

Abstract.— For fish populations in equilibrium, estimates of Z/K and L_∞ can be obtained solely from length-frequency data by using the Wetherall length-based method. Robustness of this method to departures from equilibrium conditions is examined with a population simulation model. During disequilibrium conditions following either the initiation of a fishery or a 1-year perturbation in recruitment, estimates of both parameters, especially Z/K , are severely biased. Three-year averaging of catch length-frequencies does not substantially reduce this bias. A test, based on the chi-square statistic, is proposed for detecting population disequilibrium. Provided that the sizes of length-frequency samples are sufficiently large, the test is an effective way to detect population disequilibrium and thereby avoid biased parameter estimates.

Robustness of the Wetherall Length-based Method to Population Disequilibria

David A. Somerton
Donald R. Kobayashi

Honolulu Laboratory, Southwest Fisheries Science Center
National Marine Fisheries Service, NOAA
2570 Dole Street, Honolulu, Hawaii 96822-2396

Assessments of tropical fish stocks have increasingly turned to length-based methods rather than the more traditional age-based methods, because of the difficulty in ageing tropical fishes (Pauly and Morgan 1987). Although all length-based methods use length-frequency data, they are designed to measure different biological parameters and vary in their requirements for additional types of data. Wetherall et al. (1987) examined one class of length-based methods—those designed to estimate both growth and mortality parameters without the need for additional data—and developed two new estimators: a maximum likelihood estimator and a regression estimator. The second estimator, because of its simplicity, has received the most widespread use (Wetherall 1986, Arellano 1989, Polovina 1989, Rawlinson 1989).

The regression estimator of Wetherall et al. (1987), henceforth referred to simply as the Wetherall method, estimates two parameters. The first (Θ) is the ratio of the instantaneous total mortality rate (Z) to the rate constant of the von Bertalanffy growth function (K). The second (L_∞) is the asymptote of the von Bertalanffy growth function. These parameters are estimated by regressing the mean length (\bar{l}_i) of all fish $\geq l_{ci}$ on l_{ci} , a cutoff length ranging from the first length category that is fully selected by the fishery (L_c) to the largest length category. Estimates of Z/K and L_∞ are then calcu-

lated as $Z/K = B/(1 - B)$ and $L_\infty = A/(1 - B)$, where A is the intercept and B is the slope of the regression.

Like many other length-based methods, the Wetherall method requires the population to be in equilibrium, a condition often not fulfilled because of variations in both the environment and the fishery (Csirske et al. 1987, Ralston 1989). Although Wetherall et al. (1987) clearly have cautioned potential users that biases may result if populations are in disequilibria, the likely magnitudes of such biases have not been addressed. This has prompted us to investigate the sensitivity of the method to two common types of perturbations: (1) a rapid increase in effort during the fishing-up stage of a fishery, and (2) a fluctuation in recruitment. In this paper, we examine the temporal patterns and magnitudes of biases associated with these disequilibria and present a simple statistical test that can be used with catch length-frequencies to detect population disequilibria and help minimize the consequent potential for biased parameter estimates.

Materials and methods

The performance of the Wetherall method can be assessed by applying it to catch length-frequencies generated by a length-based population simulation model that accounts for growth, mortality, and recruitment. The model is configured to

simulate a population of opakapaka *Pristipomoides filamentosus*, a Hawaiian deepwater snapper. Mean length-at-age (L_t) is described by a von Bertalanffy growth equation:

$$L_t = L_\infty (1 - e^{-K(t-t_0)}); \quad (1)$$

where L_∞ , K , and t_0 are parameters. Estimates for these parameters ($L_\infty = 66$ cm, $K = 0.24$ /year, and $t_0 = -0.78$ year) are from Ralston and Miyamoto (1983). As a computational convenience, however, growth is considered an incremental process:

$$L_t = L_1 + \sum_{i=1}^{t-1} DL_i, \quad (2)$$

where the annual growth increment (DL_t) is expressed as a time-differenced form of Equation (1):

$$DL_t = L_{t+1} - L_t = L_\infty e^{-K(t-t_0)} (1 - e^{-K}). \quad (3)$$

Variance in length-at-age (V_t) is described by an asymptotic function of age:

$$V_t = C(1 - e^{-Dt}), \quad (4)$$

where C and D are parameters. Estimates of these parameters can be obtained by fitting this function to size and age data obtained from Ralston and Miyamoto (1983), but the predicted variance is unrealistically large. Consequently, the parameters $C = 10$ and $D = 0.1$ are chosen to best fit the predicted-to-observed length-frequency distributions. Variance in annual growth increment is described as a time-differenced form of Equation (4):

$$DV_t = V_{t+1} - V_t = C(1 - e^{-D}) e^{-Dt}. \quad (5)$$

Recruitment is assumed to occur instantaneously at age-1 and, unless experimentally manipulated, to be identical each year. The length distribution of recruits is assumed to be normal with a mean and variance defined by Equations (1) and (4), respectively. The length distribution of the surviving members of this age-class in subsequent years (y) is estimated with the recursive relationship:

