ABSTRACT

We have developed annual indexes of larval abundances for five species sampled by the CalCOFI ichthyoplankton surveys from 1951 to 1985: northern anchovy, Pacific sardine, Pacific mackerel, jack mackerel, and Pacific whiting (or hake). We used a similar method to calculate annual indexes of egg abundances for a sixth species—Pacific saury. We pooled observations of larval densities within six geographic regions off southern California and northern Baja California, summa-
rizing the data into region-month-year cells. We then used a multiplicative model of independent region, month, and year factors to predict the abundance in each cell. Under log-transform, the model becomes additive (i.e., log-linear) and was solved by analysis of variance (ANOVA). The ANOVA solution has the added advantage of accommodating missing values (43% of the possible 2160 cells lack observations). The antilogged month factors provided estimates of the mean temporal distribution of larvae over the months of the year. Antilogged region factors similarly provided the mean geographic distribution. Antilogged year factors from an ANOVA weighted by the monthly and regional distributions provided the annual indexes of abundance.

Abundance indexes for all six species show large fluctuations over the 35-year period. The index of northern anchovy larvae increased 50-fold from 1951 to the mid 1960s; since then it has fluctuated widely but without a trend. Indexes of Pacific sardine and Pacific mackerel decline from 1951 to the mid 1970s, and rise sharply thereafter, consistent with known increases in abundance. Indexes of jack mackerel and Pacific whiting are highly variable, and trends are difficult to identify. The index of Pacific saury eggs is high in the 1950s, is slightly lower for the next two decades, and drops severely in the 1980s. Changes in indexes of the latter three species are probably associated with environmental influences that have produced large-scale geographic shifts in spawning.

RESUMEN

Indices anuales de producción de larvas han sido desarrollados para 5 especies muestreadas por las expediciones de ictioplancton de CalCOFI: anchoveta del Norte (Engraulis mordax), sardina del Pacífico (Sardinops sagax), macarela del Pacífico (Scomber japonicus), jurel (Trachurus symmetricus), y merluza (Merluccius productus). Un método similar fue usado para calcular índices anuales de abundancia de huevos para una sexta especie, Cololabis saira. Las observaciones de densidad de larvas dentro de 6 regiones geográficas frente al sur de California y la región norte de Baja California fueron combinadas y resumidas en categorías de región-mes-año. En cada categoría se usó un modelo multiplicativo con factores independientes de región, mes y año para predecir su abundancia. Con una transformación logarítmica, el modelo se volvió aditivo (i.e., lineal en escala logarítmica) y fue solucionado con un análisis de varianza (ANOVA). La solución ANOVA tiene la ventaja adicional de permitir la ausencia de datos (un 43% de las 2160 categorías posibles carece de observaciones). La función antilogarítmica de los factores mensuales entregó estimaciones de la distribución temporal promedio de larvas durante los meses del año. Similarmente, la función antilogarítmica de los factores regionales entregó la distribución geográfica promedio. La función antilogarítmica de los factores anuales de un ANOVA, ponderados por las distribuciones mensuales y regionales proporcionó índices anuales de abundancia.

Los índices de abundancia para las 6 especies muestran grandes fluctuaciones durante el periodo de 35 años. El índice de las larvas de la anchoveta del Norte aumentó 50 veces entre 1951 y mediados de la década de 1960; desde entonces ha fluctuado enormemente pero sin una tendencia clara. Los índices de la sardina y la macarela del Pacífico disminuyeron desde 1951 hasta mediados de la década de 1970, para después aumentar abruptamente, en forma consistente con aumentos en abundancia previamente establecidos. Los índices de T. symmetricus y M. productus son altamente variables y es difícil establecer un patrón. El índice de los hue-
vos de *C. saira* es alto en la década de 1950, levemente inferior en las dos siguientes décadas, y cae considerablemente en la década de 1980. Los cambios en los índices de las últimas 3 especies están probablemente relacionados con factores ambientales que han producido desplazamientos geográficos de gran escala en el desove.

**INTRODUCTION**

The California Cooperative Oceanic Fisheries Investigations (CalCOFI) have sampled fish eggs and larvae off the coast of California and northern Baja California since 1951. One of the goals of these and similar ichthyoplankton surveys has been to compile time series of abundances of the parent fish species (Moser et al. 1987; Smith and Moser 1988). In developing such abundance estimates there are many difficulties, ranging from uncertainty about the physiology of spawning activity to the statistics of highly patchy populations and incomplete sampling.

