9264

Indices of Relative Abundance from Fish Spotter Data based on Delta-Lognormal Models

N. Chyan-huei Lo, Larry D. Jacobson, and James L. Squire

National Marine Fisheries Service, Southwest Fisheries Science Center, P.O. Box 271, La Jolla, CA 92038, USA

Lo, N. C. H., L. D. Jacobson, and J. L. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci. 49: 2515–2526.

Fish spotters are pilots in small aircraft employed by commercial fishermen to locate, identify, estimate the size of, and direct boats toward pelagic fish schools. Data describing species, location, and size of schools can be inexpensively obtained from fish spotters but are difficult to interpret statistically. We developed an index of relative abundance from fish spotter data based on extended delta-lognormal models and applied the method to data for northern anchovy (*Engraulis mordax*). In contrast with previous approaches, our method used all available data, provided an index for northern anchovy that was proportional to abundance, and explicitly modeled factors (pilots, regions, seasons, and time of day) that affected observations by fish spotters. We also included information about mixed layer depth and sea surface temperature in models for a reduced study area and found that environmental data, where available, can be used to improve estimates of relative abundance from fish spotter data. Simulation results indicated that our approach is a cost-effective way to improve biomass estimates for pelagic species like northern anchovy.

Les observateurs de poissons sont des pilotes de petits avions employés par des pêcheurs commerciaux pour localiser, identifier, évaluer la taille et diriger les bateaux vers les bancs de poissons pélagiques. Des données indiquant l'espèce, l'emplacement et la taille des bancs peuvent être obtenues à peu de frais des observateurs de poissons, mais elles sont difficiles à interpréter statistiquement. Nous avons élaboré un indice d'abondance relative à partir des données des observateurs de poissons basé sur des modèles delta-lognormaux étendus et nous avons appliqué la méthode à des données sur l'anchois du Nord (*Engraulis mordax*). Contrairement à celles utilisées précédemment, notre méthode utilise toutes les données disponibles, fournit un indice pour l'anchois du Nord qui est proportionnel à l'abondance et modélise explicitement les facteurs (pilotes, régions, saisons et heure) qui influent sur les observations effectuées par les observateurs de poissons. Nous avons également inclus de l'information sur la température en surface et en profondeur des couches mélangées dans les modèles pour une aire d'étude réduite et nous avons découvert que les données environnementales, lorsque disponibles, peuvent être utilisées pour améliorer les estimations d'abondance relative faites à partir des données fournies par les observateurs de poissons. Les résultats des simulations ont indiqué que notre méthode constitue une façon économique d'améliorer les estimations de la biomasse d'espèces pélagiques comme l'anchois du Nord.

Received March 18, 1991
Accepted June 25, 1992
(JA951)

Reçu le 18 mars 1991
Accepté le 25 juin 1992

Information about changes in abundance of fish stocks is an important part of information required for fisheries management. Catch-per-unit-effort (CPUE) data from fisheries have traditionally been used to measure changes in relative abundance, but accumulated evidence indicates that interpreting CPUE data from commercial fisheries is difficult, particularly for pelagic fish, because commercial CPUE data are affected by changes in fishing efficiency over time and tend to decline more slowly than abundance (Bannerot and Austin 1983; MacCall 1984). A number of alternative fishery-independent procedures for estimating relative abundance of fishes have been developed including scientifically designed acoustic surveys, trawl surveys, egg and larva surveys, aerial surveys, and tagging experiments (Ulltang 1977). These approaches may provide good results but are expensive. In contrast, aerial "fish spotters" are a potential source of large quantities of inexpensive information about relative abundance of pelagic fish that can be obtained from existing commercial operations. To date, fish spotter data have received relatively little attention from fishery managers.

Fish spotters are pilots in small aircraft employed by commercial fishermen to locate, identify, estimate the size of, and direct boats towards pelagic fish schools (Squire 1961; Maughan and Marmelstein 1971). Fish spotters are used in the Australian southern bluefin tuna (*Thunnus maccoyii*) fishery; the southern California purse-seine fishery for northern anchovy (*Engraulis mordax*), Pacific (chub) mackerel (*Scomber japonicus*), jack mackerel (*Trachurus symmetricus*), Pacific bonito (*Sarda chiliensis*), bluefin tuna (*Thunnus thynnus*), and Pacific sardine (*Sardinops sagax*); the South African purse-seine fishery for anchovy (*Engraulis capensis*) and pilchard (*Sardinops ocellata*); the Japanese sardine (*Sardinops melanostica*) fishery; the U.S. Atlantic menhaden (*Brevoortia tyrannus*) and swordfish (*Xiphias gladius*) fisheries, and other pelagic fisheries around the world (Squire 1961, 1972, 1983; Maughan and Marmelstein 1971; Agenbag 1980; Anonymous 1981; Williams 1981; Habib et al. 1982; Agenbag et al. 1984; Hara 1985).

It is important to emphasize the differences between data from fish spotters and data from aerial surveys conducted for scientific purposes. Methods are available for using data from scientifically designed aerial surveys, usually in connection with acoustic data, to estimate biomass or relative abundance of pelagic fishes (Cram and Hampton 1976; Rivas 1978; Hampton et al. 1979; Hara 1986; Scott et al. 1989; Hara 1990). Data collected during such surveys are relatively easy to interpret but costly. In contrast, data collected opportunistically during the search for fishable schools by fish spotters are harder to interpret but inexpensive (Anonymous 1981).

Accuracy and precision of fish spotter data are reinforced because inaccurate species identification and poor estimates of school size inconvenience fishermen who employ fish spotters. Difficulties with interpretation of fish spotter data arise, however, because of factors that are unique to fish spotter data (imprecise nature of visual biomass estimates and differences among fish spotters) and other factors (patchy distribution of pelagic fish, lack of systematic sampling design, environmental variables, differences among regions, seasons, and time of day) that hinder interpretation of other types of data (e.g. commercial CPUE) as well. Flights are conducted only when environmental conditions are good enough for fishing, but environmental conditions still affect observations by fish spotters. Changes in fish spotters employed by a commercial fishery are another important problem because changes in abundance may be confounded by changes in fish spotters who collect the data. Despite problems in interpretation, fish spotter data can be useful and cost-effective tools for management purposes (Barnes et al. 1992), particularly, as in our analysis, when adjustments are made for differences among fish spotters, regions, and other important factors.

An important advantage of fish spotter data over commercial CPUE data is absence of saturation (increased abundance but no increase in CPUE due to limited hold capacity, trip limits for catch, market demand, or processing capacity) because fish spotters can record the size and location of all schools seen. This important characteristic makes it relatively easy to construct indices of abundance based on fish spotter data that are linear measures of fish biomass (i.e. change in proportion to biomass). Another advantage is that technical improvements to equipment used by fish spotters, such as aircraft and navigational gear, have probably not substantially increased their efficiency in locating fish. Fish spotters may purchase new aircraft and better navigational or radio equipment, but schools are still located visually while flying at reduced speeds and low altitudes. In contrast, changes in commercial CPUE due to changes in fish abundance are often confounded by technical improvements to fishing gear such as nets and acoustic equipment (Kimura 1981; Jacobson et al. 1987).

We developed delta-lognormal linear models (Shimizu 1988) for fish spotter data based on the delta distribution (Aitchison and Brown 1957; Pennington 1983) and lognormal linear models (Bradu and Mundlak 1970; Kerlinger and Pedhazur 1973; McCullagh and Nelder 1983). Our approach extends delta-lognormal models and breaks new ground in fisheries because we model components of the delta distribution as functions of factors and covariates in lognormal linear models (Lambert 1992) and provide approximate variances for estimates. Lognormal linear models are similar to analysis of variance models fit using linear regression and often used by fishery scientists and managers to derive a single index of fishing effort for a group of heterogeneous vessels or to derive a single index of relative abundance for a fish stock from two or more types of data (Gulland 1956; Robson 1966; Kimura 1981, 1988; MacCall and Prager 1988; Hilborn and Walters 1992). Although we used lognormal linear models for components of the delta distribution, other linear or nonlinear models based on other statistical distributions could be used instead.