$$N_{l,t+1,y+1} = \sum_{s=l_0}^1 N_{s,t,y} e^{-(M+QF)} P(1-s|t); \quad (6)$$

where M and F are the instantaneous rates of natural and fishing mortality, Q is a size-dependent selectivity coefficient, l_0 is the smallest individual in each age

group and $P(1-s|t)$ is the conditional probability of growing an amount $(1-s)$ at t years of age. The growth-increment probability function is assumed to be normal with a mean and variance computed from Equations (3) and (5). An estimate of natural mortality ($M = 0.3$) is obtained from Ralston (1987); estimates of fishing mortality ($F = 0.3$ and 0.6) are chosen somewhat arbitrarily to bracket the probable true value for opakapaka. The selectivity coefficient Q is represented by a reparameterized logistic function of length:

$$Q = \frac{1}{1 + e^{-E(l-l_{50})}}, \quad (7)$$

where E is a parameter controlling the steepness of the function, and l_{50} is a parameter controlling the length of 50% selectivity. The parameter values of $E = 0.5$ and $l_{50} = 45$ cm are chosen arbitrarily. The population length-distribution is calculated by summing across all age-classes, i.e.,

$$N_{l,..y} = \sum_t N_{l,t+1,y}, \quad (8)$$

and the catch length-frequency distribution is calculated as

$$C_{l,..y} = N_{l,..y} \left(\frac{QF}{M+QF} \right) (1 - e^{-(M+QF)}). \quad (9)$$

The parameters Z/K and L_∞ are estimated from the simulated catch length-frequency distribution as follows. First, the length at full vulnerability (L_c) is estimated as one length interval (1 cm) larger than the rightmost mode in the catch length-frequency distribution (Polovina 1989). Second, \bar{l}_i , the mean length of all fish $\geq l_{ci}$, is calculated for each $l_{ci} > L_c$. Third, \bar{l}_i is regressed on l_{ci} by using weighted linear regression with weights equal to the sample size (i.e., $\sum C_i$ for $l > l_c$). Fourth, Z/K and L_∞ are then computed from the previously specified functions of the regression coefficients.

The disequilibrium experiments are conducted as follows. For the experiments examining the fishing-up phase of a fishery, the initial population length-frequency distribution is set at the equilibrium distribution in the absence of fishing. This length distribution is generated by running the model with $F = 0.0$ until the population is in equilibrium (i.e., until the total population size is identical on two successive iterations). Fishing mortality at each of the two specified levels is then applied instantaneously, and the simulation is run until the population is again at equilibrium. For the experiments examining the effect of perturbation dur-

ing recruitment, the initial population length-frequency distribution is set at the equilibrium distribution with a fishery producing $F = 0.6$. Two types of recruitment perturbations are tested: (1) a year with twice the normal recruitment, and (2) a year with a complete absence of recruitment. After each recruitment perturbation is introduced, the simulation is run until the population is again in equilibrium. Since disequilibrium bias may be only one component of the total bias of a parameter, the disequilibrium bias is isolated by expressing it as a percentage difference relative to the equilibrium value rather than the known value of a parameter.

Population equilibrium is statistically tested by comparing a catch length-frequency distribution in any one year to those from the previous two years by using a chi-square test of independence. Rationale for this type of test is that a population in equilibrium should, except for sampling variability, produce catch length-frequencies that are identically distributed over time. Significance of the chi-square test therefore indicates that the population is not in equilibrium. Performance of the chi-square test first requires the construction of 3-by- N frequency tables, where N is the number of size categories that are jointly defined over the 3-year sequence. To do this, the largest and smallest size categories in the catch containing at least five individuals in each of the three years are determined. Then for each year, all size categories less than the lower bound of the joint interval are pooled, as are all size categories greater than the upper bound. Once the table is constructed, the chi-square test can be performed according to the procedures described in Conover (1971).

The statistical power of the test is examined in a Monte Carlo experiment in which the test is applied to catch length-frequencies generated by the population simulation model. For each disequilibrium experiment, annual catch length-frequencies are subsampled, with replacement, with subsample sizes of 100, 500, 1000, and 5000 fish. At each level of subsampling, a chi-square test of independence is then performed on each 3-year sequence of subsampled length-frequencies. Subsampling and testing are replicated 100 times, and the power is estimated as the proportion of the 100 tests that is significant at the 5% level.