The present attempt to derive abundance time series from the CalCOFI data emphasizes development of a uniform methodology that can be applied to all species, independent of the amount of supplemental knowledge. We emphasize that these time series represent a first attempt for most of these species, and additional knowledge should lead to improved estimates for individual species. For example, Lo (1985) used information on growth rates and the size-frequency distribution of anchovy larvae to estimate their mortality rates and rates of parental egg production. The latter provided an improved index of adult abundance. However, we lack such supplementary information for most species. By comparing this “information-poor” analysis with complex, more reliable analyses of the same and/or supplementary data for well-studied species, we can infer the relative performance of this simple approach.

**DATA AND METHODS**

We summarized the CalCOFI egg or larval data by six geographical regions (Figure 1), for each month and year in which samples were taken. Although these regions were originally defined for purposes of stratifying distributions of anchovy larvae (Smith 1972; see Lo 1985 for current region definitions), they suffice for initial investigations of other species. It is important to note that the geographic regions used in this study do not necessarily cover the full spawning range of some of the species treated here. Because many sets of CalCOFI cruises began in November, we used a sampling year of November through October in this analysis.

Following the approach of Smith (1972), we calculated the average larval density for each of these year-month-region strata, or “cells,” as the mean of the individual density estimates for all stations occupied within the cell1 (see Kramer et al. 1972 for a general description of sampling methods; detailed data and changes in sampling are described

---

1 In cases where a cruise extended slightly into a second month, with very small numbers of stations occupied, we nominally combine the data under the adjacent principal month. These months were 11–62, 12–67, 5–68, 11–72, 11–74, 11–77, 4–82, 2–83, and 11–84. Data from 11–75, 12–75, and 11–85 were not used in the analysis.
in a series of publications beginning with Ambrose et al. 1987). During the study period 1951 to 1985, only six years (1970, 1971, 1973, 1974, 1976, and 1977) were not sampled at all. Many years were sampled thoroughly, especially 1951 through 1960. In the 30 “survey years” during the study period, a total of 1238 cells were sampled, which is about 57% of the possible 2160 cells. Missing values in the remaining 43% of the cells present statistical difficulties in developing reliable abundance estimates.

A good solution to the problem of missing values is the method of maximum likelihood, but that method requires developing a complicated probabilistic model based on individual observations. The sheer number of observations (nearly 12,000) would present difficult computational problems. A simpler approach is to use an established model and statistical method that naturally accommodate missing values. For example, previous abundance indexes have used summation of data aggregated on a quarterly basis (Smith 1972), and linear interpolation has been proposed as a basis for estimation. Use of such a constant is a necessary evil in this log-linear approach and can distort the values of the estimated factors, especially for seasonal or yearly with very low larval abundances. The added constant of 0.01 larvae/10 m² was near or slightly smaller than the smallest nonzero larval density measurable for a cell at typical sampling densities. Because the model is estimated under log transformation, the factors reflect geometric mean effects.

A further limitation of the simple ANOVA model is its disregard for interactions among factors. These interactions, such as geographic shifts in spawning areas as the season progresses, are known to occur (e.g., Kramer and Smith 1970). They can be addressed by a model such as

\[ \ln(N_{ik} + 0.01) = \ln(C_i) + \ln(T_k) + \text{constant} + \text{error} \]  

where \( C_i \) is a combined month-region factor (\( i = 1, \ldots, 72 \)). We used this model to examine month-region interactions for the peak months of spawning.

Moreover, the estimated time series of abundance indexes (\( T_k \)) were compared for the two models, providing a check on the amount of distortion introduced by the lack of interaction terms in equation 1. Shifts in spawning seasons or areas over the years also are known to occur, but formal recognition of these interactions in a model similar to equation 2 requires an excessive number of estimated parameters. Moreover, sample sizes for some individual factors become rather small, and a strong year-versus-region pattern of missing values makes estimation problematic; this is an area for further investigation.

In this log-linear ANOVA model with equal weighting of all cells, the year factors are highly influenced by months and regions that account for relatively few larvae. Accordingly, we used a weighted ANOVA to estimate the year factors. We chose a weighting scheme that allowed the influence of various months and regions to be made
roughly proportional to their average contribution to the total year's spawning. Thus we approximated the weighting implicit in an additive "larval census" treatment, which is a sum of larvae over regions and seasons within a year (Smith 1972). Note that this weighted ANOVA should tend to be robust to error arising from variations in the duration of the spawning season (to the extent that duration is independent of spawner abundance), but may not be robust to error arising from overall shifts in spawning season or location among years.

In determining cell weights, we did not include the number of stations contributing to the cell mean; this modification is worth further consideration, but is somewhat more complicated than would first appear. For example, individual stations can vary substantially in amount of water filtered, as well as probable extrusion and/or avoidance by the larvae. The solution to these variable sample volume problems is likely to take the form proposed by Zweifel and Smith (1981), who weighted individual observations before averaging.