The delta distribution is often used to estimate abundance of planktonic organisms whose spatial distribution is highly contagious (Dennis and Patil 1988). Survey data for planktonic organisms typically have statistical distributions with a large fraction of "zeros" (samples in which no organisms are observed) and may be so distorted that conventional methods based on the sample mean yield inefficient estimates of abundance. The delta distribution avoids problems with contagion by treating zero and nonzero data separately; final estimates of abundance are obtained from the product of the proportion and mean for nonzero observations. The delta distribution is used in many disciplines to model processes that generate more zero observations than might be expected on the basis of distributional assumptions (Lambert 1992).

Our approach used two lognormal linear models: the first to estimate the proportion of flights in which fish were seen by fish spotters and the second to estimate density (tons per area) of fish from data for flights during which fish were seen. Analogous to the delta distribution, our index of relative abundance from fish spotter data was based on the product of these quantities (adjusted for the size of the survey area as described below).

We applied our models to data for northern anchovy off the coast of southern California. In addition, we applied an expanded version of the model to a subset of the fish spotter data and two environmental variables (sea surface temperature and mixed layer depth) to determine if environmental information could potentially be used to improve estimates. Other types of environmental data (e.g. sea state and wind speed) were not available but may have been useful. The amount of data available and models used in our analysis enabled us to obtain useful estimates from fish spotter data and to account for differences among fish spotters, regions, seasons, time of day, and other variables.

Previous analyses of fish spotter data used a simple CPUE-like index (tons sighted per block area flight or T/BAF) to measure relative abundance (Squire 1961, 1972, 1983). The T/BAF index was corrected for time of day by excluding data for day flights and corrected for differences among regions by restricting the analysis to a "core" region where the species of interest was naturally abundant. A significant advance in our analysis was that all available information was used to estimate abundance; corrections for differences between time of day, among regions, and seasons were made without excluding data.

Another important difference between our analysis and previous ones was that we made corrections for differences among fish spotters. Our records indicated consistent differences among fish spotters in mean reported weight of anchovy schools and the proportion of flights during which schools were sighted (Fig. 1). Moreover, fish spotters participating in the monitoring program changed over time so that, for example, some fish spotters participated only during the 1960's and 1970's, while others did not begin to participate until the late 1980's. If indices of abundance from fish spotter data were not corrected for differences among fish spotters, then temporal changes in abundance may have been confounded by changes in fish spotters. As shown below, effects of fish spotters were important for anchovy.

The original motivation for our work was to improve spawning biomass estimates used to manage the fishery for northern anchovy. Data for biomass estimates were reduced in recent years and precision and accuracy of biomass estimates deteriorated (Jacobson and Lo 1990). For this reason, we conducted simulation experiments to determine if our new index based on fish spotter data could be used to improve biomass estimates and management advice for northern anchovy.

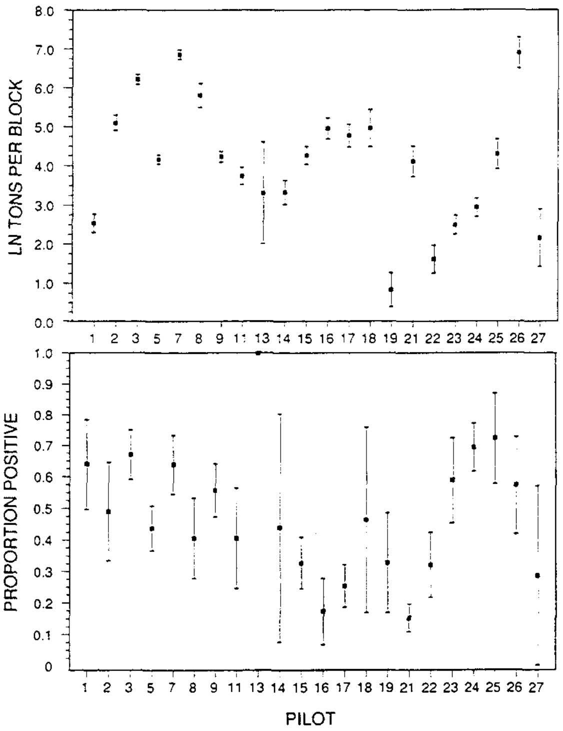

FIG. 1. Mean log tons of anchovy sighted per block (top panel) and proportion positive blocks (bottom panel) for fish spotters in region 2 (see Fig. 2) during 1963–90. Some fish spotters did not participate during the entire time period. Vertical bars show means ±2 SE.

Materials and Methods

The anchovy data used in our analyses were collected by fish spotters paid a nominal wage ($1 to $4 per hour) to participate in a pelagic resource monitoring program initiated in 1962 by the Tiburon Marine Laboratory, U.S. Fish and Wildlife Service (now Southwest Fisheries Science Center, National Marine Fisheries Service, or NMFS), in central and southern California (Squire 1961, 1972). Data were collected during the course of routine fish spotting and involved little additional effort by fish spotters. The data set (called the "complete" data set below) included records from about 16 000 flights conducted during 1963–90 (records for fish spotters with less than 20 flights were omitted). The subset of fish spotter data used with environmental data consisted of records from about 3700 flights over a smaller geographic area (records for fish spotters with less than 10 flights were omitted).

Fish spotters recorded the species (most pelagic schooling fishes of commecial importance can be identified from the air), location, time of day, and estimated weight of all fish schools encountered during flights in special flight logs (Squire 1961, 1972). Locations were specified by "blocks" which are 10′ latitude by 10′ longitude. Time of day was either day or night (fish schools are visible at night due to bioluminescence; Squire 1972). Flight log data for each sighting were checked for errors and entered into a computer database.

The area covered by fish spotters was divided into six regions (Fig. 2). During 1963–90, fish spotters collected data from the area between 27 and 37°N latitude, but anchovy were most common in the area north of region 6 (30°33′N latitude). Data for region 6 were excluded from our analysis because anchovy were seldom seen there.

Sea surface temperature and mixed layer depth were estimated for each block using data collected during routine California Cooperative Oceanic Fisheries Investigations (CalCOFI) research cruises (L. Eber, NMFS, Southwest Fisheries Science Center, P.O. Box 271, La Jolla, CA, 92038 USA, pers. comm.). CalCOFI cruises were conducted throughout the study period (with some interruptions) on a quarterly basis (Hewitt 1988). Environmental data could be obtained for regions 2 and 3 (Fig. 2) only because CalCOFI cruises have been limited to these regions since 1987. Years with no CalCOFI environmental data were omitted from our models when environmental data were included.

Sea surface temperature and mixed layer depth were treated as categorical, rather than continuous, variables because initial estimates of the functional form of the relationships between anchovy biomass and sea surface temperature or mixed layer depth were erratic when environmental data were treated as continuous variables. Although preliminary estimates of functional forms were uninterpretable, it was clear that anchovy abundance varied with sea surface temperature and mixed layer depth when the environmental data were treated as categorical variables. Sea surface temperature data were, therefore, aggregated in 2°C categories and mixed layer depth data were aggregated into 20-m categories. Intervals for the categories were chosen so that there was a sufficiently large number of observations in each category.

Models

As described above, relative abundance of northern anchovy was expressed as the product of density and a measure of area:

(1) $I = DA$

where I is the index of relative abundance for a given year (tons), D is density of anchovy (tons per block), and A is the area (blocks) covered by fish spotters. We assumed that fish spotters flew over an area that was at least as large as the area occupied by the anchovy stock in each year. Units for the index (I) are tons of anchovy sighted by fish spotters.

Density of anchovy (D) for each year was

(2) $D = dP$

where d is a standardized measure of anchovy density (tons per block) for positive flights (flights during which anchovy were seen) and P is a standardized measure of the proportion of blocks that were covered by positive flights (referred to as proportion positive). As described above, density of anchovy (D) was calculated from the product of density in positive blocks (d) and proportion positive (P) in order to avoid problems that arise from including a large number of zeros. Moreover, area of the stock (A), density for positive flights (d), and proportion positive (P) are all useful measures of relative abundance for pelagic species (MacCall 1990; Mangel and Smith 1990; Smith 1990), while area (A) and proportion positive (P) provide information about spatial distribution.