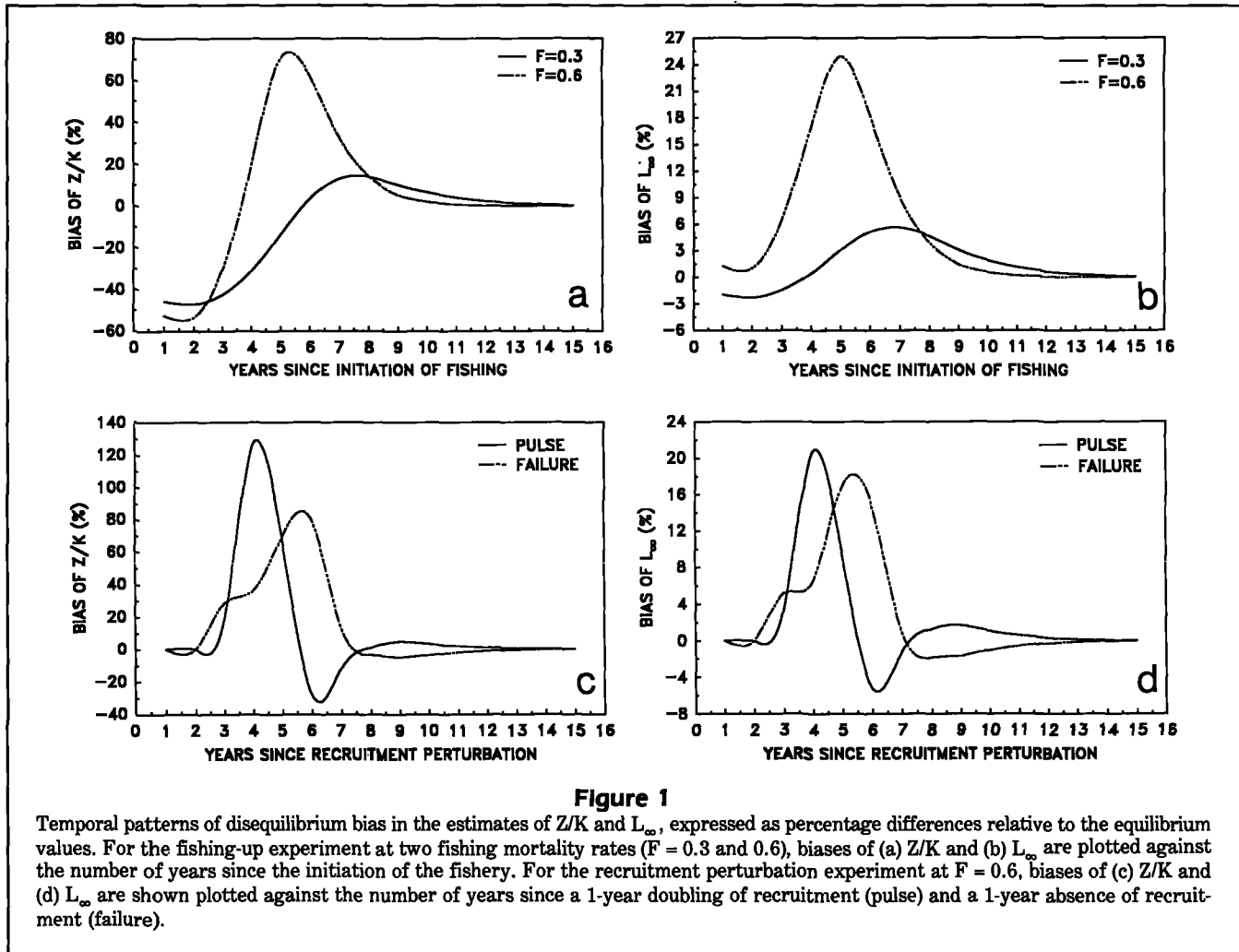
To determine whether disequilibrium bias could be reduced by averaging a series of catch length-frequencies over time, the Wetherall method is also applied to simulated catch length-frequencies after they have been time-averaged. First, the catch length-frequencies for each year are converted to proportions by length. Second, the time-series of catch proportions for each size interval is smoothed by using a 3-year centered running average. Third, the smoothed catch proportions at length for each year are then multiplied

by the actual catch (in numbers) for that year to recover the true sample size.

In addition to disequilibrium bias, also examined are two types of bias that can influence Wetherall estimates of Z/K and L_∞ even when the population is in equilibrium. The focus of this examination is on the variation in these biases as a function of the chosen value of L_c . To do this, the model first is run until equilibrium conditions are reached. Then for each chosen value of L_c , the Wetherall method is applied to the catch length-frequency distribution, which experiences size selection by the fishery, and to the population length-frequency distribution. Type I bias in the estimates of Z/K is expressed as the percentage difference between the estimates based on catch length-frequencies and those based on population length-frequencies. Type II bias in the estimates of Z/K is expressed as the percentage difference between the estimates based on population length-frequencies and the known values of Z/K .

