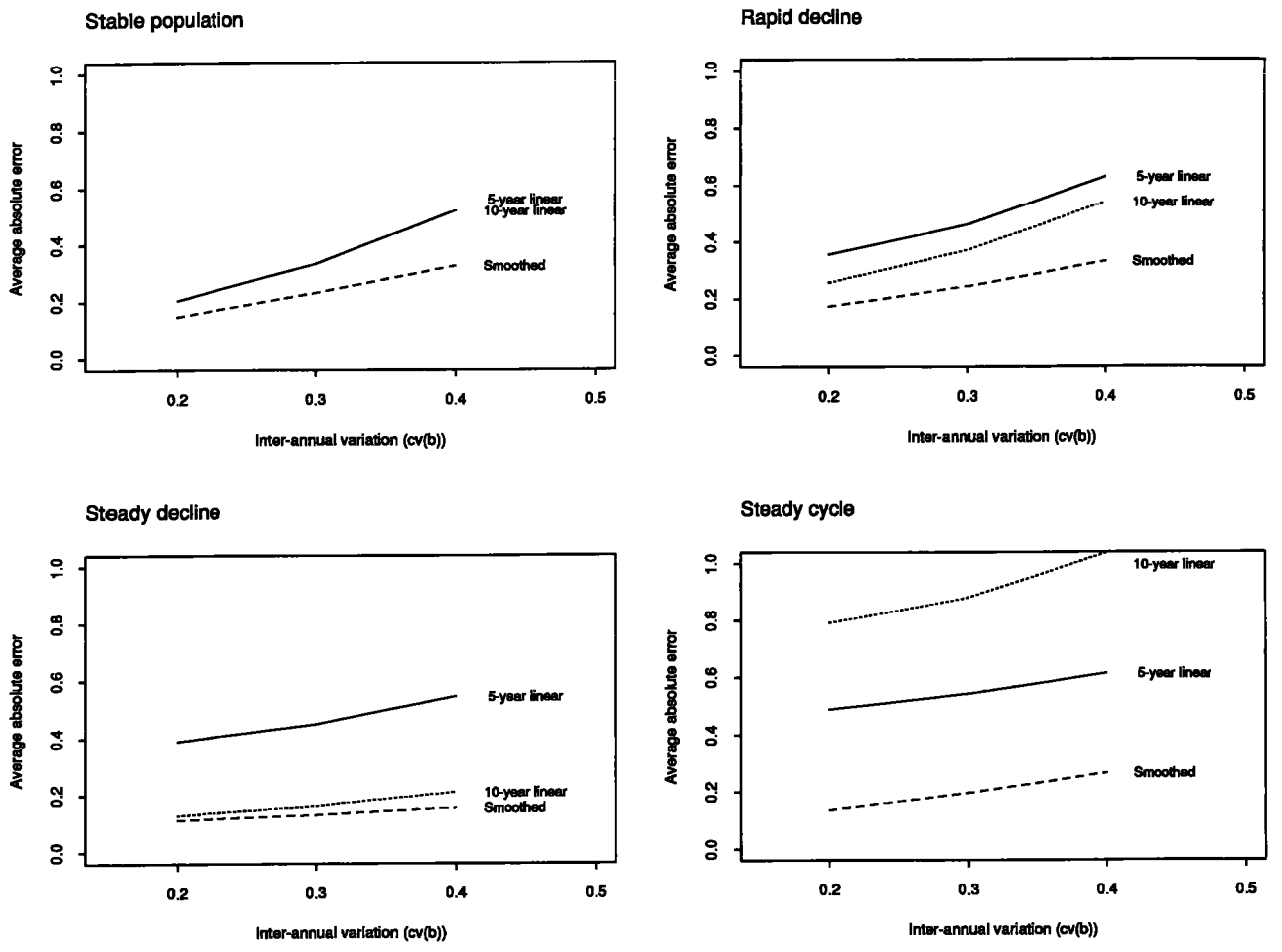

Figure 8

Average absolute error as a function of different degrees of interannual variation, represented by CV_ε, for each of the four scenarios simulated.

