Power to detect linear trends in dolphin abundance: Estimates from tuna-vessel observer data, 1975-89

Elizabeth F. Edwards
Peter C. Perkins
Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA
P.O. Box 271, La Jolla, California 92038-0271

Trends in abundance of dolphin stocks affected by the tuna purse-seine fishery in the eastern tropical Pacific Ocean (ETP) are of intense interest to a number of organizations concerned about the stocks' continued survival (Hammond and Laake 1983, Gerrodette 1987, Holt et al. 1987, Buckland and Anganuzzi 1988, Anganuzzi and Buckland 1989, Anganuzzi et al. 1991). The most straightforward method for estimating such trends is linear-regression analysis of relative abundance indices across time (e.g., Anganuzzi and Buckland 1989). Such abundance indices can be derived from data collected by observers aboard the tuna vessels (Buckland and Anganuzzi 1988, Anganuzzi and Buckland 1989, Anganuzzi et al. 1991). Linear trends in abundance over successive 5-year periods have been reported by Buckland and Anganuzzi (1988), Anganuzzi and Buckland (1989), and Anganuzzi et al. (1991).

Power analysis provides a method to quantify the probability of not detecting low rates of change in abundance over a specified time-period. It also provides a method, in cases where no statistically-significant trends are apparent, for determining the steepness of change necessary for its statistical detection given observed variability in the data, i.e., detectable trend (Gerrodette 1987, Peterman 1990). We use power analysis here to assess the efficacy of weighted linear-regression analysis for estimating linear trends in abundance of eight stocks of ETP dolphins. While it is instructive to evaluate the power of conclusions about observed trends, it is perhaps even more important to determine the magnitude of change required for detection of a trend, given observed variability in the dolphin abundance estimates. Therefore, we also calculate detectable trends, in addition to power of observed trends.

We present here estimates of observed trend, power to detect trends, and detectable trends for eight stocks of ETP dolphins, over time-series of 5, 8, and 10-years, assuming a two-sided hypothesis with $\alpha=0.10$, using the noncentral t-distribution for the alternative hypothesis. Detectable trends were estimated assuming Type I (α) and Type II (β) error levels equal 0.10. We estimated power and detectable trends for all three sets of time-series to determine how much improvement might be expected by increasing the number of years included in the trend estimate. We did not include longer time-series, as it is unlikely that even a population with reproductive and individual growth rates as relatively slow as ETP dolphins would follow a linear trend for more than a decade, if that long.

Methods

Relative abundance indices and their associated bootstrap standard errors for eight stocks of ETP dolphins during the years 1975–89 (Table 1, Fig. 1) formed the database for the regression analyses presented here. Indices and standard errors for 1975–87 were taken from Anganuzzi and Buckland (1989), and for 1988 and 1989, from Anganuzzi et al. (1991). The eight dolphin stocks included northern offshore and southern offshore stocks of the pantropical spotted dolphin *Stenella attenuata*, the eastern spinner dolphin *Stenella longirostris orientalis*, northern and southern stocks of whitebelly spinner dolphin (hybrid/intergrades between *Stenella l. orientalis* and *Stenella l. longirostris* [Perrin 1990]), and northern, central and southern stocks of the common dolphin *Delphinus delphis*.

Observed trends

We estimated linear trends in relative abundance for each of the eight stocks over sequential series of 5, 8, and 10 years, using standard weighted least-squares regression (Wilkinson 1989). The slope of the regression (b) estimates the trend in abundance. The estimated standard error of the estimated trend (s_b) indicates the variability associated with the trend estimate. Weights were the reciprocal of the square of the bootstrap standard errors (Buckland and Anganuzzi 1988). We eliminated the estimate for 1983 from all analyses, because the presence of a very strong El Niño that year caused biologically unreasonable estimates of abundance for many of the stocks, in particular for northern offshore spotted dolphins, the stock affected by the fishery in greatest numbers (Buckland and Anganuzzi 1988, Anganuzzi et al. 1991). In the absence of any objective criteria for choosing which stocks were (or were not) affected by the El Niño, we elected to treat

Manuscript accepted 1 June 1992.
Fishery Bulletin, U.S. 90:625–631 (1992).