The complete data set was used to obtain standardized estimates of anchovy density for positive blocks (d) and proportion positive (P) for each year using lognormal linear models

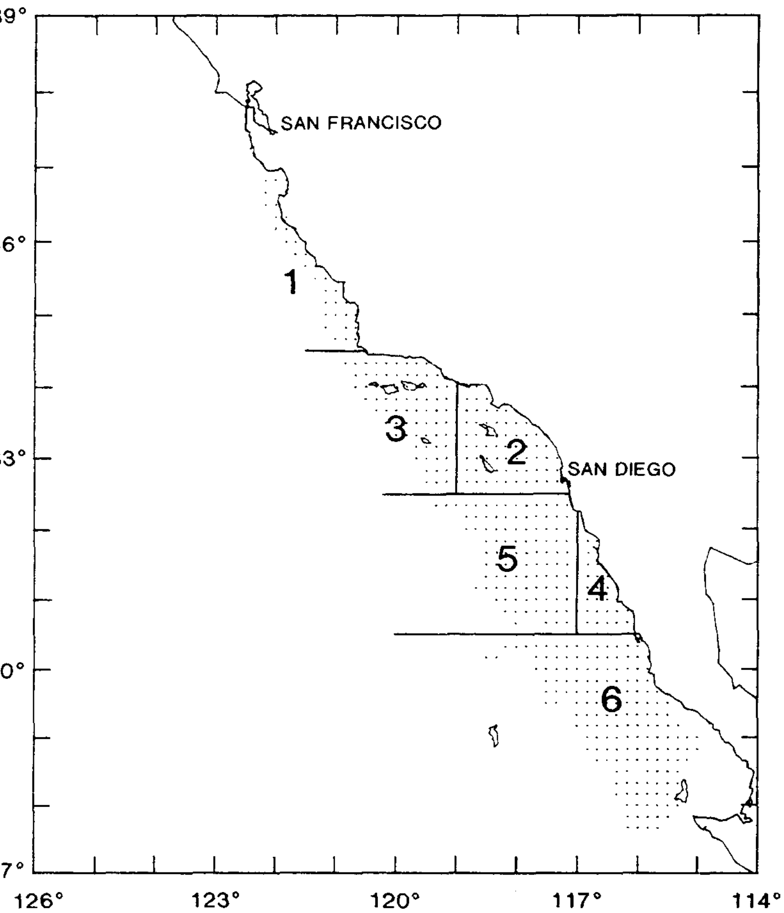

FIG. 2. Study area, regions, and blocks covered by fish spotters in 1989. Regions are outlined and denoted by numbers. Blocks are denoted by dots.

described below. Five factors were treated as main effects: years, fish spotters, regions, seasons, and time of day (day or night). Factors were allowed to cross in the model so that any combination of years, fish spotters, regions, seasons, and time of day could be accommodated. In addition to main effects, first-order interactions between regions, seasons, and time of day were included in the models. Other first-order and higher order interactions were not included to avoid overparameterization of models and problems with parameter estimation as well as to limit the amount of computer memory required.

Data used to estimate density of anchovy for positive flights (d) were aggregated by flight. The model was

$$\ln(\boldsymbol{d}) = \boldsymbol{X\beta} + \boldsymbol{Z\alpha} + \boldsymbol{\epsilon}$$

where $\boldsymbol{d}$ is a vector for observations, $\boldsymbol{X}$ is the design matrix for main effects and $\boldsymbol{\beta}$ is the parameter vector for main effects, $\boldsymbol{Z}$ is the design matrix for interactions and $\boldsymbol{\alpha}$ is the parameter vector for interactions, and $\boldsymbol{\epsilon}$ is a vector of independent normally distributed errors with expectation zero and variance σ^2.

For record k, we have

$$(3) \quad \ln(d_k) = \ln(y_k/b_k) = f(\boldsymbol{\beta}, \boldsymbol{\alpha}, X_k) + \epsilon_k$$

and

$$(4) \quad f(\boldsymbol{\beta}, \boldsymbol{\alpha}, X_k) = \beta_0 + \sum_{i=1}^{5} \sum_{w=1}^{J_i} \beta_{iw} X_{iwk} + \sum_{i=1}^{5} \sum_{w=1}^{J_i} \sum_{L>i}^{5} \sum_{v=1}^{J_L} \alpha_{iwLv} X_{iwk} X_{Lvk}$$

where d_k is tons of anchovy sighted per block for positive flight k, b_k is the *total number of blocks* searched during positive flight k (not the number of positive blocks), y_k is total tons of anchovy sighted, and ϵ_k is a normally distributed error with expectation zero and variance σ^2. The number of positive blocks (b_k) was incremented each time a fish spotter left one block and entered another so that an individual block searched twice during a flight would be counted twice.

Subscripts i and L in (4) index main effects (years, fish spotters, regions, seasons, and time of day), w and v index factor levels (e.g. spring, summer, and autumn for seasonal effects), J_i is the number of levels for factor i, J_L is the number of levels for factor L, and k is used to index flights or records. The coefficient β_0 is the mean log density ($\ln(d_k)$) of anchovy sighted

TABLE 1. Symbols and descriptions for main effect and interaction parameters in lognormal linear models. The symbols "w" and "v" are used to subscript factor levels. The symbols "α" and "β" are used in models for density, while "γ" and "τ" are used in models for proportion positive.

Parameter	Description
$\boldsymbol{\alpha}$ or $\boldsymbol{\gamma}$	Vector for interaction parameters
$\boldsymbol{\beta}$ or $\boldsymbol{\tau}$	Vector for main effect parameters
β_0 or τ_0	Mean log density ($\ln(d_k)$) or mean log proportion positive ($\ln(P_k + 1)$) of anchovy sighted for a reference year (1963), region (region 2), season (winter, January–March), time of day (nighttime), fish spotter (fish spotter 1), and, if applicable, set of environmental conditions
β_{1w} or τ_{1w}	Year effects; w = 1–27 for years 1964–90
β_{2w} or τ_{2w}	Region effects; w = 1–4 for regions 1, 3, 4, and 5
β_{3w} or τ_{3w}	Seasonal effects; w = 1–3 for spring (April–June), summer (July–September), and autumn (October–December)
β_{4w} or τ_{4w}	Day/night effect; w = 1 for daytime
β_{5w} or τ_{5w}	Fish spotter effects; w = 1–21 for fish spotters numbered 2–27 (some fish spotters with less than 20 flights excluded)
β_{6w} or τ_{6w}	Temperature effects; w = 1–4 for temperatures $\leq$12°C, >12 and $\leq$14°C, >16 and $\leq$18°C, and >18°C
β_{7w} or τ_{7w}	Mixed layer depth effects; w = 1–3 for mixed layer depths $\leq$20 m, >20 and $\leq$40 m, and >60 m
α_{2w3v} or τ_{2w3v}	Interaction between regions and seasons; w = 1–4 and v = 1–3
α_{3w4v} or τ_{3w4v}	Interaction between seasons and time of a day; w = 1–3 and v = 1
α_{2w4v} or τ_{2w4v}	Interaction between regions and time of a day; w = 1–4 and v = 1

for a reference year (1963), region (region 2), season (winter, January–March), time of day (nighttime), and fish spotter (fish spotter 1). The vectors $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ hold parameters for main effects and interactions (see Table 1 for details). The vector X_k holds the dummy variables X_{iwk} (Weisberg 1980) for main effect i, level w, and positive flight k. The number of parameters and dummy variables for each main effect was equal to the number of categories minus 1. There were, for example, three parameters for seasonal effects with three dummy variables that were 0,0,0 for winter (January–March) flights 1,0,0 for spring (April–June) flights, 0,1,0 for summer (July–September) flights, and 0,0,1 for autumn (October–December) flights.