when the amount of interannual variability exceeds precision of the estimates. In other words, in the inevitable trade-off between Type-I and Type-II errors, the smoothed test has lower Type-II error rates at the expense of an increase in the Type-I error rate. From the management point of view, this is a safer compromise than the one posed by the linear test, since the probability of failing to detect a significant trend is smaller with the smoothed test. The increase in Type-I error rates can be related to the amount of smoothing done by the particular algorithm chosen. An algorithm that would smooth the estimates more would be less prone to this problem, but it would have less power to detect trends in the estimates in certain situations. Such an algorithm would also induce more correlation between smoothed estimates, and the separation in time between them would have to be greater in order not to compromise validity of the comparisons. An alternative would be a smoothing algorithm that can adaptively change the amount of smoothing done on the estimates, either by cross-validation techniques or by controlling the amount of smoothing through incorporating auxiliary information, such as birth and death rates, in the procedure; in other words, by building a model of the population dynamics.

Acknowledgments

I would like to thank Bill Bayliff, Steve Buckland, Martin Hall, and Tim Gerrodette for their valuable comments.

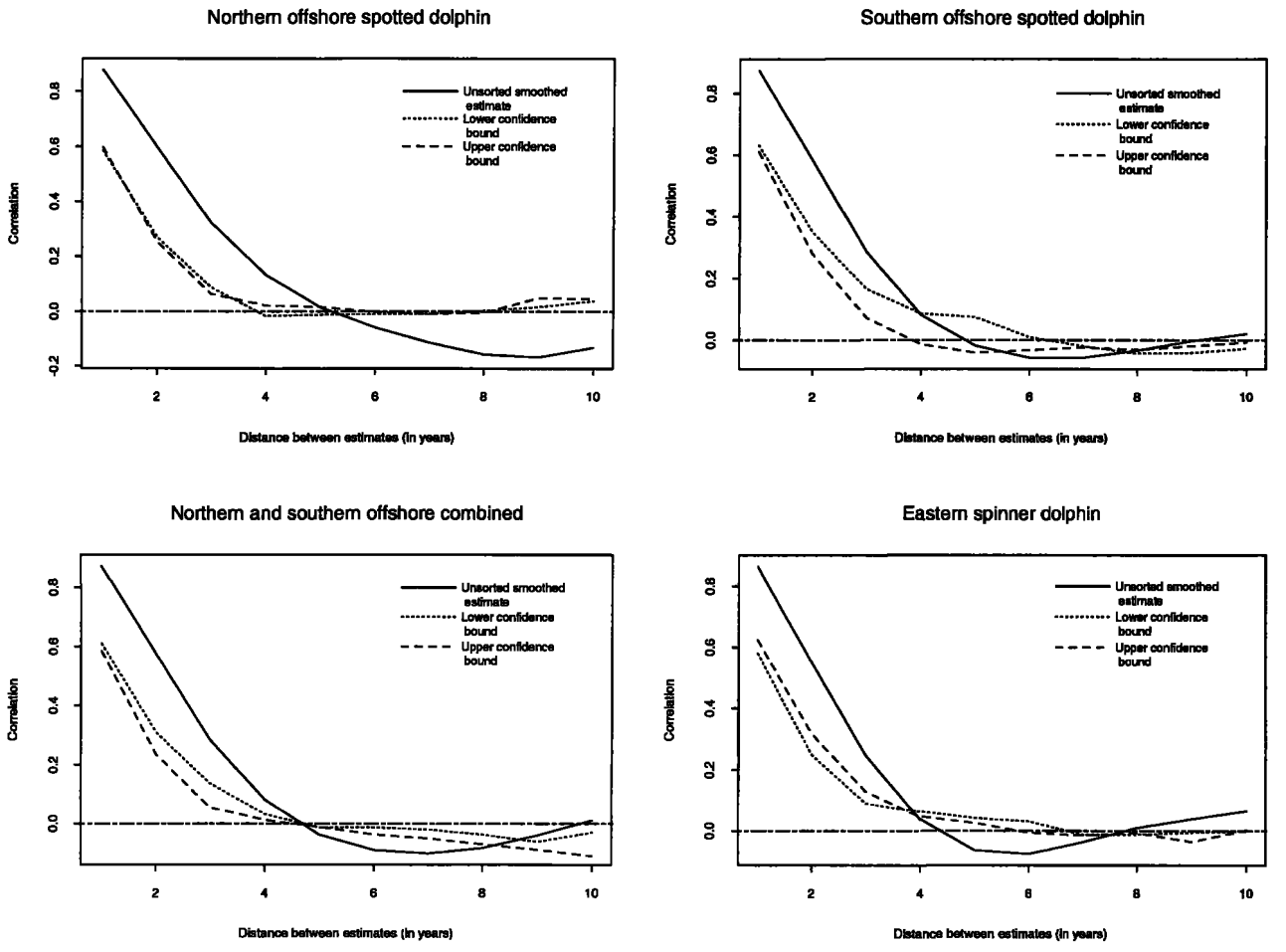

Figure 9

Correlation between smoothed estimates as a function of the separation in years between estimates for four stocks of dolphin in the eastern Pacific Ocean: Offshore spotted *Stenella attenuata* and spinner dolphin *Stenella longirostris*.