Table 1

Estimated abundances (first row) and bootstrapped standard errors (second row) for eight stocks of dolphins affected by the tuna purse-seine fishery in the eastern tropical Pacific Ocean. Values are in thousands. Data for 1975–87 from Buckland and Anganuzzi 1988; data for 1988–89 from Anganuzzi and Buckland 1989. Dashes indicate no values reported for that year. Asterisks indicate values from 1983 omitted as anomalous. SOPS boundaries (Anganuzzi and Buckland 1989) used in estimating southern offshore spotted dolphins. See text for scientific names.

	Dolphin stocks							
	1	2	3	4	5	6	7	8
1975	3949	—	599	490	—	122	404	385
	996	—	197	181	—	51	138	227
1976	4253	574	535	1205	115	344	287	586
	908	212	190	293	62	130	80	223
1977	3828	924	514	588	203	637	466	229
	751	332	152	163	87	200	90	162
1978	3212	584	395	613	65	358	329	—
	543	302	124	150	47	126	105	—
1979	2950	1040	428	366	—	650	644	—
	559	394	202	183	—	248	287	—
1980	3335	260	447	342	124	512	251	230
	582	168	112	99	86	191	85	139
1981	2536	199	255	694	124	513	111	435
	443	83	165	287	80	330	35	142
1982	2550	591	202	416	100	—	232	103
	557	180	91	132	47	—	101	90
1983	*	*	*	*	*	*	*	*
	*	*	*	*	*	*	*	*
1984	2158	244	340	253	182	—	71	—
	362	115	85	72	71	—	91	—
1985	2884	238	586	648	247	—	265	249
	352	68	124	128	64	—	105	318
1986	3165	154	584	451	—	475	169	—
	302	60	108	95	—	237	50	—
1987	2953	—	384	650	—	304	60	—
	293	—	87	105	—	123	18	—
1988	2689	79	717	484	88	323	241	253
	326	30	110	92	33	93	50	100
1989	2910	560	389	515	190	243	125	179
	275	140	71	78	69	107	24	47

1 = northern offshore spotted dolphin
2 = southern offshore spotted dolphin
3 = eastern spinner dolphin
4 = northern whitebelly spinner dolphin
5 = southern whitebelly spinner dolphin
6 = northern common dolphin
7 = central common dolphin
8 = southern common dolphin

all eight stocks similarly by eliminating the 1983 estimate.

For each series, we calculated regressions using as many data points as existed for each species for that number of years. In some cases, this resulted in as few as three data points contributing to the regression. Because we omitted data from 1983, the 10-year series contained at most nine data points. Because some years were omitted or were missing abundance indices, not all year-series comprised strictly consecutive x-values (year values).

We reexpressed the slope estimate of trend (b) in terms of a change parameter r, where

$$r = b/A_1$$

and A_1 (estimated abundance in first year of series) is calculated from the estimated slope and intercept for each year-series. For these linear regressions, the parameter r expresses the annual rate of change as a fraction of the estimated initial abundance (Gerrodette 1987). Linear regressions were calculated only for series with at least three data points.

Power

We estimated power of statistical conclusions about the significance of each slope by assuming a two-sided alternative hypothesis and using the non-central t (nct) distribution. In all cases, we assumed error levels $\alpha = \beta = 0.10$. We used a two-sided hypothesis test to be consistent with earlier estimates of 5-year trends in abundance (Buckland and Anganuzzi 1988, Anganuzzi and Buckland 1989, Anganuzzi et al. 1991).

To calculate power using the nct distribution, we utilized a series of programs (available upon request) designed to return power estimates as a function of three input variables:

$t_{\alpha,\mathrm{df}}$ = normal t statistic given a level of α and degrees of freedom,

IDF = degrees of freedom, and

$\delta = b/s_b$.

Degrees of freedom were $n-2$ where n is the number of years for which abundance estimates existed in a series. Values for b and s_b were calculated from the weighted linear regressions.