In addition to weighting the ANOVA to obtain the year factors, we restricted the months used in the ANOVA to peak spawning months. These months tend to have more and larger nonzero observations, reducing distortion due to addition of a constant for logarithmic transformation. On the other hand, using a small number of months increases difficulties arising from missing observations. Except for Pacific whiting, which has a brief spawning period, we included at least four months in the nominal peak spawning period in order to balance these two considerations.

The antilogged month and region factors from the previous unweighted ANOVA were used to calculate the weights for the weighted ANOVA, as well as to identify the peak spawning months. In accordance with equation 1, weights were calculated as the product of a month-weighting contribution and a region-weighting contribution. The month contribution consisted of the antilogged factors for the peak months, taken from the unweighted ANOVA and scaled to a unit sum. The region contribution was calculated similarly, except that antilogged region factors were first multiplied by the area of the corresponding region, in order to represent relative abundance of larvae in keeping with the analogy to a larval census (Smith 1972). The regions are shown in Figure 1, and their areas (taken from Lo 1985) are given in Table 1. Because of the month and region weights, observations in some cells are more important than others. A rough index of "effective sample size" is provided by the sum of cell weights for nonmissing values in each year (Table 2).

The added constant of 0.01 larvae/10 m² created a minimum possible abundance index, i.e., an index value obtained if no larvae were observed in any cell. We determined the approximate value of this minimum by supplying the ANOVA with an additional year of artificial data containing no positive observations. We assigned exceedingly small weights to these artificial data to avoid influencing the ANOVA results for the actual data.

The comparison of time series of year factors (T_k) as estimated by the simple and interaction models in equations 1 and 2 followed a procedure identical to that described above: We used an unweighted "main effects" SAS GLM procedure (ANOVA) to identify the peak months, and to provide cell weights accordingly. Weighted ANOVAs then produced time series of abundance indexes for the two models. We used the same cell weights for both models. Our use of peak months was especially important in this comparison, because the added constant in the logarithmic transformations in equations 1 and 2 would tend to create erroneous interactions in months and regions where larval densities were low. For a similar reason, the comparison was restricted to years that

<table>
<thead>
<tr>
<th>Region</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>11</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern anchovy</td>
<td>0.828</td>
<td>0.024</td>
<td>0.087</td>
<td>0.121</td>
<td>0.020</td>
<td>0.074</td>
</tr>
<tr>
<td>Pacific sardine</td>
<td>0.478</td>
<td>0.053</td>
<td>0.078</td>
<td>0.173</td>
<td>0.144</td>
<td>0.074</td>
</tr>
<tr>
<td>Pacific mackerel</td>
<td>0.273</td>
<td>0.090</td>
<td>0.153</td>
<td>0.106</td>
<td>0.219</td>
<td>0.159</td>
</tr>
<tr>
<td>Jack mackerel</td>
<td>0.050</td>
<td>0.084</td>
<td>0.272</td>
<td>0.033</td>
<td>0.318</td>
<td>0.243</td>
</tr>
<tr>
<td>Pacific whiting</td>
<td>0.458</td>
<td>0.128</td>
<td>0.249</td>
<td>0.072</td>
<td>0.063</td>
<td>0.028</td>
</tr>
<tr>
<td>Pacific saury</td>
<td>0.214</td>
<td>0.129</td>
<td>0.336</td>
<td>0.036</td>
<td>0.157</td>
<td>0.128</td>
</tr>
<tr>
<td>Region area (10³ n.mi.²)</td>
<td>20</td>
<td>12</td>
<td>29</td>
<td>9</td>
<td>21</td>
<td>29</td>
</tr>
</tbody>
</table>
are relatively well sampled, i.e., that lacked no more than about one month’s observations during the peak spawning period.

RESULTS

In all six species cases, the time series of abundance indexes estimated from equation 1 and the month-region interaction model of equation 2 are nearly indistinguishable, indicating that no serious distortion results from the model in equation 1. Because of the lack of differences, we do not include detailed results of these comparisons in this report.

Northern Anchovy (Engraulis mordax)

The unweighted ANOVA showed a protracted spawning season (Figure 2), with the peak months of February through May accounting for 85% of the larvae seen during the year. The two inshore
regions account for 95% of the total larvae (Table 1). Because the southern, offshore region 14 accounts for very few larvae (less than 0.1%), we omitted it from the weighted ANOVA.