Citations

Anganuzzi, A.A., & S.T. Buckland
1989 Reducing bias in estimated trends from dolphin abundance indices derived from tuna vessel data. Rep. Int. Whaling Comm. 39:323–334.

Anganuzzi, A.A., S.T. Buckland, & K.L. Cattanach
1991 Relative abundance of dolphins associated with tuna in the eastern tropical Pacific, estimated from tuna vessel sightings data for 1988 and 1989. Rep. Int. Whaling Comm. 41.

Anganuzzi, A.A., K.L. Cattanach, & S.T. Buckland
1992 Relative abundance of dolphins associated with tuna in the eastern tropical Pacific in 1990 and trends since 1975, estimated from tuna vessel sightings data. Rep. Int. Whaling Comm. 42:541–546.

Buckland, S.T.
1984 Monte Carlo confidence intervals. Biometrics 40:811–817.

Buckland, S.T., & A.A. Anganuzzi
1988 Trends in abundance of dolphins associated with tuna in the eastern tropical Pacific. Rep. Int. Whaling Comm. 38:411–437.

Buckland, S.T., K.L. Cattanach, & A.A. Anganuzzi
1992 Estimating trends in abundance of dolphins associated with tuna in the eastern tropical Pacific Ocean, using sightings data collected on commercial tuna vessels. Fish. Bull., U.S. 90:1–12.

Edwards, E.F., & P.C. Perkins
1992 Power to detect linear trends in dolphin abundance: Estimates from tuna-vessel observer data, 1975–89. Fish. Bull., U.S. 90:625–631.

Gerrodette, T.
1987 A power analysis for detecting trends. Ecology 68:1364–1372.

IWC (International Whaling Commission)
1992 Report of the sub-committee on small cetaceans. Rep. Int. Whaling Comm. 42:178–228.