δ is the offset of the alternative distribution (the nct) standardized by the standard error of the offset. In all cases, we assumed as the alternative distribution the observed trend for a series. δ is thus the distance, expressed as standard deviation units, between the mean of the null distribution (taken here to be zero slope) and the mean of the alternative distribution (the slope estimated from regression of the data).

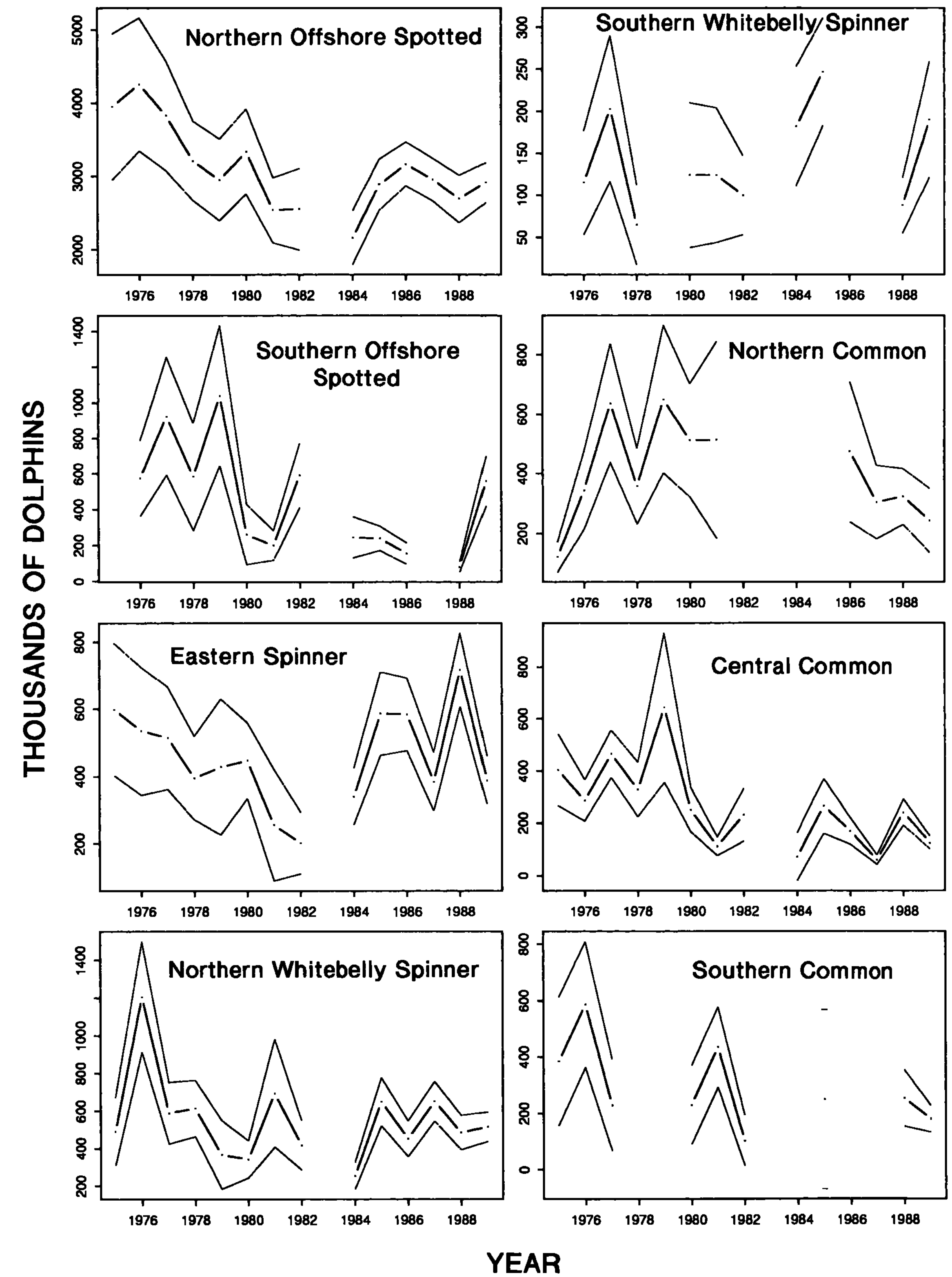

Figure 1

Estimated abundances (dashed line) of dolphins bounded by 1 bootstrap standard error (solid lines). Years for which no estimates were reported are omitted; estimates for 1983 are omitted as anomalous.