We used a stepwise linear regression program (BMDP2R; Dixon et al. 1988) to obtain maximum likelihood estimates for parameters in (3) and (4). The stepwise regression program sequentially applied forward selection and backward elimination algorithms to avoid problems associated with using one or the other exclusively (Weisberg 1980; Draper and Smith 1981), and statistical significance was established if the F-value for parameters associated with a main effect or interaction exceeded 4.0. The stepwise regression program made it easy for us to identify and select groups of parameters for main effects and interactions in the model that were statistically significant (i.e. explained a significant amount of variance in the dependent variable; see Kerlinger and Pedhazur 1973). In a few cases, we forced the program to include an insignificant main effect or effects associated with a significant interaction so that, for example, both main effects A and B would be included if the interaction between A and B was significant. Changes in the coefficient of determination (R^2) and order in which the stepwise regression procedure included variables and interactions were used as crude measures of their relative importance.

A standardized and unbiased measure of anchovy density in positive flights during year j is $d_j = \exp(\beta_0 + \beta_{1j} + \sigma^2/2)$ where β_0 is mean log density ($\ln(d_k)$) for the reference year, region, season, time of day, and fish spotter and β_{1j} is the main effect for year j. This measure was estimated by

$$(5) \quad \hat{d}_j = \exp(\hat{\beta}_0 + \hat{\beta}_{1j})\, \Psi_{\hat{d}_j}$$

where "hats" (^) denote estimates from fitting the lognormal linear model (3) to data and Ψ_d is a correction for bias. The correction for bias and variance of d_j were calculated as described in Appendix 1.

Data used to estimate proportion positive (P) were aggregated by year, fish spotter, region, time of day, and season in order to reduce the size of the data set. Preliminary analyses indicated that effects of different approaches to aggregating data were minimal. A lognormal linear model similar to (3) and (4) was used to estimate proportion positive:

$$(6) \quad \ln(P_k + 1) = \ln(f_k/B_k + 1) = f(\boldsymbol{\tau}, \boldsymbol{\gamma}, X_k) + \zeta_k$$

where P_k is proportion positive for record k. Positive blocks (f_k) is the total number of blocks searched during *positive* flights and B_k is the total number of blocks searched during *all* flights that contribute to record k. The additive constant 1 in (6) allowed logarithmic transformation when P_k was zero. Choice of additive constant can effect results from analysis of variance models (Berry 1987). The additive constant 1 was a reasonable choice for our data because it usually reduced and never increased skewness after the transformation, while smaller values (e.g. 0.1) sometimes increased skewness. The linear function $f(\ )$ in (6) is the same as in (3) and (4) except that the vectors $\boldsymbol{\tau}$ and $\boldsymbol{\gamma}$ hold parameters for main effects and interactions (Table 1). The term ζ_k was assumed to be a normally distributed error with expectation zero and variance δ^2.

The assumption of normally distributed errors (ϵ_k and ζ_k) in (3) and (6) was reasonable because of the Central Limit Theorem (values of P_k were averages usually computed from a large number of observations) and because scatter diagrams of residuals were symmetric. A natural alternative would have been to assume binomial-distributed errors and estimate proportion positive (P_k) by logistic regression or other likelihood-based procedures (McCullagh and Nelder 1983). We chose to assume normally distributed errors because software for lognormal linear models facilitated computation of approximate variances for proportion positive and our index of relative abundance (Appendix 1). The assumption of constant variances in (3) and (6) was reasonable because scatter diagrams of residuals did not show any trend when plotted against predicated values or independent variables.

A standardized and unbiased measure of proportion positive during year j is $P_j = \exp(\tau_0 + \tau_{1j} + \delta^2/2) - 1$ where τ_0 is the mean of log proportion positive ($\ln(P_k + 1)$ for a reference year, region, season, time of day, and fish spotter and τ_{1j} is the main effect for year j. This measure was estimated by

$$(7) \quad \hat{P}_j = \exp(\hat{\tau}_0 + \hat{\tau}_{1j})\, \Psi_{\hat{P}_j} - 1$$

where $\Psi_{\hat{P}}$ is a correction for bias and estimates were obtained by fitting the lognormal linear model (6) to data. The correction for bias and variance of $\hat{P}_j$ were calculated as described in Appendix 1.

Substituting (2) into (1) and replacing parameters with their estimates from (5) and (7) gives the following estimate of relative abundance for anchovy in year j:

$$(8)\quad \hat{I}_j = \hat{d}_j \hat{P}_j A_j$$
$$= [\exp(\hat{\beta}_0 + \hat{\beta}_{1j})\,\Psi_{\hat{d}_j}]\,[\exp(\hat{\tau}_0 + \hat{\tau}_{1j})\,\Psi_{\hat{P}_j} - 1]\,A_j.$$

Variance of (8) was estimated as described in Appendix 1. Models and procedures used with environmental and fish spotter data were similar to those used with the complete data set.

The structure of models (3) and (6) allowed us to estimate abundance of anchovy in the last year of the time series from data for only one season in that year. The estimates obtained in our analysis for 1990, for example,were based on data for January–March, while estimates for all other years were based on data for January–December. Use of data for one season or partial season in the last year decreases precision but makes it possible for managers to make estimates, which may be important in some circumstances, of relative abundance during the most recent or current fishing season. Another approach is to use annual periods other than calendar years so that the last season with data is the end of the last annual period (Jacobson and Lo 1991).

Results and Discussion

For the complete data set, the stepwise regression procedure included the same variables and interaction terms in the model (3) for density (d) and as in the model (6) for proportion positive (P) (Table 2). Effects due to time of day, regions, seasons, and fish spotters were all more pronounced than year effects (which are used in (8) to estimate relative abundance) in both models (Table 2). Final estimates for all indices (corrected for bias) are given with standard errors in Table 3.

Results for the environmental and subset of the fish spotter data were similar to those for the complete data set except that the stepwise regression procedure included sea surface temperature in the model (3) for density (d) and both temperature and mixed layer depth in the model (6) for proportion positive (P). Although effects due to sea surface temperature and mixed layer depth were less pronounced than year effects in both models, these results indicate that environmental data were useful in interpreting fish spotter data. Final estimates for all indices (corrected for bias) are given with standard errors in Table 4.

Additional information about the importance of environmental variables can be obtained by examining results for 1983 which was an El Niño year with unusually warm water temperatures in the study area (Fiedler et al. 1986). The estimate of relative abundance (I) for anchovy during 1983 from the complete data set was implausibly low relative to the estimates for 1981, 1982, and 1984 due to low density (d) and proportion positive (P) (Fig. 3; Table 3), while the coefficient of variation (CV) for 1983 was higher than for adjacent years (Table 3). In contrast, the estimate of relative abundance (I) during 1983 and the coefficient of variation from fish spotter and environmental data (Fig. 3; Table 4) were more similar to estimates for 1981 and 1984 (no estimate for 1982 was available because of no environmental data). This result indicates that environmental data may be important in interpreting fish spotter data, particularly, as in El Niño years, when unusual environmental conditions exist.

Total area covered by fish spotters in each year (A) increased from 180 blocks in 1963 to around 350 blocks in the late 1980's (Table 3; Fig. 4). The number of blocks in which anchovy were sighted averaged about 100 blocks during 1963–84 and about 150 blocks during 1985–90. Since the number of blocks covered by fish spotters in each year was larger than the number of blocks in which anchovy were seen, we concluded that there was little bias in our estimates of relative abundance (I) due to fish spotters failing to cover the entire range of the anchovy stock.

Evaluation of Indices

We evaluated indices of relative abundance from fish spotter data by comparing them with estimates of total anchovy bio-

TABLE 2. Summary of lognormal linear models for density in positive blocks (d) and proportion positive (P) fit to the complete fish spotter data set for northern anchovy. Columns give the order of entry and change in R^2 for main effects and interactions included in the models. Numbers in parentheses indicate the number of parameters estimated for each variable or interaction. The row labeled "Records" gives the total number of flights (in the model for density) or records (in the model for proportion positive) used to estimate parameters.