Results and discussion

As Wetherall et al. (1987) have cautioned, population disequilibrium indeed leads to biased estimates of Z/K and L_∞ , but the temporal pattern and magnitude of this bias vary greatly with the type of equilibrium perturbation. In the fishing-up experiment, disequilibrium bias in the estimates of Z/K is initially large and negative, but later becomes positive before decaying to zero (Fig. 1a). Increasing the fishing mortality rate increases the magnitude of the positive bias and shortens the time in which the maximum positive bias is reached. The magnitude of fishing mortality, however, appears to have little effect on the magnitude of the initial negative bias. Disequilibrium bias in the estimates of L_∞ changes with time and with fishing mortality rate in a manner similar to the bias in Z/K , but the bias in L_∞ is almost always positive and considerably smaller than the bias in Z/K (Fig. 1b). In the recruitment perturbation experiment, when recruitment is doubled for 1 year, the disequilibrium bias in the estimates of Z/K is oscillatory in time and has a pronounced positive peak and a less-pronounced negative peak (Fig. 1c). When recruitment is eliminated for 1 year, the disequilibrium bias in Z/K is again oscillatory, but the positive and negative peaks occur 2 years later and the magnitude of the bias, especially that of the negative bias, is smaller. Disequilibrium bias in L_∞ varies with time in a manner similar to the bias in Z/K , but again the bias is smaller in magnitude (Fig. 1d). Thus, the following two points emerge from these simulations: (1) Biases in both parameters can be large but vary tremendously as the size structure of the

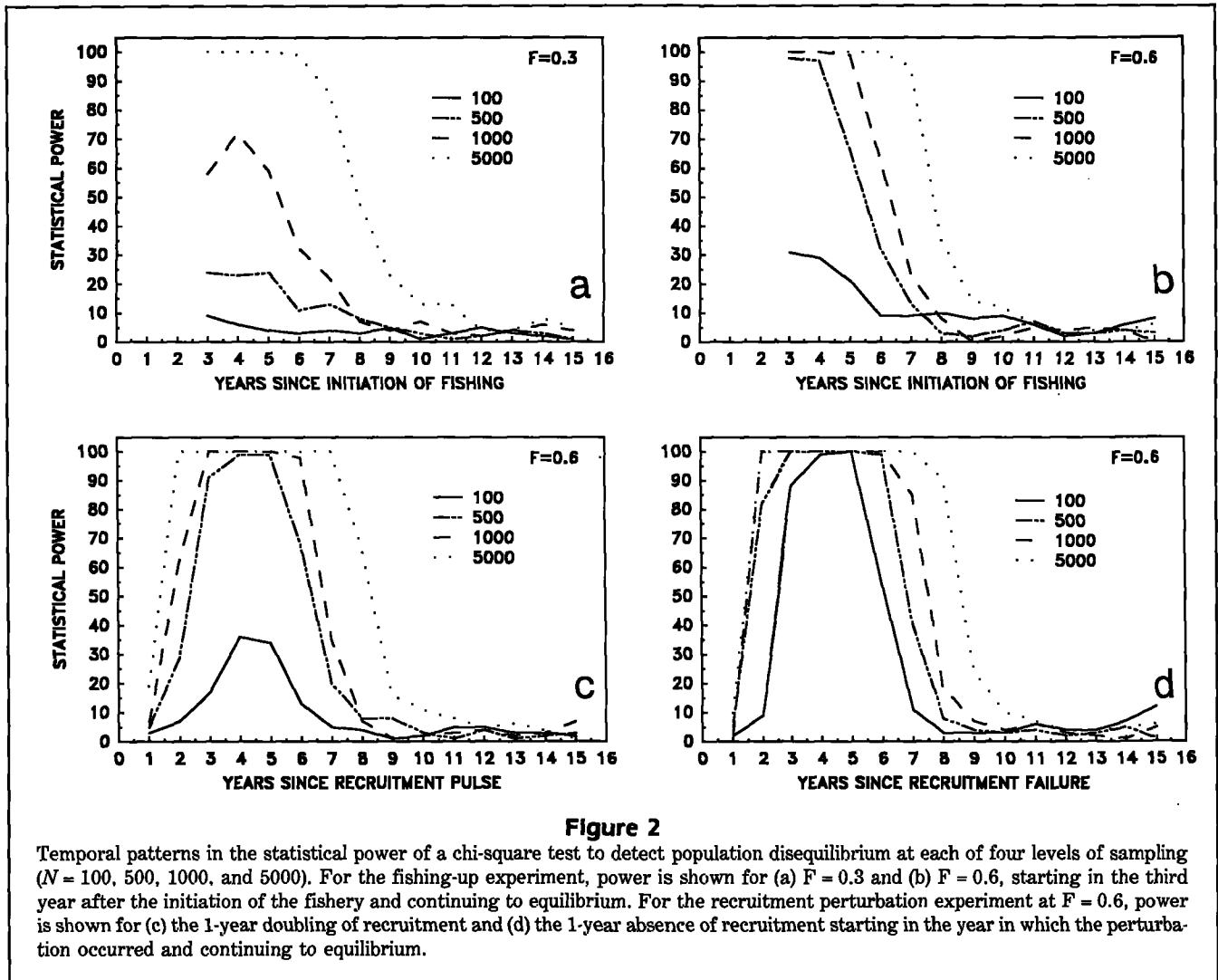


population adjusts from one equilibrium to the next, and (2) bias in Z/K estimates is considerably larger than bias in L_{∞} estimates.