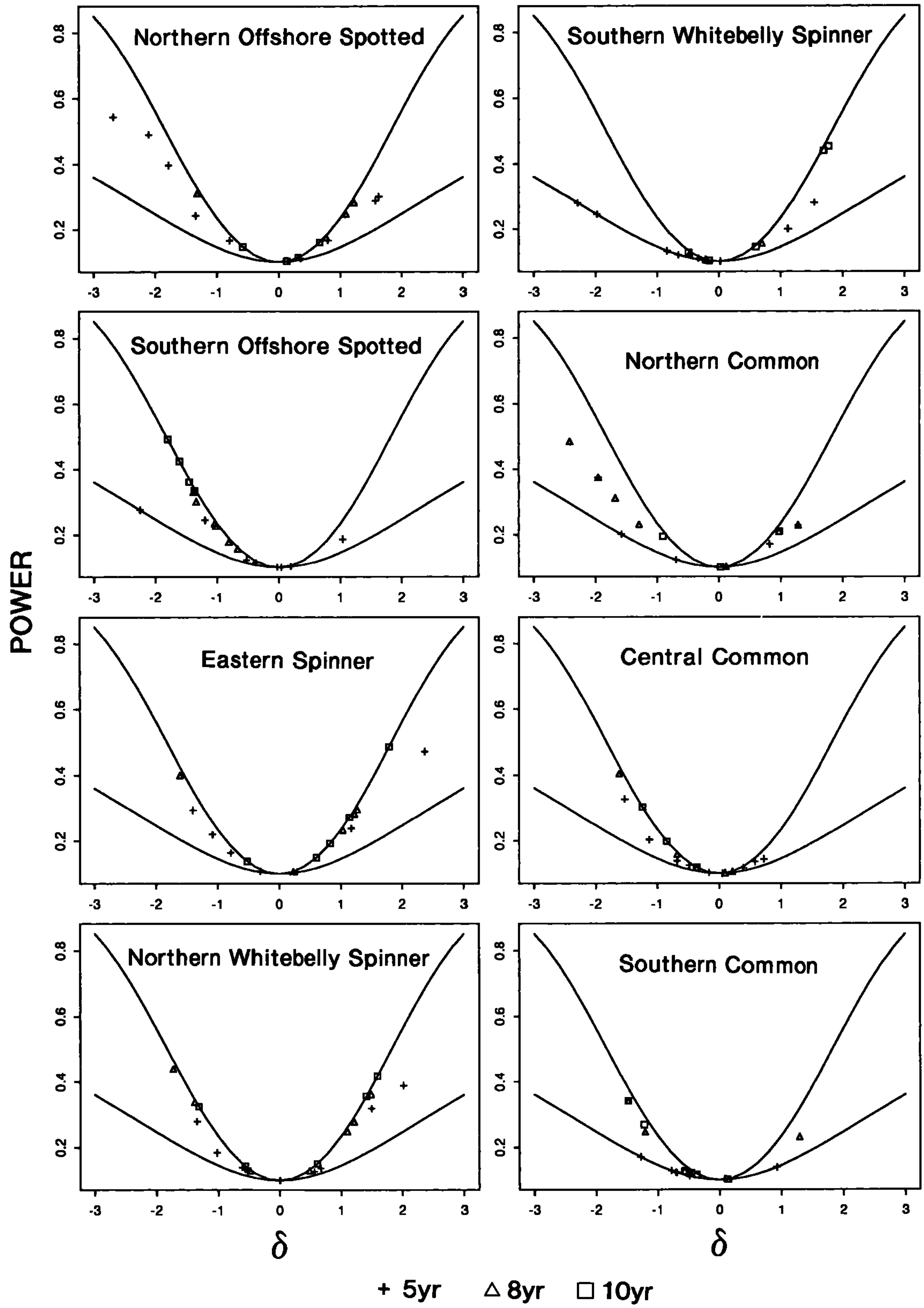

Figure 2
Power of conclusions about lack of statistical significance for estimated 5, 8, and 10-year trends in dolphin abundance. Power calculated for two-tailed noncentral t, assuming $\alpha = 0.10$. Solid lines indicate maximum (7 df) and minimum (1 df) power envelope. δ is the noncentral t parameter.