	Density (d)		Proportion positive (P)	
	Order	Change in R^2	Order	Change in R^2
Intercept		Included in all models		
Time of day	1	0.088 (1)	3	0.0016 (1)
Regions	3	0.028 (4)	1	0.056 (4)
Fish spotters	2	0.26 (21)	4	0.105 (21)
Seasons	5	0.0127 (3)	2	0.025 (3)
Years	6	0.058 (27)	5	0.058 (27)
Region by time of day interaction	4	0.0217 (3)[a]	6	0.0172 (4)
Season by time of day interaction	8	0.0016 (3)	7	0.009 (3)
Region by season interaction	7	0.015 (10)[a]	8	0.011 (10)[a]
Records	6793		3733	
Total R^2 (adjusted)		0.48		0.24

[a]Some coefficients were not estimated because of collinearity with other independent variables.

TABLE 3. Relative abundance of anchovy during 1963–90 and other estimates from the complete fish spotter data set. A is the area (number of unique blocks) covered by fish spotters in each year, number positive is the number of blocks in which anchovy were seen. d is density (tons per block) of anchovy for positive flights. P is proportion positive. I is relative abundance of anchovy, and CV is the coefficient of variation. Estimates of density (d) and proportion positive (P) were corrected for bias as described in the text.

Year	A	Number positive	d[a]	CV(d)(%)	P	CV(P)(%)	I	CV(I)(%)
1963	180	126	13	22	0.47	15	1 129	27
1964	206	118	76	26	0.61	16	9 569	30
1965	208	112	67	25	0.56	17	7 851	31
1966	224	106	109	24	0.60	16	14 643	29
1967	200	108	132	23	0.72	14	19 105	27
1968	221	106	50	25	0.48	19	5 260	31
1969	223	90	163	23	0.51	18	18 663	29
1970	143	75	167	24	0.71	16	16 965	29
1971	175	82	156	23	0.70	15	19 082	27
1972	184	87	154	23	0.67	16	19 025	28
1973	338	146	339	23	0.64	15	73 493	28
1974	304	96	255	23	0.76	14	58 892	27
1975	280	68	392	23	0.84	14	91 789	27
1976	388	97	221	23	0.63	16	53 966	28
1977	289	99	248	23	0.71	15	50 683	27
1978	283	84	358	25	0.52	18	52 898	31
1979	288	88	590	25	0.54	19	91 366	31
1980	196	100	455	29	0.69	19	61 345	34
1981	232	110	313	29	0.46	23	33 281	37
1982	249	88	264	28	0.42	24	27 846	37
1983	340	79	78	36	0.27	33	7 125	49
1984	341	82	176	32	0.29	31	17 373	45
1985	357	151	349	30	0.44	22	54 261	38
1986	332	122	132	29	0.42	23	18 289	38
1987	352	119	114	31	0.43	24	17 103	40
1988	335	201	275	31	0.50	22	46 164	38
1989	379	142	81	32	0.47	23	14 373	40
1990	351[b]	147[b]	227	42	0.41	33	32 527	54

[a]Rounded to nearest integer.
[b]Area (A) and number positive for 1990 estimated from the average values for 1985–89.

mass during 1964–86 (Fig. 3) obtained independently by catch-at-age analysis using a stock synthesis model (Methot 1989). The stock synthesis model is a form of catch-at-age analysis that incorporates information about relative abundance of spawning or schooling anchovy from four fishery-independent surveys, commercial fishery data, and environmental data in a single maximum likelihood based procedure. Although the stock synthesis model for northern anchovy was designed to estimate spawning biomass, total biomass estimates are produced as well (Methot 1989; Lo and Methot 1989). Stock synthesis estimates of anchovy biomass have been used to set catch quotas since 1986 and were the best estimates available.

We also compared our estimates of relative abundance with updated versions of the simple T/BAF index from fish spotter data (Fig. 3) described by Squire (1972, 1983) to determine if our more complex approach was any better.

When comparing indices, it is important to remember that estimates of relative abundance from fish spotter data measure changes in *schooling* biomass rather than total, catchable, or spawning biomass. These distinctions may be important if, for example, young fish do not school or a large fraction of the stock is immature.

The first step in evaluating indices was to determine which were linear measures of relative biomass for northern anchovy. We fit a quadratic polynomial (parabolic) regression model to each index of relative abundance:

$$(9) \quad L_y = \iota_1 B_y + \iota_2 B_y^2 + \epsilon_y$$

where L_y is an index that measures biomass in year y (e.g. T/BAF), B_y is anchovy biomass during year y estimated by Methot (1989) using a stock synthesis model, ι_1 and ι_2 are parameters, and ϵ_y is a normally distributed statistical error. The important features of this model are that it intersects the origin (indices have zero value at zero anchovy biomass) and nonlinearity in the relationship between an index and anchovy biomass is indicated by the sign and statistical significance of the estimate of ι_2. If the index of abundance (L) is not a linear estimate of anchovy biomass (B), then the estimate of ι_2 should be significantly different from zero. Negative values of ι_2 indicate saturation. The parabolic model (9) worked well for our data because the inflection point of the fitted line fell outside the range of the data for each index and the line was always increasing in the range from zero to the largest observed biomass level.

Results from the regression analyses (Table 5) indicate that our index of relative abundance from fish spotter data (I) and T/BAF were linear measures of relative anchovy biomass (Table 5). Parameter estimates for ι_2 were not statistically significant (p-value > 0.05) for our index of relative abundance from fish spotter data (I, with and without environmental data), T/BAF, and, in the case of fish spotter with environmental data, proportion positive (P).

To further evaluate linear measures of relative biomass for anchovy, we computed correlation coefficients between indices from fish spotter data and stock synthesis estimates of anchovy biomass:

TABLE 4. Relative abundance of anchovy during 1963–90 and other estimates from a subset of the complete fish spotter data plus environmental data. Column headings and definitions as in Table 3. Area (A in (8)) used to estimate relative abundance from fish spotter and environmental data was the same as listed in Table 3 for a complete data set. Estimates for some years are unavailable due to a lack of environmental data.

Year	d^a	CV(d)(%)	P	CV(P)(%)	I	CV(I)(%)
1963	41	33	0.19	9	1 366	34
1964	532	43	0.16	14	17 452	45
1965	368	45	0.16	15	12 076	47
1966	864	35	0.15	12	29 311	37
1967	820	40	0.17	13	28 634	42
1968	245	40	0.08	25	4 126	48
1969	943	33	0.13	13	27 583	36
1970						
1971						
1972	810	35	0.13	13	19 449	37
1973			No environmental data available			
1974	819	37	0.30	9	74 747	38
1975	2 564	34	0.25	8	179 773	35
1976	676	41	0.10	25	26 526	48
1977	645	40	0.11	21	20 911	46
1978	1 804	39	0.06	30	29 046	49
1979	1 153	77	0.06	48	19 729	91
1980	1 809	57	0.17	19	60 890	60
1981	1 057	52	0.11	22	26 877	56
1982						
1983	520	57	0.03	75	5 389	94
1984	562	57	0.04	57	8 387	81
1985	891	56	0.05	50	16 123	76
1986	794	57	0.05	52	12 687	77
1987	868	57	0.05	53	14 566	78
1988	434	63	0.02	129	3 166	144
1989	2 052	67	0.06	46	45 568	81
1990	1 445	63	0.07	48	33 333	79

[a]Rounded to nearest integer.

Index	Correlation with stock synthesis estimates
T/BAF	0.60
I (no environmental data)	0.67
I (with environmental data)	0.63

Indices of relative abundance (I) for anchovy based on delta-lognormal linear models appear superior to the T/BAF index because correlations with stock synthesis estimates of total biomass were slightly higher and because the T/BAF index dropped to implausibly low levels after 1982 (Fig. 3).

The correlation between stock synthesis estimates of biomass and our estimates of relative abundance (I) from environmental plus fish spotter data was almost as good as the correlation with estimates from the complete data set even though the sample size for each year was smaller. This result indicates that environmental data, if readily available, could be used to improve the precision of relative abundance indices from fish spotter data.