Overall, the abundance indexes (Figure 3) were quite similar to the historical egg production time series developed by Lo (1985) from much more complete information including the size distribution of larvae. The main exception was the index value for 1978, which we estimated to be much lower than Lo’s value. However, Lo’s estimate for 1978 (unlike other years in the series) is based on an assumed preliminary value regarding egg retention by the bongo net (Lo 1983). This value was later revised based on experiments conducted in 1984 (N. Lo, SWFC, pers. comm.); the updated egg retention value was used by Methot (1986) and gives a 1978 value that is similar to the index value developed here.

**Pacific Sardine (Sardinops sagax)**

The long apparent spawning season (Figure 4) is partially an artifact of the added constant; sardine larvae have been scarce for many of the 35 years covered by the surveys. The main spawning period is April, May, and June, but a broad peak appears to extend from February through July and includes at least 85% of the larvae seen during the year. About half of the larvae occurred in inshore southern California waters (region 7), and 80% occurred in combined nearshore regions 7,11, and 13 (Table 1). As in the case of nonpeak spawning months, the importance of the offshore regions (8, 9, and 14) was probably exaggerated because of the added constant.

The minimum abundance index corresponding to “no larvae” was only slightly below the estimates obtained for the 1970s (log$_2$[minimum] = −3.99; Figure 5). Our abundance estimates for this low period are therefore not reliable, and the population may well have reached lower levels of abundance than those suggested by our analysis. However, for the first half of the time series, the indicated decline is in agreement with the trend in estimates of abundance derived from fishery catch data (MacCall 1979).

**Pacific Mackerel (Scomber japonicus)**

As in the case of the sardine, the long apparent spawning season is partially an artifact of the added constant and the low abundance of larvae experi-
enced over much of the time series. Peak spawning is in May through July (Figure 6); however, we included April as a fourth month to increase the sample size. The period April–July accounts for at least 68% of the larvae seen during the year (the true average spawning season is probably more peaked than that calculated here, because of the added constant). Larvae were distributed rather evenly among the six regions (Table 1), but again this evenness also may be exaggerated by the added constant. The index value for 1975 is very near the minimum possible estimate ($\log_{10} \text{minimum} = -4.26$; Figure 7); as in the case of the sardine, lower levels of abundance may have been reached than are indicated by our analysis. The sudden increase in 1978 and 1981 is consistent with the recent resurgence of the population (MacCall et al. 1985).
Jack Mackerel (Trachurus symmetricus)

There is a sharp peak of larval abundance in May and June (Figure 8), but for purposes of the ANOVA, we used a peak spawning season of March through July, accounting for 97% of the annual total. Larvae are mostly offshore, with regions 9, 13, and 14 accounting for 83% of the total (Table 1). At the time this study was conducted, the CalCOFI data for 1951 were incomplete, so 1951 was not included in the calculations. The time series of larval abundance indexes (Figure 9) is highly variable, and does not correspond closely to the time series of mean larval abundances compiled by MacCall and Stauffer (1983), which included information from waters off central California. We also note for jack mackerel as well as for some other species that the variability in this index must be far greater than the true variability of a population composed of long-lived individuals.

Pacific Whiting or Hake (Merluccius productus)

The very short spawning season of whiting (Figure 10) prevents extending the analysis beyond February, March, and April, during which 97% of the larvae are observed. Also, 84% of the larvae were observed in the northern three regions (Table 1). Ahlstrom (1969) notes that through 1966, the standard CalCOFI plankton samples, which reached a maximum depth of about 140 m, “evidently do not encompass the complete vertical distribution of hake eggs and larvae at some stations.” The depth of the standard CalCOFI plankton samples was extended to about 210 m beginning in 1969, at which time the netting was also changed from silk to an equivalent but more uniform nylon
The index of whiting abundance (Figure 11) is higher during the period following these sampling changes, but changes in actual abundance cannot be discounted; the index increases in 1966, immediately prior to the sampling changes, and the low levels of 1982–84 are lower than any seen previously. The index for whiting is highly variable, and does not agree with abundance estimates based on fishery data (Hallowed et al. 1987), which show a major increase during the 1980s.

**Pacific Saury (Cololabis saira)**

Egg abundance is very high in April (Figure 12), but we assumed a more extended peak spawning season of February through June, covering 90% of the annual total eggs observed. The geographic distribution of eggs is fairly even, except for a scarcity in region 11 (Table 1). There has been an irregular decline in abundance of saury eggs in the study area since the early 1950s, with eggs being nearly absent since 1980 (note that the latter indexes approach the minimum value possible in this analysis, \( \log_2 [\text{minimum}] = -5.24 \); Figure 13). We cannot determine from internal evidence if the change is due to a northward contraction of the spawning distribution, or if it is a population decline. Few years during the transition period were sampled; the southern regions are incompletely covered; and the geographic coverage of this study is very limited with respect to the overall distribution of saury spawning.