Detectable trends

We estimated detectable trends (r_d) for all data-series. All estimates of r_d assume error levels $\alpha = \beta = 0.10$. Detectable trends were estimated by determining the value of Delta ($\delta_{0.9}$) that returns a power value of 0.90 from the nct algorithm. As before, input value for IDF was $n-2$, and for s_b was the value estimated from the weighted regression. Then the value of b generating the desired power level (b_{nct}) is

$$b_{nct} = \delta_{0.9}/s_b$$

and the detectable initial trend per year is

$$r_d = b_{nct}/A_1.$$

Results

Observed trends

The majority (151/192; 79%) of the series showed no significant trend (specific data available from the authors). Of those that did, most showed decreases prior to the mid-1980s and no consistent trends since. Where population-abundance indices changed relatively regularly over time, successively longer time-series retained the same general patterns as found in shorter series. For example, observed trends were significantly negative for northern offshore spotted dolphin during the 5-year series 75/79 and 77/81, the 8-year series 75/83, 76/83, and 77/84, and the 10-year series 75/84 and 76/85. Similarly, 5-year negative trends were also reflected in 8- and 10-year series for southern offshore spotted dolphin, eastern spinner dolphin, northern whitebelly spinner dolphin, and central common dolphin.

Data were so sparse and variable for southern whitebelly spinner dolphin and southern common dolphin that little can be said about trends in these stocks. Northern common dolphin were the only species for which trends may have switched during the period of investigation (from negative during earlier years, to positive more recently); but it is obvious that here, as in the other series, the pattern in trend estimates is simply a function of the length of the series selected and its placement in time.

Power

Power to reject a false null hypothesis increases with increases in either or both of series length (as degrees of freedom increase) or δ (offset) (Fig. 2), but for TVOD the increases generally were not sufficient to be of practical use. Where no significant trends (slopes) were found, power to detect a false null hypothesis was low, averaging 20–30% in most cases and never exceeding 60%. Power for each test was small because the alternative hypothesis for these power calculations was taken to be the observed slope, which was usually fairly small, and also because scatter around the regression line tended to be large. Therefore the null and alternative distributions overlapped considerably. The low power of these tests simply means that if the true slope equaled the observed slope, the power to distinguish the true slope from a slope of zero (i.e., no change in abundance) would be quite small in most cases.

Detectable trends

The range of detectable trends decreased rapidly with increasing series-length in all cases (Fig. 3), as this increases the degrees of freedom (number of data points). The decrease is misleading in most cases, however. Although the improvement in ability to detect smaller trends with longer time-series appears dramatic, in most cases even the smallest detectable trends are still much too large to be of use.

Even with as many as 10 years of data in a series, linear trends less than about 10% per year could be detected consistently only for northern offshore spotted dolphin. For all other stocks, trends of at least 15–20% per year would be required to produce a significant result (Fig. 2). Series lengths would have to be such that populations more than doubled or decreased to zero in order for the change to be statistically detectable. This would require series lengths of at least 10 years.