Simulation of Anchovy Biomass Estimates

Two simulation experiments involving the stock synthesis model for anchovy were conducted to determine if our index of schooling biomass (I) could be used to improve spawning stock biomass estimates for northern anchovy. The stock synthesis model was used in the simulation experiment because it was convenient and actually used to manage anchovy, but any other type of catch-at-age analysis with auxiliary information (Deriso et al. 1985; Pope and Shepherd 1985; Gavaris 1988) would have been just as suitable. Our results are general enough to indicate the relative benefits of including fish spotter data in catch-at-age analysis with auxiliary information for fisheries similar to the anchovy fishery. A detailed description of the simulations is given in Appendix 2.

Simulation results indicated that fish spotter data increased the precision of biomass estimates from the stock synthesis model by about 10% on average (Fig. 5). Coefficients of variation for spawning biomass estimates with fish spotter data ranged from 7.8 to 40% in the last year and ranged from 7.6 to 54% in the last year without fish spotter data.

Fish spotter data may have reduced bias in stock synthesis estimates of spawning biomass for northern anchovy. Bias of stock synthesis estimates for each year was measured by the percentage difference between the mean estimate from the simulations and the "true" value assumed in the simulation experiment, i.e. as $(I_{\text{average}} - I_{\text{true}})/I_{\text{true}} \times 100$. Percentage difference is a reasonable measure of bias when the number of simulations employed is large but may be misleading when the number of simulations is as small as 50 because the mean of a small sample will differ from the underlying population mean due to sampling error only (Efron 1982).

The absolute value of our measure of bias ranged from 0 to 10% (0–63 000 tons) of true values in the simulation for relative abundance estimates from fish spotter data and from 0 to 12% (0–75 000 tons) for relative abundance estimates without fish spotter data (Fig. 5). As explained above, however, these

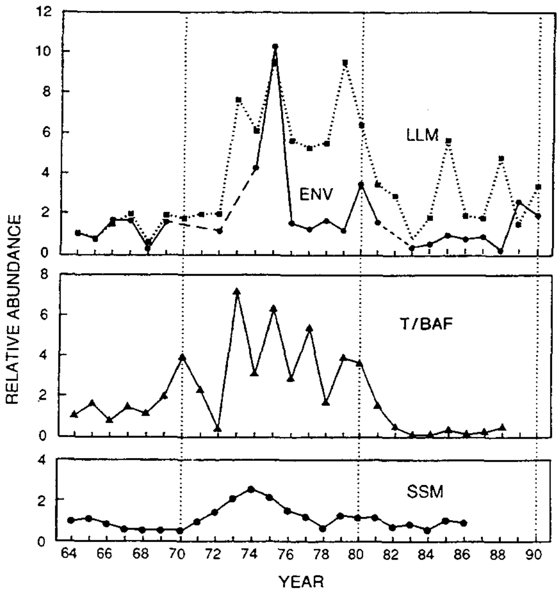

FIG. 3. Stock synthesis model (SSM) estimates of total biomass (Methot 1989), a catch-per-unit-effort-like index (T/BAF) from fish spotter data (revised and updated values similar to data given in Squire 1972), and estimates of relative abundance from delta-lognormal linear models fit to fish spotter data with (ENV) and without (LLM) environmental data. Before plotting, each series was scaled so that the value of 1964 was equal to 1.0 (i.e. the values in each series were divided by the value for 1964).

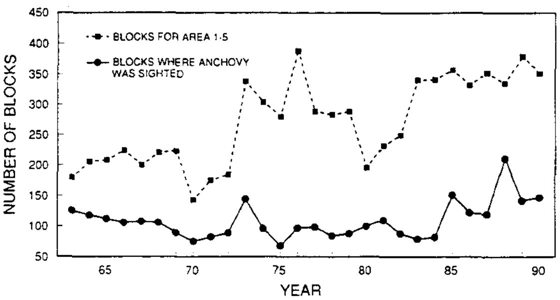

FIG. 4. Number of blocks in regions 1–5 covered by fish spotters (*A*) and number of blocks where anchovy were found during 1963–90.

results were based on relatively few simulations and may be misleading. Additional work is required to determine the nature and extent of possible bias in the stock synthesis model.

The most important effect of using fish spotter data in the simulations was increased precision and potential reduction in bias for the last year's estimate. The biomass estimate for the last year is the least precise in most stock assessments (e.g. Rivard 1989) but also the most important because it is used to set the next catch quota. Use of fish spotter data in our simulation experiment resulted in an improvement of 26% in the precision of spawning biomass estimates for the last year and may have reduced bias by 17% (i.e. [12% − 10%]/12%). These results indicate that indices of relative abundance based

TABLE 5. Sign (+ or −) for ι_2 and its p-value under the null hypothesis that ι_2 = zero for regression analyses of anchovy abundance indices (equation (9)). Negative estimates indicate saturation. Small p-values (e.g. $p < 0.05$) indicate nonlinearity.

Index	Estimate for ι_2 Sign	p-value
T/BAF	−	0.910
Area (*A*)	−	0.000
Indices from fish spotter data only		
Density for positive blocks (*d*)	−	0.011
Proportion positive (*P*)	−	0.000
Relative abundance (*I*)	−	0.317
Indices from fish spotter and environmental data		
Density for positive blocks (*d*)	−	0.049
Proportion positive (*P*)	−	0.793
Relative abundance (*I*)	+	0.266

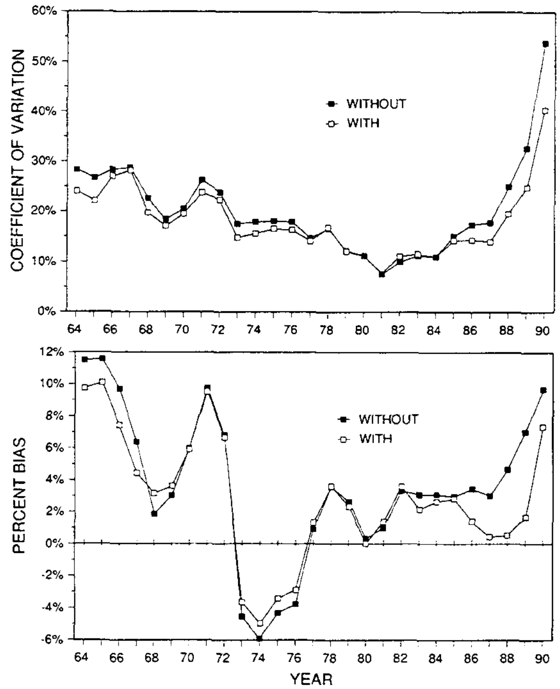

FIG. 5. Coefficients of variation (top panel) and percent bias (bottom panel) for estimates of anchovy total biomass during 1964–90 from the stock synthesis model. Coefficients of variation and percent bias were calculated by simulation for estimates made with and without fish spotter data.

on fish spotter data are cost-effective ways to improve biomass estimates used to manage the fishery for northern anchovy.

Our work extended the delta distribution models proposed by Aitchison and Brown (1957) and Pennington (1983) because both density of anchovy in positive blocks (*d*) and proportion positive (*P*) were estimated using models that allowed us to include factors (e.g. pilots) that affected estimates of density and proportion positive and increased the precision of our estimates. Our approach should be applicable whenever the statistical distribution of data is distorted by excessive zeros (e.g. ichthyoplankton surveys) and affected by external factors or covariates, but care should be taken to ensure that the distribution of nonzero data is as assumed for estimation procedures (Meyers and Pepin 1990).

Future Research

The models we developed for fish spotter data may be best suited for relatively short time series because the number of parameters will increase as data are collected for more years and pilots. At some point, the number of parameters may grow large enough to make the models cumbersome and too large for small computers. Future work could usefully be directed towards reducing the number of parameters either by restricting the number of years included in the analysis or developing alternative statistical approaches.

For sufficiently long collections of fish spotter data, it may be possible and advantageous to develop time series models (Abraham and Ledolter 1983). Our approach estimated relative abundance for each year separately and may not have fully utilized the data because relative abundance of fish in adjacent years is usually similar enough to be used for making estimates (Roff 1983).

Acknowledgements

Larry Eber (NMFS, Southwest Fisheries Science Center) compiled environmental data collected during CalCOFI sea surveys. John Hunter, Alec MacCall, and an anonymous reviewer read earlier versions of the paper and made helpful suggestions. Cindy Myer and John Lim assisted with preparation of data for analysis. Judy Ester helped with the figures. The San Diego Supercomputer Center (affiliated with the University of California, San Diego) provided time on a Cray supercomputer and other resources used to conduct the simulations.

References

ABRAHAM, B., AND J. LEDOLTER. 1983. Statistical methods for forecasting. John Wiley and Sons, New York, NY. 445 p.
AGENBAG, J. J. 1980. General distribution of pelagic fish off southwest Africa as deduced from aerial fish spotting (1971–1974 and 1977) and as influenced by hydrology. Fish. Bull. S. Afr. 13: 55–67.
AGENBAG, J. J., I. KRUGER, AND F. LE CLUS. 1984. Environmental factors associated with a distinct decline in pelagic fish abundance as detected by aerial fish spotting in February and March 1982 off northern South West Africa. S. Afr. J. Mar. Sci. 2: 93–107.
AITCHISON, J., AND J. A. C. BROWN. 1957. The lognormal distribution. Cambridge University Press, Cambridge, MA. 387 p.
ANONYMOUS. 1981. Stock assessment activities within the National Marine Fisheries Service. NOAA Tech. Mem. NMFS, NOAA-TM-NMFS-SWFC-12: 130 p.
BANNEROT, S. P., AND C. B. AUSTIN. 1983. Using frequency distributions of catch per unit effort to measure stock abundance. Trans. Am. Fish. Soc. 112: 608–617.
BARNES, J. T., L. D. JACOBSON, A. D. MACCALL, AND P. WOLF. 1992. Recent population trends and abundance estimates of the Pacific sardine. Calif. Coop. Oceanic Fish. Invest. Rep. 33: 60–75.
BERRY, D. A. 1987. Logarithmic transformations in ANOVA. Biometrics 43: 439–456.
BRADU, D., AND Y. MUNDLAK. 1970. Estimation in lognormal linear models. J. Am. Stat. Assoc. 65: 198–211.
CRAM, D. L., AND I. HAMPTON. 1976. A proposed aerial/acoustic strategy for pelagic fish stock assessment. J. Cons. Int. Explor. Mer 37: 91–97.
DENNIS, B., AND G. P. PATIL. 1988. Applications in ecology, p. 303–330. *In* E. L. Crow and K. Shimizu [ed.] Lognormal distributions: theory and applications. Marcel Dekker Inc., New York, NY.
DERISO, R. B., T. J. QUINN II, AND P. R. NEAL. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42: 815–824.
DIXON, W. J., M. B. BROWN, L. ENGELMAN, M. A. HILL, AND R. I. JENNRICH. 1988. BMDP statistical software manual. Vol. 1. University of California Press, Los Angeles, CA. 619 p.
DRAPER, N. R., AND H. SMITH. 1981. Applied regression analysis. John Wiley and Sons, New York, NY. 709 p.
EFRON, B. 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, PA. 92 p.
FIEDLER, P. C., R. D. METHOT, AND R. P. HEWITT. 1986. Effects of California El Niño 1982–1984 on the northern anchovy. J. Mar. Res. 44: 317–338.
GAVARIS, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29: 12 p.
GULLAND, J. A. 1956. On the fishing effort in English demersal fisheries. Fish. Invest. Ser. II Mar. Fish. G.B. Minist. Agric. Fish. Food 20(5).
HABIB, G., I. T. CLEMENT, K. A. FISHER, P. M. SWANSON, C. L. CAREY, AND G. J. VOSS. 1982. School fish sightings around New Zealand in 1981. N.Z. Minist. Agric. Fish. Fish. Res. Div. Occas. Publ. Data Ser. No. 10: 40 p.
HAMPTON, I., J. J. AGENBAG, AND D. L. CRAM. 1979. Feasibility of assessing the size of the South West African pilchard stock by combined aerial and acoustic measurements. Fish. Bull. S. Afr. 11: 10–22.
HARA, I. 1985. Shape and size of Japanese sardine school in the waters off the southeastern Hokkaido on the basis of acoustic and aerial surveys. Bull. Jpn. Soc. Sci. Fish. 51: 41–46.
1986. Stock assessment of Japanese sardine in the waters off the southeast coast of Hokkaido using line transect method. Bull. Jpn. Soc. Sci. Fish. 52: 69–73.
1990. Comparison of ship and aerial surveys of sardine schools. IEEE Trans. Geosci. Remote Sens. 28: 693–695.
HEWITT, R. P. 1988. Historical review of the oceanographic approach to fishery research. Calif. Coop. Oceanic Fish. Invest. Rep. 29: 27–41.
HILBORN, R., AND C. J. WALTERS. 1992. Quantitative fisheries stock assessment. Choice, dynamics and uncertainty. Routledge, Chapman and Hall Inc., New York, NY. 570 p.
JACOBSON, L. D., AND N. C. H. LO. 1990. Spawning biomass of the northern anchovy in 1990. Natl. Mar. Fish. Serv. Southwest Fish. Sci. Cent. Admin. Rep. LJ-90-20: 31 p.
1991. Spawning biomass of the northern anchovy in 1991. Natl. Mar. Fish. Serv. Southwest Fish. Sci. Cent. Admin. Rep. LJ-91-19: 53 p.
JACOBSON, L. D., W. R. MACCALLUM, AND G. R. SPANGLER. 1987. Biomass dynamics of Lake Superior lake herring (*Coregonus artedii*): application of Schnute's difference model. Can. J. Fish. Aquat. 44(Suppl. 2): 275–288.
KERLINGER, F. N., AND E. J. PEDHAZUR. 1973. Multiple regression in behavioral research. Holt, Rinehart and Winston Inc., New York, NY. 534 p.
KIMURA, D. K. 1981. Standardized measures of relative abundance based on modelling log (c.p.u.e.), and their application to Pacific ocean perch (*Sebastes alutus*). J. Cons. Int. Explor. Mer 39: 211–218.
1988. Analyzing relative abundance indices with log-linear models. N. Am. J. Fish. Manage. 8: 175–180.
LAMBERT, D. 1992. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34: 1–14.
LO, N. C. H., AND R. D. METHOT. 1989. Spawning biomass of the northern anchovy in 1988. Calif. Coop. Oceanic Fish. Invest. Rep. 30: 18–31.
MACCALL, A. D. 1984. Report of the working group on resources study and monitoring, p. 9–39. *In* J. Csirke and G. D. Sharp [ed.] Reports of the expert consultation to examine changes in abundance and species composition of neritic fish resources. FAO Fish. Rep. No. 291, Vol. 1: 102 p.
1990. Dynamic geography of marine fish populations. University of Seattle Press, Seattle, WA. 153 p.
MACCALL, A. D., AND M. H. PRAGER. 1988. Historical changes in abundance of six fish species off southern California, based on CalCOFI egg and larva samples. Calif. Coop. Oceanic Fish. Invest. Rep. 29: 81–101.
MAIS, K. F. 1974. Pelagic fish surveys in the California current. Calif. Dep. Fish Game Fish Bull. 162: 79 p.
MANGEL, M., AND P. E. SMITH. 1990. Presence–absence sampling for fisheries management. Can. J. Fish. Aquat. Sci. 47: 1875–1887.
MAUGHAN, P. M., AND A. D. MARMELSTEIN. 1971. Operational use of remote sensors in commercial fishing, p. 8–24. *In* Proceedings of the symposium on remote sensing in marine biology and fishery resources, January 25–26, 1971. Texas A&M University Remote Sensing Center, College Station, TX. TAMU-SG-71-106.
MCCULLAGH, P., AND J. A. NELDER. 1983. Generalized linear models. Chapman and Hall, New York, NY. 261 p.
METHOT, R. D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Am. Fish. Soc. Symp. 6: 66–82.
MEYERS, R. A., AND P. PEPIN. 1990. The robustness of lognormal-based estimators of abundance. Biometrics 46: 1185–1192.