**DISCUSSION**

The ANOVA method is successful in that it accommodates missing values and produces easily derived indexes of abundance. The method does have shortcomings, especially in treating geographic or seasonal shifts in distribution among years, which seems to be a major source of error. Inclusion of data from waters off central California might have clarified some northward shifts, but that area has a high incidence of missing values, and its contribution would be correspondingly imprecise. In the case of rare larvae, results are biased by the added constant in equation 1. Of course, there is little information content in years where few or no larvae are encountered.

Several improvements on this log-linear or ANOVA approach should be investigated. Missing values may be better accommodated by the EM algorithm (Haberman 1974). Even better would be the maximum likelihood approach, which would include variable sample sizes among cells, but re-
Figure 13. Time series of abundance indexes for Pacific saury eggs. Upper: arithmetic scale, mean = 1; lower: log₂ scale. Dashed line indicates minimum abundance index due to added constant.

Figure 14. Time series of annual average Scripps Pier water temperatures. Horizontal line is long-term mean.

Results would still be imprecise. Methods for accounting for region-year and region-month interactions should be investigated. The performance and robustness of various approaches could be investigated by means of simulated data. Regions could be redefined in a manner appropriate for each species, and regions off central California should be included.

Major changes in the indexes of three species—Pacific sardine (increase), Pacific mackerel (increase), and Pacific saury (decrease)—have occurred in the last decade, and there is a suggestion that the indexes for Pacific whiting and jack mackerel also have decreased. The California Current has tended to be warmer since the mid 1970s. Annual average Scripps Pier temperatures (Figure 14) have risen by about one-half a degree Celsius, ending a cool period extending from the early 1940s. The warming has been more consistent in the offshore regions of the California Current (Cole and McLain, in press). Pacific saury and jack mackerel spawn relatively northward and offshore, and we speculate that the declines in abundance indexes for these two species are probably the result of northward spawning shifts in response to this warming. For this reason, we prefer to interpret the recent declines in these indexes as shifts in distribution rather than decreases in population abundance.

The spawning area of Pacific whiting similarly shifted northward in recent years (Paul Smith, SWFC, pers. comm.), presumably also in association with the oceanic warming. We also interpret the recent decline in the abundance index for whiting as the result of a northward shift in spawning, with a substantial portion of the population now spawning off central California. Contrary to the decrease in this larva-based index, fishery-based estimates of whiting abundance show a substantial increase during the 1980s, but declining sizes at age in the U.S. fishery independently suggest a northward shift of the resource (Hallowed et al. 1987).

On the other hand, the increases in indexes for sardine and Pacific mackerel, which also coincide with the warming, are definitely associated with respective increases in population abundance (Wolf and Smith 1986; MacCall et al. 1985). Again, these increases have been accompanied by northward shifts in the spawning distributions as indicated by the larval distributions.

The time-series of these six species underscores
the importance of long-term monitoring. Had the CalCOFI surveys ceased in 1975, these time series would be far less interesting. Of the six species, only the index for anchovy shows little change after 1978, and that index shows a very large change in the first decade of the time series. Thus, a time series spanning only 25 years would risk missing major population or environmental changes, whereas a series spanning 35 years seems, at least in this case, to have detected major changes.

Isaacs (1976) emphasized the importance of "regimes" in contrast to the usual assumptions of stability (or stable distributions of environmental conditions) underlying fishery models. The temperature history in Figure 14 suggests three major temperature regimes: warm from the beginning of observations in 1916 to the early 1940s (>30 yr), cold to the mid 1970s (ca. 35 yr), and then warm again. The earlier warm regime preceded the era of physical and biological monitoring. The physical and biological changes associated with the recent warming indicate that Isaacs's concern with regimes was justified. Long-term monitoring, such as has been conducted by CalCOFI, is vital to understanding these phenomena.

ACKNOWLEDGMENTS

In addition to the many people who have contributed to the CalCOFI program over the past 35 years, we wish to thank several people who contributed particularly to the present paper. Celeste Santos Methot compiled the source data from the CalCOFI data files. Rich Charter, John Michno, and Kevin Colwell provided computer programming and data base management support. Geoff Moser and his co-workers provided summaries of larvae from recent CalCOFI cruises virtually as the identifications were being completed. Finally, we thank the Ocean Assessments Division, National Ocean Service, NOAA, for supporting this investigation.

LITERATURE CITED