In many cases, where significant trends were found, these trends were of lesser magnitude than the estimated detectable trend. This occurs because the estimated detectable trend is the expected value of the alternative distribution. Any trend value which falls below this expected value, but which also falls above the Type-I error limit for the null distribution, will be assumed significantly different from the null even though the trend could actually belong to either distribution. For example, if the Type-I error limit for the null distribution occurs at a trend value of 0.75 (i.e., if the cut-off point for values assumed to belong to the null distribution is 0.75), and the expected value (i.e., the mean) for the alternative distribution falls at 0.85, any trend value within the range 0.75–0.85 will be assigned to the alternative distribution even though it is smaller than the expected value of the alternative distribution. In practice for the ETP data, this effect is unimportant compared with the overall problem of high variability obscuring the possibility of detecting managerially-relevant trends in abundance (i.e., the

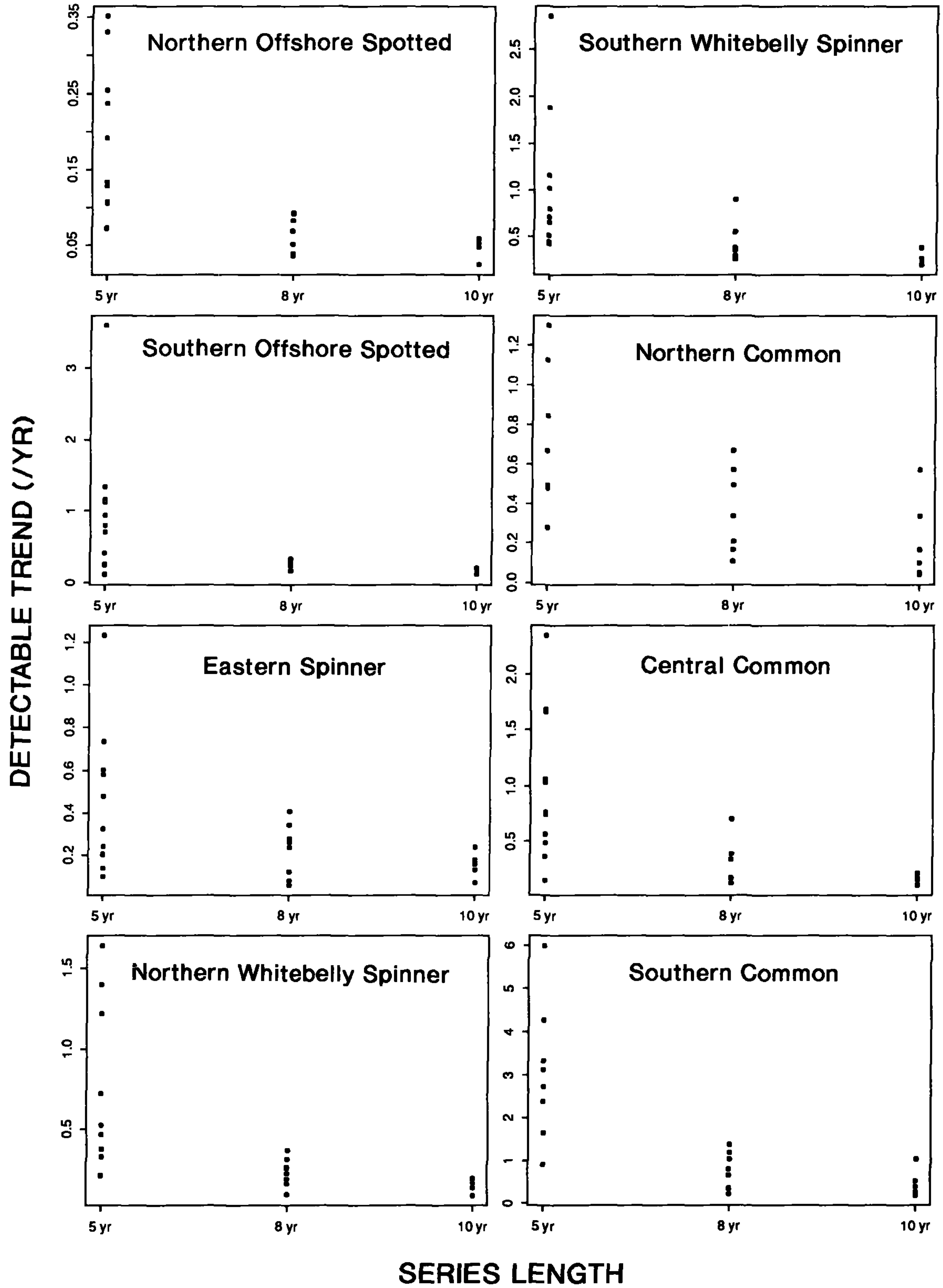

Figure 3
Estimated detectable trends for successive 5, 8, and 10-year series of tuna vessel observer data.

spread and subsequent overlap between null and alternative distributions is so great that differences between the cut-off point in the null and the mean of the alternative is effectively inconsequential).