PENNINGTON, M. 1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics 39: 281–286.
POPE, J. G., AND J. G. SHEPHERD. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Int. Explor. Mer 42: 129–151.
RIVARD, D. 1989. Overview of the systematic, structural, and sampling errors in cohort analysis. Am. Fish. Soc. Symp. 6: 49–65.
RIVAS, L. R. 1978. Aerial surveys leading to 1974–1976 estimates of the numbers of spawning giant bluefin tuna (*Thunnus thynnus*) migrating past the western Bahamas. Int. Comm. Conserv. Atlantic Tunas Collect. Vol. Sci. Pap. 7: 301–312.
ROBSON, D. S. 1966. Estimation of the relative fishing power of individual ships. Int. Comm. Northwest Atl. Fish. Res. Bull. 3: 5–14.
ROFF, D. 1983. Analysis of catch effort data: a comparison of three methods. Can. J. Fish. Aquat. Sci. 40: 1496–1506.
SCOTT, G. P., M. R. DEWEY, L. J. HANSEN, R. E. OWEN, AND E. S. RUTHERFORD. 1989. How many mullet are there in Florida Bay? Bull. Mar. Sci. 44: 89–107.
SEBER, G. A. F. 1982. The estimation of animal abundance and related parameters. Oxford University Press, New York, NY. 654 p.
SHIMIZU, K. 1988. Point estimation, p. 27–87. *In* E. L. Crow and K. Shimizu [ed.] Lognormal distributions: theory and applications. Marcel Dekker Inc., New York, NY.
SMITH, P. E. 1990. Monitoring interannual changes in spawning area of Pacific sardine (*Sardinops sagax*). Calif. Coop. Oceanic Fish. Invest. Rep. 31: 145–151.
SQUIRE, J. L. JR. 1961. Aerial fish spotting in the United States commercial fisheries. U.S. Fish Wildl. Serv. Commer. Fish. Rev. 23: 1–7.
1972. Apparent abundance of some pelagic marine fish off the southern and central California coast as surveyed by an airborne monitoring program. Fish. Bull. 70: 1005–1019.
1983. Abundance of pelagic resources off California, 1963–78, as measured by an airborne fish monitoring program. NOAA Tech. Rep. NMFS SSRF-762: 75 p.
ULLTANG, O. 1977. Methods of measuring stock abundance other than by the use of commercial catch and effort data. FAO Fish. Tech. Pap. 176: 23 p.
WEISBERG, S. 1980. Applied linear regression. John Wiley and Sons, New York, NY. 283 p.
WILLIAMS, K. 1981. Aerial survey of pelagic fish resources off southeast Australia. CSIRO Div. Fish. Ocean., Cronulla, NSW, Australia Rep. 130.

Appendix 1. Corrections for Bias and Variance Calculations

Approximate estimates for the variance of relative abundance in each year (I in (8)) were computed by the delta method (Seber 1982):

$$\text{(A1)}\quad \mathrm{Var}(\hat{I}) \approx A^2\,[\mathrm{Var}(\hat{d})\,\hat{P}^2 + \hat{d}^2\,\mathrm{Var}(\hat{P}) + 2\,\hat{d}\hat{P}\mathrm{Cov}(\hat{d},\hat{P})]$$

where hats (^) denote estimates, Var ( ) denotes variance, I is the index of relative abundance for anchovy in each year, A is area of the anchovy stock, d is the density of anchovy in positive blocks, P is the proportion positive, and $\mathrm{Cov}(\hat{d},\hat{P})$ denotes the within-year covariance of estimates for density (d) and proportion positive (P). Variances for estimates of d and P were calculated as described below. The covariance term in (A1) was estimated approximately from the correlation of d and P *among years* and the *within-year* standard deviation of d and P:

$$\text{(A2)}\quad \mathrm{Cov}(\hat{d},\hat{P}) \approx \rho_{\hat{d},\hat{P}}\,[\mathrm{SE}(\hat{d})\,\mathrm{SE}(\hat{P})]$$

where $\rho_{\hat{d},\hat{P}}$ denotes the correlation and SE( ) denotes a standard error.

Unbiased estimates of anchovy density in positive blocks (d in (5)) and proportion positive blocks (P in (7)) as well as variance estimates for d and P were calculated as described by Bradu and Mundlak (1970). The correction for bias in (5) and (7) involves a correction factor Ψ (Ψ_d for density and Ψ_P for proportion positive). Assuming lognormal-distributed errors, the correction factor is

$$\text{(A3)}\quad \Psi = g_m\left[\frac{m+1}{2m}(\hat{\xi}^2 - \hat{\xi}^2_{\hat{\eta}})\right]$$

where $g_m(\ )$ is a function described below, $\hat{\xi}^2$ is the residual variance (σ^2 in (3) for density or δ^2 in (6) for proportion positive), m is degrees of freedom for the estimate of residual variance, $\hat{\eta} = \hat{\beta}_0 + \hat{\beta}_{1j}$ for density (d) or $\hat{\eta} = \hat{\tau}_0 + \hat{\tau}_{1j}$ for proportion positive, and $\hat{\xi}^2_{\hat{\eta}}$ is the variance of $\hat{\eta}$.

The function $g_m(\ )$ in (A3) is

$$\text{(A4)}\quad g_m(t) = \sum_{p=0}^{\infty}\left[\frac{m^p\,(m+2p)}{m(m+2)\ldots(m+2p)}\left(\frac{m}{m+1}\right)^p \frac{t^p}{p!}\right]$$

where t is the argument for the function.

Variance estimates for $\hat{d}$ in (5) and $\hat{P}$ in (7) were calculated as

$$\text{(A5)}\quad \mathrm{Var}(.) = e^{2\hat{\eta}}\left\{g_m^2\left[\frac{m+1}{2m}\left(\hat{\xi}^2 - \hat{\xi}^2_{\hat{\eta}}\right)\right] - g_m\left[\frac{m+1}{m}\left(\hat{\xi}^2 - 2\,\hat{\xi}^2_{\hat{\eta}}\right)\right]\right\}$$

where Var( ) is the variance of either $\hat{d}$ or $\hat{P}$.

Appendix 2. Simulations

First, the stock synthesis model for northern anchovy was modified to include the new index of schooling biomass from fish spotter data. The original stock synthesis model for anchovy included an index of schooling biomass based on a sonar survey conducted by the California Department of Fish and Game during 1969–85 (Mais 1974; Methot 1989). Sonar and fish spotter indices both measure relative schooling biomass, so the pattern of age-specific availability of anchovy to the fish spotter index was assumed to be the same as for the sonar survey (age-specific availability patterns are used in the stock synthesis model to relate the index of schooling biomass to stock age structure). This assumption allowed us to include the new fish spotter index for anchovy with a minimum of additional programming. The fish spotter index was given the same relative weight as other types of information about abundance in the stock synthesis model.

The modified stock synthesis model was fitted to data, including the index based on fish spotter data, for 1964–90. After the model was fitted, predicted values for each type of data were calculated and stored. Standard deviations for the fit of the model to each type of relative abundance data were calculated (assuming a lognormal distribution for discrepancies between observed and predicted values).

Fifty simulated data sets with fish spotter data and 50 simulated data sets without fish spotter data were generated (the number of simulations conducted was limited by available computer time). A pseudorandom number generator, predicted values from the original fit, and standard deviations calculated as described above were used to generate relative abundance data for the simulation experiment. Age composition data were generated from predicted age compositions by assuming multinomial sampling errors. The stock synthesis model was fit to each of the simulated data sets and coefficients of variation for estimates of spawning biomass during 1964–90 were calculated to evaluate precision. This approach to estimating coefficients of variation for spawning biomass estimates is a "parametric bootstrap" approach in the language of Efron (1982).