Even though the simulated changes in fishing mortality and recruitment may be more abrupt than those typically experienced by the population, the simulations clearly demonstrate that the Wetherall estimates of Z/K and L_{∞} can be quite biased when the population is not in equilibrium. As a consequence, the method apparently can be used with confidence only when population equilibrium can be demonstrated. Such a demonstration of equilibrium, or lack thereof, can be achieved by comparing successive catch length-frequencies with a chi-square test of independence. The utility of the chi-square test can best be judged by examining its statistical power to detect the disequilibrium conditions that lead to biased parameter estimates. Since the statistical power of the proposed test will increase as between-year differences in catch length-frequencies

increase, it will tend to vary positively with the magnitude of parameter bias. Provided that sample sizes are sufficiently large, the test can therefore be used as a data screening device to detect when parameter bias is likely. One important consideration, then, is the relationship between sample size and the undetectable level of disequilibrium bias. This relationship can be seen by comparing the time trajectories of power (Fig. 2) with those of bias (Fig. 1).

Consider first the fishing-up experiment and bias in estimates of Z/K . For $F = 0.3$ and a sample size of 5000 fish, the statistical power of detecting disequilibria is high (i.e., ≥ 0.75) from years 3 to 7 (Fig. 2a). Although this interval includes most of the period in which bias is high (Fig. 1a), it does not include the first 2 years after the initiation of the fishery when bias is also high, because the proposed test requires 3 successive years of data. Furthermore, the interval of high statistical power does not include years 8–10 when bias is nearly



13%. For $F = 0.6$ and a sample size of 5000 fish, the interval of high power is almost identical to that at $F = 0.3$. At lower sample sizes, however, the chi-square test is considerably more powerful when $F = 0.6$ than when $F = 0.3$.

In the recruitment perturbation experiment, statistical power is high over a broader time-interval than in the fishing-up experiment. For the case of a 1-year doubling of recruitment at a sample size of 5000 fish, power is high between years 2 and 8, an interval that includes the entire period in which disequilibrium bias is $>5\%$. For the case of a 1-year absence of recruitment, power is high between years 2 and 8, an interval that again includes the entire period in which disequilibrium bias is $>5\%$. Unlike the situation in the fishing-up experiments, power tends to remain relatively high with reductions in sample size. These experiments indicate that, in terms of its ability to detect the

likelihood of disequilibrium bias, the chi-square test performs best on a 1-year doubling of recruitment, second best on a 1-year absence of recruitment, and worst on the fishing-up disequilibrium. The fishing-up case is worst because the test cannot be applied to the first 2 years of the time-series when bias is high and larger sample sizes are needed to detect disequilibrium.

In addition to the type of disequilibrium perturbation and size of the length-frequency sample used, the statistical power of the chi-square test also varies with the number of length-distributions included. Thus, power is calculated for tests including two, three, and four length-distributions. Since the gain in power is substantial when the number of length-distributions is increased from two to three but only minor when the number is increased from three to four, all further power simulations are based on three length-distributions.