Discussion

Given the observed variability in TVOD, it does not appear that linear trends in dolphin abundance estimated from these data with these techniques will be detectable at any levels practical for management purposes, except possibly for the stock of northern spotted dolphin. Even for this stock, survey series at least a decade in length are required.

Our estimates of 5-year trends agree with previously published estimates derived by the same methods, where 1983 has been omitted from the analyses (Anganuzzi and Buckland 1989). Our results imply that for time-periods as short as 5 years, considerably larger trends than those observed would be necessary to produce estimates of significant change with any reasonable power. Except for northern offshore spotted dolphin, this applies also to all other stocks, even for the maximum time-series tested (10 years in this study). It appears that other methods must be found to determine whether trends truly exist in dolphin abundance in the ETP. For management purposes, longer time-series must be monitored, for which linearity cannot be assumed. Other regression procedures making greater use of the precision estimates (standard errors) of the indices could have more power, but for linear analyses, at least, it is uncertain whether the increase in power could overcome the inherent variabiity and probable nonlinearities in the data.

A more effective approach to estimating trends in dolphin abundance is probably represented by the sophisticated smoothing method applied recently to these data by Buckland et al. (1992). The method reduces the relatively-scattered abundance estimates to smoothly-changing estimates of abundance, but with the advantage of producing confidence limits about the smoothed trend and generating a more biologically-reasonable result (abundance of natural populations rarely changes linearly). However, simulation experiments will be required to determine the circumstances under which the smoothed trends do, or do not, reflect accurately the true underlying dynamics of the stocks. Such simulations are currently underway, but results are as yet unavailable (Alejandro Anganuzzi, Inter-Am. Trop. Tuna Comm., La Jolla, pers. commun., July 1991).

Regardless of the results of the tests of various smoothing methods, it appears fruitless, based on the results presented here, to use linear-regression techniques to estimate trends in abundance of dolphin stocks in the ETP, even for periods as long as a decade. The power to detect ecologically (or managerially) relevant trends, given the observed variability in the data, is simply not sufficient.

Future efforts should, as suggested by Buckland et al. (1992), focus on developing or applying robust, curvilinear smoothing techniques that are reasonably responsive to the underlying processes or mechanisms controlling actual changes in dolphin abundance.

Acknowledgments

This study has benefited greatly from technical assistance by Cheryl Glick and extensive helpful discussions with Tim Gerrodette and Doug DeMaster.

Citations

Anganuzzi, A.A., and 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, and 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. 41:497–506.

Buckland, S.T., and A.A. Anganuzzi
1988 Estimated 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, and A.A. Anganuzzi
1991 Estimating trends in abundance of dolphins associated with tuna in the eastern tropical Pacific, using sighting data collected on commercial tuna vessels. Fish. Bull., U.S. 90: 1–12.

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

Hammond, P.S., and J.L. Laake
1983 Trends in abundance of dolphins involved in the purse-seine fishery for tuna in the eastern tropical Pacific Ocean, 1977–1981. Rep. Int. Whaling Comm. 33:565–588.

Holt, R.S., T. Gerrodette, and J.B. Cologne
1987 Research vessel survey design for monitoring dolphin abundance in the eastern tropical Pacific Ocean. Fish. Bull., U.S. 85:435–446.

Perrin, W.F.
1990 Subspecies of *Stenella longirostris* (Mammalia: Cetacea: Delphinidae). Proc. Biol. Soc. Wash. 103(2):453–463.

Peterman, R.M.
1990 The importance of reporting statistical power: The forest decline and acidic deposition example. Ecology 71:2024–2027.

Wilkinson, L.
1989 SYSTAT: The system for statistics. Systat, Inc., Evanston.