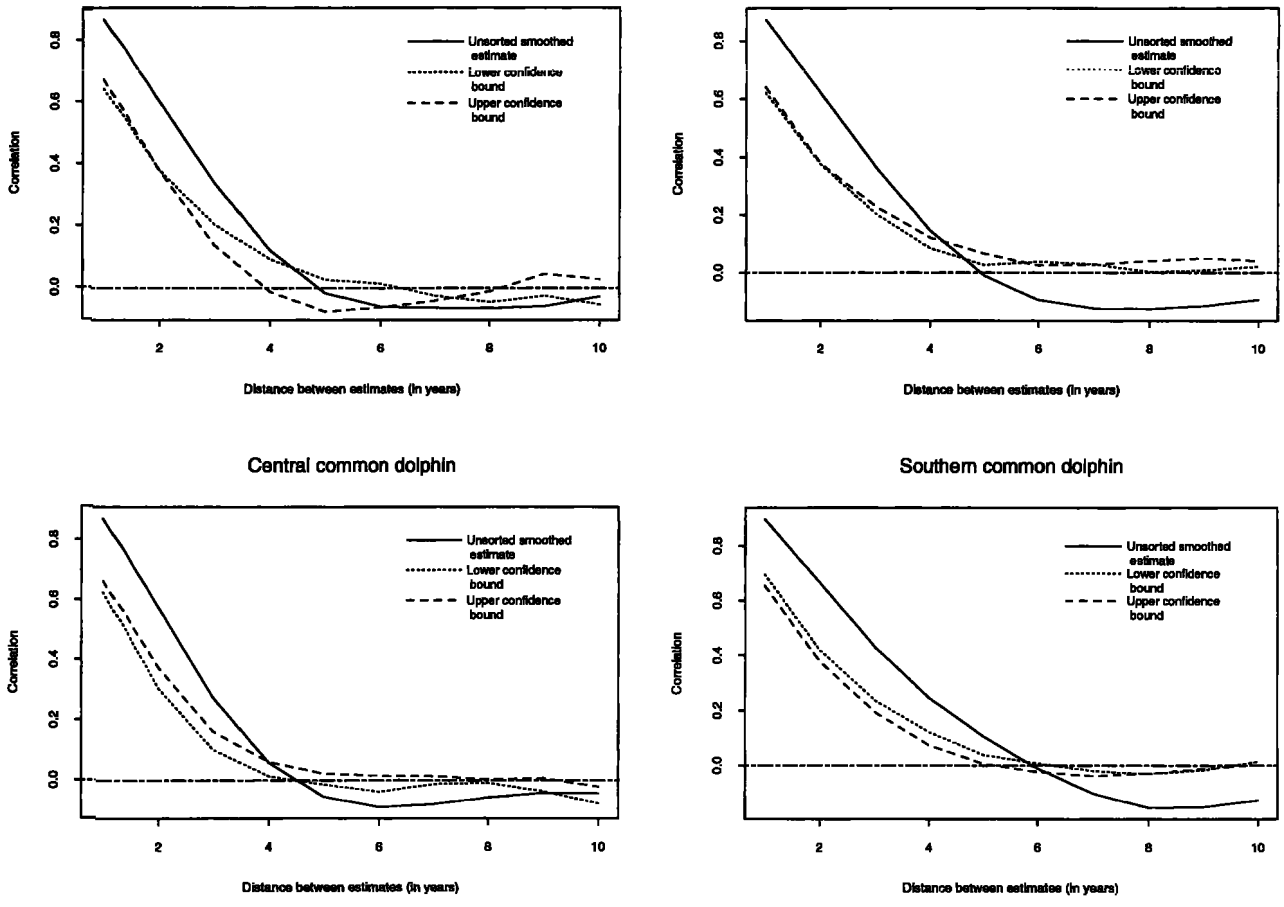

Figure 10

Correlation between smoothed estimates as a function of the separation in years between estimates of four stocks of dolphin in the eastern Pacific Ocean: Spinner dolphin *Stenella longirostris* and common dolphin *Delphinus delphis*.

Velleman, P.F., & D.C. Hoaglin

1981 Applications, basics and computing of exploratory data analysis. Duxbury Press, London.

Wade, P.R., & T. Gerrodette

1992 Estimates of dolphin abundance in the tropical Pacific: Preliminary analysis of five years of data. Rep. Int. Whaling Comm. 42:533–539.