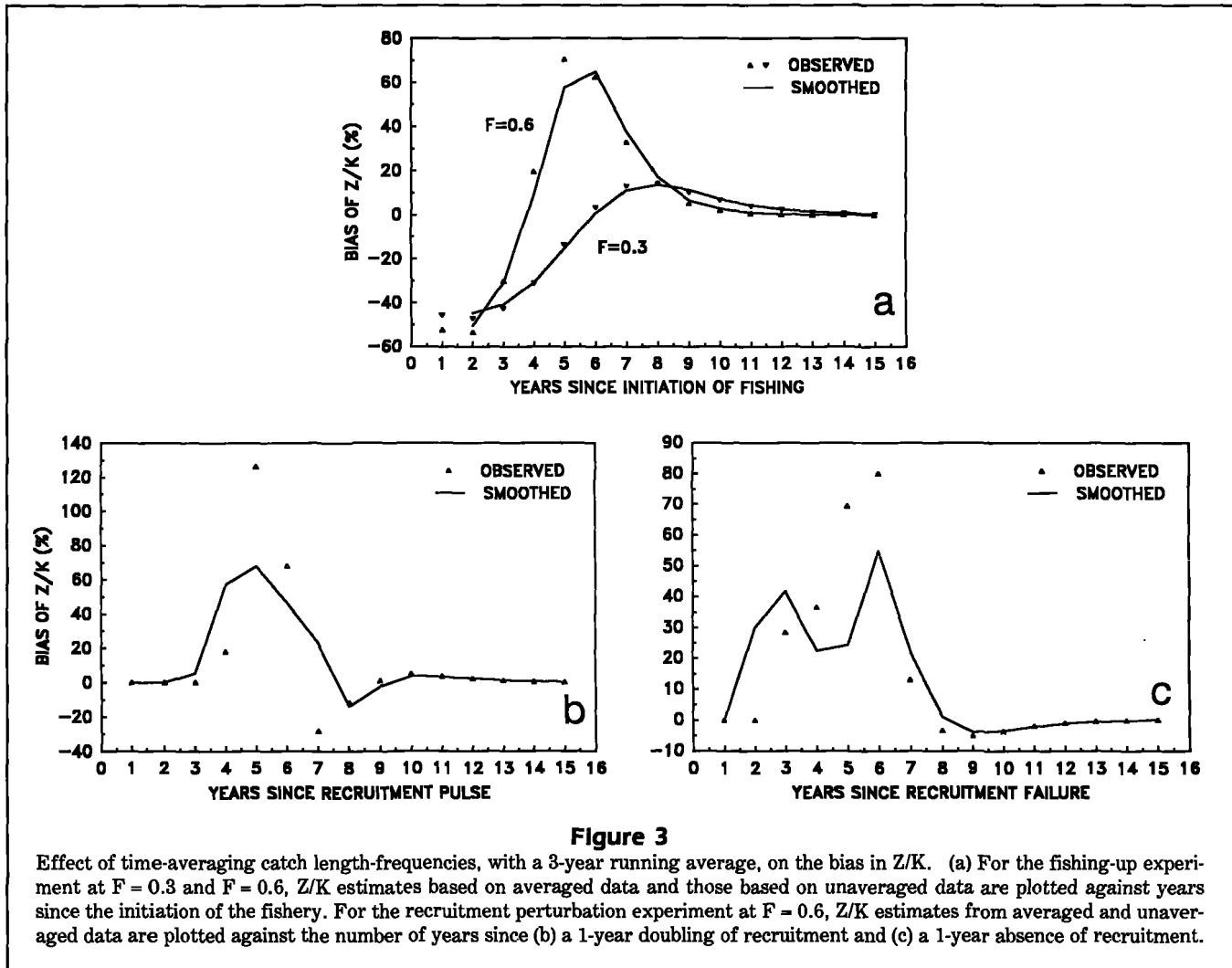


Figure 3

Effect of time-averaging catch length-frequencies, with a 3-year running average, on the bias in Z/K. (a) For the fishing-up experiment at $F = 0.3$ and $F = 0.6$, Z/K estimates based on averaged data and those based on unaveraged data are plotted against years since the initiation of the fishery. For the recruitment perturbation experiment at $F = 0.6$, Z/K estimates from averaged and unaveraged data are plotted against the number of years since (b) a 1-year doubling of recruitment and (c) a 1-year absence of recruitment.

If a population is found to be in disequilibrium, one can either discard the parameter estimates or find some means to reduce or eliminate the disequilibrium bias. Following the second approach, we have examined the utility of a technique often used with such age-based procedures as catch curve analysis to reduce the effects of year class variability. This technique consists of averaging the catch length frequencies over time. The effectiveness of time-averaging depends on the type of perturbation that creates the disequilibrium. In the fishing-up experiment, time averaging has essentially no effect when $F = 0.3$ and only a slight effect when $F = 0.6$ (Fig. 3a). In the recruitment variation experiment, time averaging has a somewhat greater effect, especially during the periods of maximum bias when reductions in bias are as much as 30% (Fig. 3b). Time averaging, however, also increases bias during some periods. Although time averaging is somewhat more effective when disequilibrium is due to recruitment per-

turbation, the 3-year time averaging employed in our study appears to be an ineffective way of reducing disequilibrium bias.

Besides disequilibrium bias, the Wetherall method is subject to two types of biases that may occur even when the population is in equilibrium. The first type of equilibrium bias (herein referred to as Type I or selection bias) is always negative and occurs when fish, equal in length to the smallest l_c values used in the regression, are not fully vulnerable to the fishery (Fig. 4a). Since, in practice, the exact form of the fishery selection curve will rarely be known, choice of the initial l_c value is somewhat arbitrarily based on the shape of a catch length-frequency histogram. Our approach follows Polovina (1989) in choosing the initial l_c to be one size-interval greater than the rightmost mode on a length-frequency histogram. This choice, however, can result in selection bias. For example, when $F = 0.3$, the initial l_c (51 cm) is identical to the length of 95%

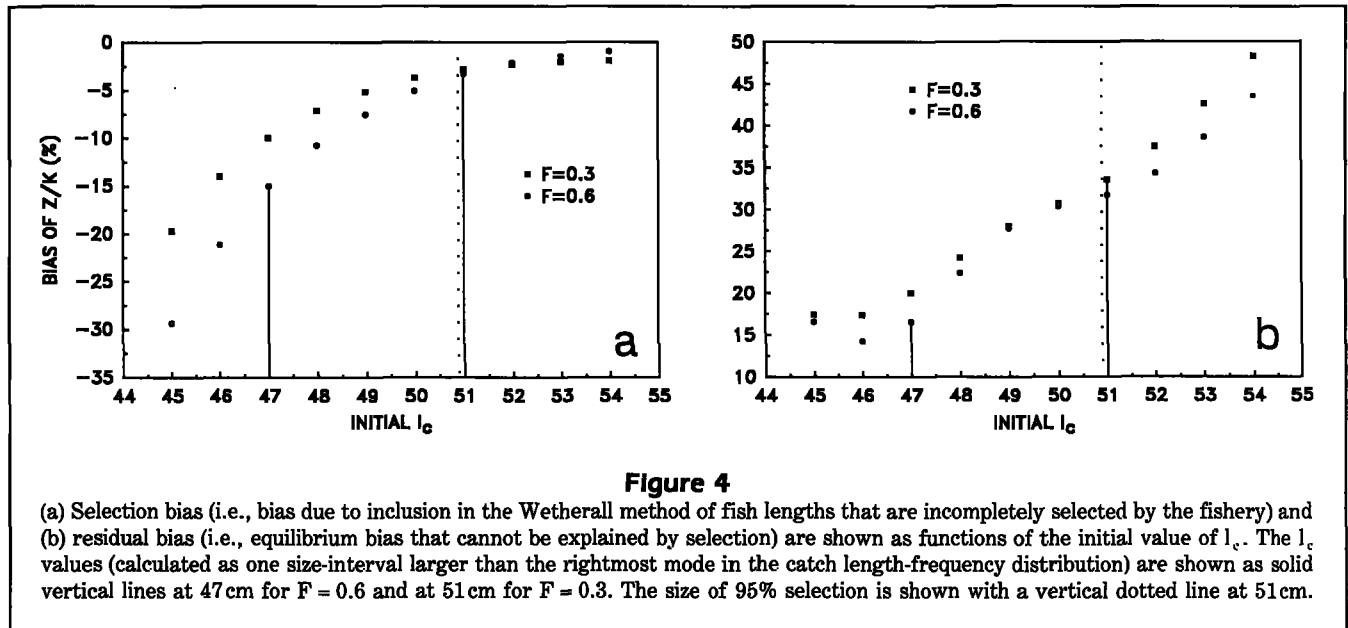


Figure 4

(a) Selection bias (i.e., bias due to inclusion in the Wetherall method of fish lengths that are incompletely selected by the fishery) and (b) residual bias (i.e., equilibrium bias that cannot be explained by selection) are shown as functions of the initial value of l_c . The l_c values (calculated as one size-interval larger than the rightmost mode in the catch length-frequency distribution) are shown as solid vertical lines at 47 cm for $F = 0.6$ and at 51 cm for $F = 0.3$. The size of 95% selection is shown with a vertical dotted line at 51 cm.

selection, and selection bias is small (3%). But when $F = 0.6$, the initial l_c (47 cm) is considerably less than the size at 95% selection, and selection bias is consequently larger (15%).

Since selection bias asymptotically approaches zero with increasing value of the initial l_c (Fig. 4a), one logical way of reducing equilibrium bias would be to choose a larger initial l_c . This strategy, however, would increase the second type of equilibrium bias (herein referred to as Type II or residual bias), because residual bias increases with initial l_c (Fig. 4b). For example, when $F = 0.6$ and the initial $l_c = 47$ cm, residual bias is about 16%. But when $F = 0.3$ and $l_c = 51$ cm, residual bias is 34%. Although the causes of residual bias are unknown, it could be related to either of two factors. First, as Wetherall et al. (1987) demonstrate, the estimates of Z/K and L_∞ are asymptotically unbiased as the sample size increases. Since increasing the size of the initial l_c reduces sample size, it follows that bias will increase. Second, as Laurec and Mesnil (1987) demonstrate, individual variability in growth, which has been incorporated into our simulation model, leads to bias in length-based estimates of mortality.

In the foregoing, we have demonstrated that disequilibrium bias in the Wetherall estimates of Z/K and L_∞ can be large and that time averaging does not remove such

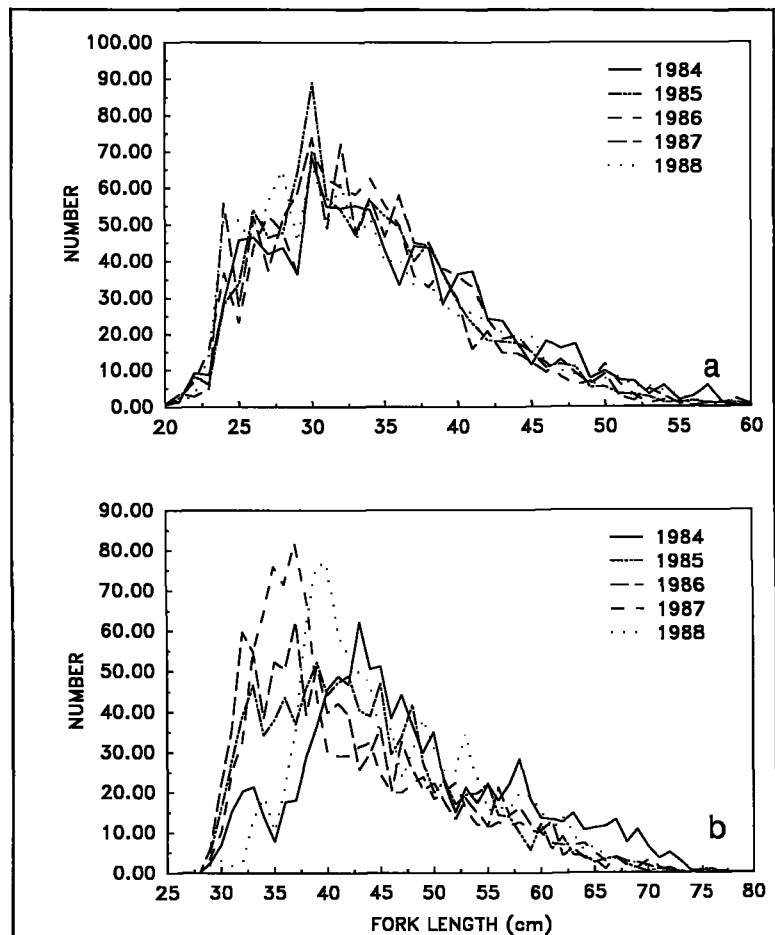


Figure 5

Catch length-frequency distributions for (a) ehu and (b) opakapaka based on 3000 fish samples collected in each year (1984–88).

bias. In light of this, the general utility of the Wetherall method clearly depends upon whether fish populations actually reach equilibrium, or conditions close enough to equilibrium, so that the consequent bias will be small. We believe, however, that equilibrium conditions may prevail more frequently than generally perceived. For example, ehu *Etelis carbunculus*, another deepwater snapper caught by Hawaii's bottomfish fishery, displays considerably less variability in size between years than opakapaka. Chi-square tests, performed on annual length-frequency samples of 3000 fish collected in 1984-88 (Fig. 5), are always not significant for ehu (1986, $P = 0.19$; 1987, $P = 0.40$; 1988, $P = 0.11$) but are always highly significant for opakapaka (1986-88, $P < 0.001$). Therefore, the ehu population is apparently close enough to equilibrium that using the Wetherall method would be appropriate. In general, however, the Wetherall method should be used with extreme caution on new or changing fisheries or species experiencing recruitment fluctuations. In all cases, its use should be preceded by a statistical verification of population equilibrium.

Citations

- Arellano, R.V.**
1989 Estimation of growth parameters in *Panulirus penicillatus* using a Wetherall plot and comparisons with other lobsters. *Fishbyte* 7(2):13-15.
- Conover, W.J.**
1971 Practical nonparametric statistics. John Wiley, NY, 462 p.
- Csirske, J., J.F. Caddy, and S. Garcia**
1987 Methods of size-frequency analysis and their incorporation in programs for fish stock assessment in developing countries: FAO interest in receiving advice. In Pauly, D., and G.R. Morgan (eds.), Length-based methods in fisheries research, p. 1-6. ICLARM Conference Proceedings 13, International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.
- Laurec, A., and B. Mesnil**
1987 Analytical investigations of errors in mortality rates estimated from length distributions of catches. In Pauly, D., and G.R. Morgan (eds.), Length-based methods in fisheries research, p. 239-282. ICLARM Conference Proceedings 13, International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.
- Pauly, D., and G.R. Morgan (editors)**
1987 Length-based methods in fisheries research. ICLARM Conference Proceedings 13, International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait, 468 p.
- Polovina, J.J.**
1989 Density dependence in spiny lobster, *Panulirus marginatus*, in the Northwestern Hawaiian Islands. *Can. J. Fish. Aquat. Sci.* 46:660-665.
- Ralston, S.**
1987 Mortality rates of snappers and groupers. In Polovina, J.J., and S. Ralston (eds.), Tropical snappers and groupers: Biology and fisheries management, p. 375-404. Westview Press, Boulder.
1989 Effect of seasonal recruitment on bias of the Beverton-Holt length-based mortality estimator. *Am. Fish. Soc. Symp.* 6:190-197.
- Ralston, S., and G.T. Miyamoto**
1983 Analyzing the width of daily otolith increments to age the Hawaiian snapper, *Pristipomoides filamentosus*. *Fish. Bull., U.S.* 81:523-535.
- Rawlinson, N.**
1989 Population dynamics of the commercially important bait-fish species *Stolephorus heterolobus* in Solomon Islands. *Fishbyte* 7(1):12-17.
- Wetherall, J.A.**
1986 A new method for estimating growth and mortality parameters from length-frequency data. *Fishbyte* 4(1):12-14.
- Wetherall, J.A., J.J. Polovina, and S. Ralston**
1987 Estimating growth and mortality in steady-state fish stocks from length frequency data. In Pauly, D., and G.R. Morgan (eds.), Length-based methods in fisheries research, p. 53-74. ICLARM Conference Proceedings 13, International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait.