NOAA Technical Memorandum NMFS-AFSC-61

Search Algorithms for Computing Stock Composition of a Mixture from Traits of Individuals by Maximum Likelihood

by
J. Pella, M. Masuda, and S. Nelson

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

March 1996

NOAA Technical Memorandum NMFS

The National Marine Fisheries Service's Alaska Fisheries Science Center uses the NOAA Technical Memorandum series to issue informal scientific and technical publications when complete formal review and editorial processing are not appropriate or feasible. Documents within this series reflect sound professional work and may be referenced in the formal scientific and technical literature.

The NMFS-AFSC Technical Memorandum series of the Alaska Fisheries Science Center continues the NMFS-F/NWC series established in 1970 by the Northwest Fisheries Center. The new NMFS-NWFSC series will be used by the Northwest Fisheries Science Center.

This document should be cited as follows:

> Pella, J., M. Masuda, and S. Nelson. 1996. Search algorithms for computing stock composition of a mixture from traits of individuals by maximum likelihood. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-61, 68 p.

Reference in this document to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

NOAA Technical Memorandum NMFS-AFSC-61

Search Algorithms for Computing Stock Composition of a Mixture from Traits of Individuals by Maximum Likelihood

by
J. Pella, M. Masuda, and S. Nelson

Alaska Fisheries -Science Center
Auke Bay Laboratory
11305 Glacier Highway
Juneau, AK 99801-8626

U.S. DEPARTMENT OF COMMERCE
Ronald H. Brown, Secretary
National Oceanic and Atmospheric Administration
D. James Baker, Under Secretary and Administrator
National Marine Fisheries Service
Rolland A. Schmitten, Assistant Administrator for Fisheries

March 1996

This document is available to the public through:

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

www.ntis.gov

Notice to Users of this Document

This document is being made available in .PDF format for the convenience of users; however, the accuracy and correctness of the document can only be certified as was presented in the original hard copy format.

ABSTRACT

The conditional maximum likelihood method of estimating stock-mixture composition is described for discrete characters. Computer programs were developed for several general-purpose, nonlinear optimization algorithms, specialized to searching for the conditional maximum likelihood estimate (CMLE) ; and their performances were compared for hypothetical and real-world stock mixtures. Measures of performance were search time, failure rate, and stability of CMLE distributions as the criterion for stopping search (guaranteed percent achieved of the maximum of the likelihood function, or GPA) was increased.-,,

-Programs based on the conjugate gradient (with square root transform of stock composition) and expectation maximization algorithms were superior in reliability and speed. Iteratively-reweighted least squares programs produced the most stable CMLE distributions because their terminal GPAs typically exceeded that specified by more than other programs.

THIS PAGE INTENTIONALLY LEFT BLANK

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The practice of evaluating stock contributions to mixed-stock fisheries from the traits of individuals in catches and escapements has become commonplace in North American fisheries for Pacific salmon (Oncorhynchus spp.). Growth patterns on scales (e.g., Fukuhara et al. 1962; Cook and Lord 1978; Marshall et al. 1984), genotypic variation identified by electrophoresis of enzymatic proteins (e.g., Grant et al. 1980, Utter et al. 1987, Shaklee et al. 1990, Bartley et al. 1992, Wilmot et al. 1994), presence of parasites (Moles et al. 1990), and combinations of these characters (Wood et al. 1989, Pella et al. in press) have been used in such assessments. Concerns regarding evaluations of stock contributions in Pacific salmon fisheries from a coastwide viewpoint (Shaklee and Phelps 1990) and recent applications to other species and locales with latest genetic techniques for identifying genotypes (Wirgin et al. 1993, Bowen et al. 1995) illustrate the expanding breadth of this research area.

The preferred method of estimating composition of stock mixtures from traits of individuals is that of maximum likelihood (Fournier et al. 1984, Millar 1987, Pella and Milner 1987). The maximum likelihood estimate is the stock composition (proportions or percentages) for which the likelihood function for traits of individuals observed in the stock and catch samples is maximized with respect to the unknown stock proportions from the potentially contributing, or baseline, stocks composing the catch. Traits may include discrete characters and continuous variables. A discrete trait may be qualitative or quantitative; but, in either case, individuals classify to one of its types or categories (e.g., genotypes of a genetic locus, presence or absence of parasite infection, and numbers of circuli on scales). Individuals sharing the same type for each trait classify to the same multiple character type. Continuous traits are quantitative and categories occur only because of limited accuracy of the measuring device (e.g., length, weight, and distance between points on the body or scale). A continuous trait may be converted to a discrete trait by partitioning its range into intervals, but this change, is not necessary nor even recommended. However, for simplicity, only discrete characters with limited numbers of types are used in this study.

The likelihood maximization considered will treat the relative frequencies of multiple character types in the baseline stocks, the baseline character distributions, as known and equal to their estimates from the baseline samples. Although the baseline character distributions could also be treated as unknowns to be estimated concurrently with stock composition (Fournier et al. 1984, Pella and Milner 1987), this relaxation of assumptions would make computation of mixture composition more complex because unique mixture composition estimates could not be guaranteed (Smouse et al.

1990, Xu et al.. 1994). The maximum likelihood estimate of stock composition may be said to be conditional on the baseline distributions of characters.

Computations for determining the conditional maximum likelihood estimate (CMLE) and its precision are extensive. Just the calculation of the CMLE from the observed catch and baseline samples requires a high-speed computer for virtually all practical applications. Although standard error in the CMLE caused by sampling of the catch can be evaluated with minimal computation by using Millar's infinitesimal jackknife estimator (Millar 1987), any sampling errors in the baseline character distributions are ignored. To evaluate variation in stock composition estimates caused by both sampling of the catch and imperfect knowledge of the baseline character distributions, the bootstrap method (Efron 1982) is recommended (Fournier et al. 1984). Bootstrap resampling of baseline and catch samples with computation of the CMLE for each set of derived samples accounts for sampling error from both sources. The empirical CMLE distributions from bootstrapping are used to estimate CMLE standard errors, and also to provide confidence intervals for stock composition. Efron and Tibshirani (1986) suggest roughly 25' resamplings are sufficient to calculate standard errors for statistics such as the CMLE, but that confidence intervals require 250 to 1,000 resamplings. Clearly, bootstrapping for precision of the CMLE increases the amount of computation of CMLEs manyfold, but other studies also require,.heavy computation of CMLEs.

Numerical experiments are advisable prior to actual catch sampling to judge the feasibility of a research program for determining stock composition with satisfactory accuracy and precision (Pella and Milner 1987). These experiments should include simulated baseline sampling with baseline. character distributions from available baseline samples and simulated sampling of their mixtures of known stock composition. The CMLE of stock composition of any simulated mixture is computed for each set of baseline and mixture samples. The simulated samplings are repeated many times to generate empirical distributions of the CMLEs for comparison with the known stock composition. The decision to begin sampling catches composed of the baseline stocks depends on finding satisfactory precision and accuracy of the CMLEs as judged from their distributions for the simulated situations.

After catch sampling begins, further numerical experiments are prudent to check that fit of the mixture model (Mulligan et al. 1988, Pella et al. in press) is satisfactory. Lack of fit could indicate that important stocks in the catches are not included in the baseline or that baseline character distributions have changed between times of sampling baseline stocks and catches.

Efficient algorithms for finding CMLEs are needed to make good use of computer resources. Simple determination of the CMLE from observed samples requires substantial computation, and bootstrapping, feasibility studies, and tests of fit increase computation of CMLEs manyfold. The demand for computer usage can limit the research. Therefore, a report follows describing the specialization of several general mathematical procedures to CMLE computation, together with an evaluation of their performances when encoded in computer programs written by us. No claim is made that the efficacy of the computer programs cannot be improved, but considerable effort was directed to this end. Therefore, our comparison of program performances provides a measure of what can be achieved by each algorithm.

The following material is organized into four sections. First, generic samples and their likelihood function are described for which the maximizing stock composition must be found. Second, algorithms for finding this CMLE are discussed. Third, the specific samples and performance criteria for comparing the algorithms are described. Finally, performances of computer programs for the algorithms are reported.

THE LIKELIHOOD FUNCTION

The situation modeled is a mixture of c stocks or populations, composed of proportions p_1, p_2, ., ., and p_c from the baseline stocks. The feasible values for $p^T = (p_1, p_2, \ldots p_c)$ (note that any vector is a column and its transpose, denoted by superscript T, is a row) constitute the stock composition space; to be feasible, the elements of p^T must lie between 0 and 1 and their sum must equal 1.

To begin our analysis, a random sample of individuals is obtained from the mixture, and they are compared using a set of characters. [The situation in which characters are missing for some sampled individuals will not be considered here, but the methods below have been extended elsewhere (e.g., Pella 1986) to deal with this circumstance.] Some individuals of the mixture sample may be similar or share the same types for all characters; other individuals may be unique in their character combinations. Suppose H distinct character combinations are observed in the mixture sample and that m_1, m_2, ..., and m_H individuals of the total m individuals composing the mixture sample have these multiple character types. Let the true relative frequency of the hth character combination in the ith stock be g^{hi}. (In practice, the relative frequencies of character combinations in stocks are unknown, and estimates obtained from baseline samples are substituted for unknown values.) Then the relative frequency of the hth character combination in the mixture is

$$\sum_{i=1}^{c} p_i g_{hi} \ .$$

The likelihood function of the mixture sample is that-of a multinomial probability function for the H categories,

$$L(\boldsymbol{p}) = \prod_{h=1}^{H} \left(\sum_{i=1}^{c} p_i g_{hi} \right)^{m_h} \ . \qquad (1)$$

The natural logarithm of the likelihood function, or the support function (Edwards 1972), is

$$\log L(\boldsymbol{p}) = \sum_{h=1}^{H} m_h \log \left(\sum_{i=1}^{c} p_i g_{hi} \right) . \qquad (2)$$

The same values for stock proportions maximize both the likelihood and support functions. The stock composition vector, $p^T = (p_1, p_2, \ldots, p_c)$, which maximizes the likelihood function is called the conditional maximum likelihood estimate (CMLE) and will be denoted as

$$\hat{\boldsymbol{p}}^T = (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_c) \ .$$

Any types of individuals not observed in the mixture sample are not involved in the likelihood function, and we will aggregate such types under an index value of H+l $(m_{H+1} = 0)$, for future reference. In particular, the relative frequency of this category in the ith stock is

$$g_{H+1,i} = 1 - \sum_{h=1}^{H} g_{hi}, \qquad i = 1, 2, \ldots, c . \qquad (3)$$

Next, the meaning and implications of a geometric property of the support function called concavity (Fournier et al. 1984, Millar 1987) are examined: first, to approximate the support function along an arbitrary line in the stock composition space; and second, to bound the amount by which the likelihood, function, evaluated at an arbitrary point in the stock composition space, might increase at the CMLE. The approximation and bound are for use with the algorithms later.

The approximation and bound are developed by considering any two points in the stock composition space, say $p_0 = (p_{1(0)}, p_{2(0)}, \ldots, p_{c(0)})^T$ and $\boldsymbol{p}_1 = (p_{1(1)}, p_{2(1)}, \ldots, p_{c(1)})^T$. Let $\Delta\boldsymbol{p} = \boldsymbol{p}_1 - \boldsymbol{p}_0$. Then the line in the stock composition space passing through p_0 and p_1 can be written as

$$\boldsymbol{p} = \boldsymbol{p}_0 + \lambda \Delta \boldsymbol{p},$$

where the scalar, λ, assumes values such that p is feasible.

Along the line in Equation 4, derivatives of the support function of any order with respect to λ, evaluated at $\lambda = 0$, can be written as

$$a_n = \left.\frac{\partial^n \log L}{\partial \lambda^n}\right|_{\lambda=0}$$

$$= (-1)^{n-1}(n-1)! \sum_{h=1}^{H} m_h \left(\frac{\sum_{i=1}^{c} \Delta p_i g_{hi}}{\sum_{i=1}^{c} p_{i(0)} g_{hi}} \right)^n . \qquad (5)$$

The second derivative, $a_{,,}$ is always negative provided certain conditions on the relative frequencies are met which are described shortly. If these conditions are met, we know from extrema theory of elementary calculus that the support function along the line is concave downward. The term concave *downward* means that a tangent line drawn to a point on the support function and in the plane determined by that point and the line passing through $p_{0.}$ and p_1 would lie above the support function except at the point of tangency. The feasible points, p_0 and $p_{l,}$ were arbitrary, and so the support function is everywhere concave downward in the stock composition space. Fournier et al. (1984) and Millar (1987) demonstrated this concavity property by other means.

The concavity property is certain only if the second derivative, a_2, is negative, which requires that

$$\sum_{i=1}^{c} \Delta p_i g_{hi}$$

cannot equal zero for all h = 1, 2, . . . H. If we write the relative frequencies of the H observed character combinations in each stock as a column of an H by c matrix, G, this requirement can be shown to be equivalent to linear independence of the columns of G. In more tangible terms, column independence for G means that the relative frequencies of the H observed character combinations for any one of the stocks cannot be written as a linear combination of the corresponding relative frequencies for the remaining stocks. As an example, such a failure would occur if the relative frequencies of the H observed character combinations were identical for two stocks.

The implications of the concavity property are far reaching. The existence of a single point, $\hat{\boldsymbol{p}}_{,}$, in the stock composition space that maximizes the likelihood function is guaranteed because of concavity(Fournier et al. 1984, Millar 1987). That a single CMLE occurs is easily seen from the

previous description of the behavior of the support function between arbitrary points in the stock composition space. If two local maxima of the support function occurred, an intervening minimum along the line joining the two local maximizing points in the stock composition space would be necessary. Again we know from elementary calculus that the second derivative of the support function would have to become positive at the minimum; but the second derivative is always negative.

A Taylor's series is often used to approximate a function. The Taylor's series approximation of arbitrary order, N, for the support function at any point, p, along the line passing through two points, p_0 and p_1, is

$$\log L(\boldsymbol{p}) \approx \log L(\boldsymbol{p}_0) + \sum_{n=1}^{N} a_n \lambda^n / n! , \qquad (6)$$

where λ was defined in Equation 4 and a_n was defined in Equation 5. The profile of the support function along an arbitrary line in the stock composition space can be approximated to an arbitrary degree of accuracy by increasing the value of N.

Furthermore, the concavity property can be used to determine how much the value of the likelihood function at an arbitrary point in the stock composition space might be increased at the CMLE. The tangent plane to the support function at p_0 contains all tangent lines at p_0. The support function is concave downward at p_0 in any direction and so lies entirely below any of the tangent lines. Therefore, the level of the tangent plane to the support function at p_0 must equal or exceed the value of the support function everywhere in the stock composition space.

The tangent plane at p_0 is given by

$$\log L_{TAN}(\Delta \boldsymbol{p}) = \log L(\boldsymbol{p}_0) + \sum_{i=1}^{c} \left. \frac{\partial \log L}{\partial p_i} \right|_{\boldsymbol{p}_0} \Delta p_i , \qquad (7)$$

where Δp_i is the change in the ith stock proportion from that at p_0. The first partial derivative of the support function at p with respect to the ith stock proportion is called the gradient or score for the ith stock and is given by

$$s_i(\boldsymbol{p}) = \left. \frac{\partial \log L}{\partial p_i} \right|_{\boldsymbol{p}} = \sum_{h=1}^{H} m_h \frac{g_{hi}}{\sum_{j=1}^{c} p_j g_{hj}} . \qquad (8)$$

This derivative is evaluated at $p_0 = (p_{1(0)}, p_{2(0)}, \ldots p_{c(0)})^T$ in the expression for the tangent plane. The maximum value attainable by $\log L_{TAN}$ over feasible values of Δp occurs when $\Delta_i p_i$ is set to $1 - p_{i^*(0)}$, where i^* is the stock index for which the partial derivative of the support function at p_0 is greatest. The Δp_i for the indices not equal to i* must be set to $-p_{i(0)}$, for the new point to be' in the stock composition space. The corresponding point, p, in the stock composition space is that for which only the i*th stock occurs. The value of $\log L_{TAN}$ for this choice of Δp, which represents an upper limit for the maximum value of the support function, is

$$\log L(\boldsymbol{p}_0) - m_\cdot + \sum_{h=1}^{H} m_h \frac{g_{hi^*}}{\sum_{i=1}^{c} p_{i(0)} g_{hi}} \qquad (9)$$

$$= \log L(\boldsymbol{p}_0) - m_\cdot + s_{i^*}(\boldsymbol{p}_0) .$$

Then we know from Equation 9 that the likelihood function value at p_0 is within $\exp(m - s_{i^*}(p_0))$ 100% of the maximum possible value, that is,

$$\frac{L(\boldsymbol{p}_0)}{L(\hat{\boldsymbol{p}})} \geq \exp\left(m_\cdot - s_{i^*}(\boldsymbol{p}_0)\right) . \qquad (10)$$

This bound for proximity to the maximum value of the likelihood function will be referred to as the guaranteed percent achieved (GPA).

The bound for possible increase in the support function, GPA, established here by elementary argument, generally agrees closely with, but is smaller (less precise) than, that developed through more technical mathematics by Lindsay (1983) (also see Roeder et al. 1989), namely,

$$\frac{L(\boldsymbol{p}_0)}{L(\hat{\boldsymbol{p}})} \geq \left(\frac{s_{i^*}(\boldsymbol{p}_0)}{m_\cdot}\right)^{-m_\cdot} .$$

The difference between the bounds decreases either with increase in sample size, decrease in $s_{i^*(p_0)}$, or magnitude of either bound. For fishery applications with mixture samples of at least 25 fish, the discrepancy (≤ 1 percentage point) between bounds has no practical significance when GPA is 10% or more.

ALGORITHMS

Algorithms described here for computing the CMLE of a set of baseline and mixture-samples begin at an arbitrary guess of the mixture composition. In theory, search progresses toward the CMLE through an iterative sequence of successive feasible guesses, each derived from the preceding guess, and each with corresponding increased likelihood value until the CMLE is found.

Specializations of four general-purpose optimization 'procedures for finding the CMLE are considered. The first three algorithms are called steepest ascent, expectation maximization, and iteratively reweighted least squares; each conducts the search directly in the original stock composition space. The fourth and last algorithm is the conjugate gradient, which we use to search both in the logarithm-transformed stock composition space as was done by Fournier et al. (1984) as well as in the square root-transformed stock composition space.

Steepest ascent, iteratively reweighted least squares, and conjugate gradient perform two steps at each iteration of the search: first, a line of movement from the present guess of the CMLE in the stock composition space or its transform is determined; and second, the optimal distance of movement along this line is approximated. The expectation maximization algorithm does not explicitly separate operations for direction and distance.

Descriptions of the algorithms are outlined next, but details of the computations are left for the appendix.

Steepest Ascent

The simplest algorithm in terms of principles underlying the search for the CMLE is steepest ascent. We believe our application of the algorithm to maximizing the stock mixture likelihood function as well as the specifics of constraining the search to the feasible stock composition space is original. At each point, p, occurring in the search for the CMLE, the direction chosen is that for which the local increase in the support function is greatest, provided that movement in that direction is within the stock composition space. If the boundaries of the stock composition space are not interfering with movement, the direction of greatest increase is a vector pointing along the mean-corrected gradient vector,

$$\boldsymbol{s}^{\circ} = (s_1 - \bar{s}, s_2 - \bar{s}, \ldots, s_c - \bar{s})^T, \tag{11}$$

where s_i is given in Equation 8 evaluated at p and $\bar{s}$ is the arithmetic average of the s_i. The demonstration that the mean-corrected gradient vector is the direction of greatest increase is a simple extension of showing $S = (s_1, s_2, \ldots s_c)^T$ is the direction of greatest increase without constraints (e.g., see Rao 1984); we must add the constraint that the sum of the p_i equals 1.

Boundaries interfere with movement if one or more of-the stock proportions have been driven to near zero during the search and the corresponding gradients remain positive. Boundary stocks are disengaged from the remaining interior stocks at least temporarily in order to continue searching efficiently. Permissible direction of greatest increase is 'along a modified, mean-corrected gradient vector with zeros for boundary stocks; the modified, mean-corrected gradients for the interior stocks are computed as the differences of their gradients in Equation 8 and the arithmetic average of these interior stock gradients. The guess of stock proportions for the boundary stocks is temporarily fixed at the values reached when the boundaries were approached. The search continues parallel to such boundaries until the gradient of the support surface for one of the boundary stocks is greater than the average of the gradients of the interior stocks. Thereafter, the search is allowed to move away from this stock boundary to the interior again.

The present guess of the CMLE and the direction of movement determine a line in the stock composition space. The optimal distance of movement along the line from the present guess in the direction chosen is then determined as that corresponding to the maximum value of the Taylor's series approximation to the support function described earlier. In the approximation Equation 6, the present guess, p, is taken as p_0, and the mean-corrected gradient vector or its modification at boundaries is taken as $\Delta_j p$. The maximum value of the approximation occurs when the first derivative of the approximation with respect to λ equals zero. This derivative is a polynomial equation in λ of order N-l. The minimum, positive, real number solution is chosen as the optimal step size, $\lambda_{, opt}$. The next guess is computed as $p = p_0 + \lambda \Delta_j p$, with λ equal to the smaller of $\lambda_{, opt}$ or λ_{max}, the largest permissible step size before encountering a boundary. The order of the approximation, N, is set equal to the smallest of integers 2, 3, or 4, for which an increase in the support function is found.

Iteratively Reweighted Least Squares

Two versions of the iteratively reweighted least squares algorithm-basic and difference-are possible, depending upon the mode of incorporating the constraint that the stock

proportions sum to one. The basic version makes clear the regression nature of conditional maximum likelihood estimation of stock-mixture composition and is described here -for the first time.

The basic version is obtained by use of the method of Lagrange from calculus for solving extremal problems with constraints., We begin by forming the Lagrange function from the support function,

$$F(\boldsymbol{p}) = \log L(\boldsymbol{p}) + \theta\left(\sum_{i=1}^{c} p_i - 1\right). \tag{12}$$

The scalar constant, 8, is the multiplier for the constraint. The derivatives of the Lagrange function with respect to the stock proportions, can be written as

$$\frac{\partial F(\boldsymbol{p})}{\partial p_j} = s_j + \theta, \qquad j = 1, 2, \ldots, c, \tag{13}$$

where s_j is the gradient from Equation 8 evaluated at p.

The conditions (Kuhn and Tucker 1951) that must be satisfied by the CMLE and guarantee that the solution, $\hat{\boldsymbol{p}}$, satisfying them is the CMLE are as follows:

$$\begin{aligned} &\text{if } \hat{p}_j > 0, \ \left.\frac{\partial F(\boldsymbol{p})}{\partial p_j}\right|_{\hat{\boldsymbol{p}}} = 0 \quad \text{or} \\ &\text{if } \hat{p}_j = 0, \ \left.\frac{\partial F(\boldsymbol{p})}{\partial p_j}\right|_{\hat{\boldsymbol{p}}} \le 0. \end{aligned} \tag{14}$$

The value for the multiplier at the CMLE is obtained by multiplying each derivative of the Lagrange function by the corresponding p_j and summing these products. The sum of these products is $m_\cdot + \theta$ and must equal zero at the CMLE by the Kuhn-Tucker conditions. Therefore, the value of the multiplier at the CMLE equals the negative of $m_\cdot$, and the vector of gradients of the Lagrange function at the CMLE can be written as S-m-1, where 1 is a column vector of c "1"s. The gradient vector, S, for arbitrary p can be written as

$$\boldsymbol{S} = m_\cdot^2 \boldsymbol{G}^T \boldsymbol{R}(\boldsymbol{y} - \boldsymbol{G}\boldsymbol{p}) + m_\cdot \boldsymbol{1}, \tag{15}$$

where

$y_h = m_h / m_\cdot$,
$\boldsymbol{G} = (g_{hi})$, and
$\boldsymbol{R}$ is a diagonal matrix with $r_{hh} = 1/(m_\cdot \sum_{i=1}^{c} p_i g_{hi})$,
$h = 1, 2, \ldots, H+1$; $i = 1, 2, \ldots, c$.

Therefore, we see from Equation 14 that the vector of gradients of the Lagrange function at the CMLE must satisfy

$$m_{\cdot}^{2}\boldsymbol{G}^{T}\boldsymbol{R}(\boldsymbol{y}-\boldsymbol{G}\hat{\boldsymbol{p}}) \leq \boldsymbol{0}. \tag{16}$$

The solution for the CMLE from the vector of gradients of the Lagrange function is related to weighted least squares and Fisher's method of scoring. If equality held for each row in Equation 16 and R were known, the vector of gradients of the Lagrange function at the CMLE would take the form of the system of normal equations for the weighted least squares regression problem: find the unconstrained value of p that minimizes the weighted sum of squares,

$$SS(\boldsymbol{p}) = (\boldsymbol{y}-\boldsymbol{G}\boldsymbol{p})^{T}\boldsymbol{R}(\boldsymbol{y}-\boldsymbol{G}\boldsymbol{p}) , \tag{17}$$

or equivalently, which maximizes- $Q(\boldsymbol{p}) = \boldsymbol{y}^{T}\boldsymbol{R}\boldsymbol{G}\boldsymbol{p} - \tfrac{1}{2}\boldsymbol{p}^{T}\boldsymbol{G}^{T}\boldsymbol{R}\boldsymbol{G}\boldsymbol{p}$. The solution would be

$$\hat{\boldsymbol{p}} = (\boldsymbol{G}^{T}\boldsymbol{R}\boldsymbol{G})^{-1}(\boldsymbol{G}^{T}\boldsymbol{R}\boldsymbol{y}) . \tag{18}$$

Jennrich and Ralston (1978) noted that more generally the gradients of the support function at maximum likelihood estimates for the regular exponentialfamily of distributions, which includes the multinomial distribution used here, take the form of normal equations. The observed relative frequencies of the multiple character types in the mixture sample, y_h, are fit to their expected values,

$$\sum_{i=1}^{c} p_i g_{hi} ,$$

using as weights the inverses of the expected numbers of the multiple character types in the mixture sample. The basic version of the algorithm successively approximates the CMLE by constrained solution to the least squares regression problem described, using for unknown elements of the matrix, R, at any iteration, values computed from the immediately preceding approximation of the CMLE.

Furthermore, if we sought to iteratively determine unconstrained p that maximizes the support function, Equation 2, by Fisher's method of scoring, the solution, Equation 18, also would be a new approximation of the CMLE when R was computed from the previous approximation. The matrix $G^{T}RG$ is Fisher's information matrix, and $G^{T}Ry$ is the score vector. Therefore, we shall be finding successive constrained approximations for the CMLE by the method of scoring. As such, the algorithm uses both the first and second order derivatives of the support function.

The vector of gradients of the Lagrange function satisfies the set of constraints, Equation 16, rather than corresponding equalities. These constraints are more difficult to solve than equalities from which p could be computed directly by Equation 18. Therefore, the constraints are converted to another equation system by introducing nonnegative variables called slack variables in linear programming because they take up the slack of the inequalities to modify them to equalities. The Kuhn-Tucker conditions for' maximizing $Q(p)$ for feasible p are equivaient to the following equation system:

$$\begin{aligned} \sum_{j=1}^{c} a_{ij}p_j - t_i &= b_i, \\ p_i &\geq 0, \\ t_i &\geq 0, \text{ and} \\ p_i t_i &= 0, \\ i &= 1,2,\ldots,c, \end{aligned} \tag{19}$$

where $= \boldsymbol{G^T R G} = (a_{ij})$: c x c ,
$\boldsymbol{B} = \boldsymbol{G^T R y} = (b_i)$: c x 1 , and
$t_1, t_2, \ldots, t_c$ are slack variables.

The constraint that the stock proportions must sum to one was incorporated earlier in the Lagrange function, Equation 12, and is not repeated here.

The solution to this system can be determined by a modification of the simplex algorithm of linear programming *(e.g.,* see Rao 1984), provided an initial feasible solution is available. No such solution will ordinarily be evident. Therefore, we introduce artificial variables z_1, z_2, . . . z_c, which provide an obvious initial solution to an expanded equation system and which are later forced to zero for the original system to be solved. The final problem to solve is the following:

$$\begin{aligned} &\text{maximize} \quad -\sum_{i=1}^{c} z_i \\ &\text{subject to} \\ &\sum_{j=1}^{c} a_{ij}p_j - t_i + b_i z_i = b_i, \\ &p_i \geq 0, \\ &t_i \geq 0, \\ &p_i t_i = 0, \text{ and} \\ &z_i \geq 0, \\ &i = 1,2,\ldots,c. \end{aligned} \tag{20}$$

This system has the obvious initial solution $p_1 = p_2 = \ldots = p_c = 0$, $t_1 = t_2 = \ldots = t_c = 0$, and $z_1 = z_1 = \ldots = z_c = 1$. Now the solution to the original equation system, Equation 16, can be determined by application of the simplex algorithm with one modification: t_i is not permitted to become a basic variable whenever p_i is already a basic variable, and vice-versa, for i = 1, 2, ..., c. This modification ensures the product $P_i t_i = 0$. The optimal solution satisfies the Kuhn-Tucker conditions and is guaranteed to be the CMLE if the support function is concave downward and R is known. The procedure described above is an application of Wolfe's (1959) algorithm for quadratic programming described in Hillier and Lieberman (1967).

The solution found would be the CMLE if R were known. However, the diagonal elements of R depend on the unknown CMLE. Therefore, finding the CMLE requires an iterative procedure. We begin with a guess of the CMLE, compute R from the guess, and then find the CMLE conditioned on the guess. The procedure is prone to overshooting, or passing beyond the CMLE, if the guess is far from the CMLE. Therefore, we use the line search based on the Taylor's series approximation to the support function along the line joining the initial guess with the CMLE conditioned on the guess to find the next guess. The line search is conducted as described under the steepest ascent algorithm. However, now $\Delta \boldsymbol{p}$ equals $p_1 - p_0$ if we call p_0 the initial guess at an iteration and p_1 the CMLE conditioned on that guess. The value of the next guess is used to compute R once again, and the process is continued until convergence.

The second version of the iteratively reweighted least squares algorithm, the difference version, was first described in detail by Pella (1986) and in broad outline by Pella and Milner (1987), and it is easily described now that the first version is familiar. The stock proportion p_c is eliminated to remove the redundancy that the stock proportions sum to one, thereby obviating the Lagrange function. Then the likelihood function is written as

$$\log L(\boldsymbol{p}) = \sum_{h=1}^{H+1} m_h \log \left[\sum_{i=1}^{c-1} p_i g_{hi} + \left(1 - \sum_{j=1}^{c-1} p_j\right) g_{hc} \right]$$

$$= \sum_{h=1}^{H+1} m_h \log \left[\sum_{i=1}^{c-1} p_i d_{hi} + g_{hc} \right],$$

$$\text{where } d_{hi} = g_{hi} - g_{hc}, \qquad (21)$$

$$d_{H+1,i} = \sum_{h=1}^{H} g_{hc} - \sum_{h=1}^{H} g_{hi}, \text{ and}$$

$$p_c = 1 - \sum_{i=1}^{c-1} p_i .$$

Now the gradient can be written in the following form:

$$\boldsymbol{S} = m_{\cdot}^{2}\boldsymbol{D}^{T}\boldsymbol{R}(\tilde{\boldsymbol{y}}-\boldsymbol{D}\boldsymbol{p}),$$

$$\text{where } \tilde{y}_h = \frac{m_h}{m_{\cdot}} - g_{hc}, \quad h = 1, 2, \ldots, H, \tag{22}$$

$$\text{and } \tilde{y}_{H+1} = \sum_{h=1}^{H} g_{hc} - 1.$$

The vectors p and S now have c-l elements, D has H+1 rows and c-l columns, and R is a square H+1 matrix defined earlier.

Again the gradient vector at the CMLE must satisfy the Kuhn-Tucker conditions:

$$m_{\cdot}^{2}\boldsymbol{D}^{T}\boldsymbol{R}(\tilde{\boldsymbol{y}}-\boldsymbol{D}\boldsymbol{p}) \le \boldsymbol{0}. \tag{23}$$

The CMLE is found by converting this system of constraints to a linear programming problem by adding slack and artificial variables:

$$\text{maximize } -\sum_{i=1}^{c-1} z_i$$

$$\text{subject to}$$

$$\sum_{j=1}^{c-1} a_{ij}p_j - t_i + t_c + b_i z_i = b_i, \quad i = 1, 2, \ldots, c-1;$$

$$\sum_{i=1}^{c} p_i = 1,$$

$$p_i \ge 0, \tag{24}$$

$$t_i \ge 0,$$

$$p_i t_i = 0,$$

$$i = 1, 2, \ldots, c; \text{ and}$$

$$z_1 \ge 0,\ z_2 \ge 0,\ \ldots,\ \text{and } z_{c-1} \ge 0.$$

The second form of the algorithm can again be shown to be a constrained form of weighted regression, as well as a constrained form of Fisher's method of scoring.

Expectation Maximization

The expectation maximization algorithm is a general method of solving for maximum likelihood estimates (Dempster et al. 1977). The method was first applied to the maximization of the stock mixture likelihood function,

Equation 1, by Milner et al. (1981). The idea underlying the algorithm in this application is that the CMLE of stock proportions would be obvious and could be calculated explicitly if the stock origins of the individuals in the mixture sample were known. In fact, if the numbers of individuals from each stock in the mixture sample were known, the CMLE would simply be the observed proportions of the mixture sample from each stock, or

$$\hat{p}_i = \frac{\sum_{h=1}^{H} m_h^{(i)}}{m_\cdot}, \qquad i = 1, 2, \ldots, c, \tag{25}$$

where

$m_h^{(i)}$ is the number of individuals of character combination h originating from stock i, and

m is the total number of individuals in the mixture sample.

The information regarding the characters is no longer useful for estimating stock composition-because the numbers from each stock in the mixture sample are sufficient statistics, as can be easily shown by the factorization theorem [e.g., see Lehmann (1983) for a general discussion of the factorization theorem].

Of course, the values of the $m_h^{(i)}$ are unknown, but their expected values conditioned on the present guess of stock composition, $\boldsymbol{p}_{(0)} = (p_{1(0)}, p_{2(0)}, \ldots, p_{c(0)})^T$, can easily be computed. Bayes's theorem provides the probability that an individual of type h is from stock i, and this probability multiplied by m_h equals the expected number from stock i,

$$E[m_h^{(i)}] = m_h \frac{p_{i(0)} g_{hi}}{\sum_{j=1}^{c} p_{j(0)} g_{hj}}. \tag{26}$$

If the expected value at Equation 26 is substituted for the unknown value of $m_h^{(i)}$ in Equation 25, the resulting equation for the new guess of stock proportions, $p_{(1)} = (p_{1(1)}, p_{2(1)}, \ldots p_{c(1)})^T$, is

$$p_{i(1)} = \frac{1}{m_\cdot} \sum_{h=1}^{H} m_h \frac{p_{i(0)} g_{hi}}{\sum_{j=1}^{c} p_{j(0)} g_{hj}} = \frac{1}{m_\cdot} p_{i(0)} s_i(\boldsymbol{p}_{(0)}), \qquad i = 1, 2, \ldots, c, \tag{27}$$

where s_i was defined at Equation 8. Next, the numerical values of the old guess ($p_{(0)}$) are replaced by those of the new guess ($p_{(1)}$) for use on the right-hand side of Equation 27 to compute yet another new guess. The process is guaranteed to converge

to the CMLE provided the initial guess comprises only positive (>0) proportions and the concavity property holds for the likelihood function (Redner and Walker 1984).

Description of three algorithms that search directly in the stock composition space is complete, and next considered is an algorithm that searches in transformations of that space.

Conjugate Gradient

The search path of the steepest ascent algorithm tends to zigzag in the parameter space rather than follow a direct route to the maximizing point. The reason for this zigzag behavior is that the search proceeds stepwise from each point of the path in the direction of the local gradient, reaching a new point corresponding to a local maximum along that direction;- The local gradient at this new point must be perpendicular to the preceding direction of movement; otherwise, the new point could not have been a local maximum. Therefore, the next direction chosen from this new point is at right angles to the preceding direction. This search pattern can be inefficient, and reducing zigzagging to a more direct path would presumably speed the search.

The conjugate gradient algorithm of Fletcher and Reeves (1964) also uses only the gradient to conduct the search but is more efficient than steepest ascent if the likelihood surface is approximately quadratic. Instead of using only the local gradient at the present point to determine the new direction, the conjugate gradient algorithm uses a linear sum of local gradients at all points so far included in the search. In effect, the direction chosen is moderated by previous directions. Rounding errors, however, accumulate in the linear sum of local gradients, so it is necessary to restart the algorithm periodically; Fletcher and Reeves (1964) recommend a restart every c+l iterations if c baseline stocks occur.

Fournier et al. (1984) first used the conjugate gradient algorithm to search for the CMLE and noted that the search is theoretically unconstrained in the logarithm-transformed stock composition space:

$$u_i = \log(p_i), \quad i = 1, 2, \ldots, c. \tag{28}$$

Although the transformation has the theoretical advantage that the search in the u-space is unconstrained, in practice, guesses in the u-space must be constrained because the logarithms of stock proportions near zero become negative and arbitrarily large. Fournier et al. (1984) also observed that in retransforming to the stock composition space,

$$p_i = \exp(u_i) / \sum_{j=1}^{c} \exp(u_j), \quad i = 1, 2, \ldots, c, \tag{29}$$

the u_i can be scaled arbitrarily to produce the same point, p. Therefore, the constraint was imposed that

This constraint was effected by adding to the support function a penalty term that increased with the discrepancy from the equality. The search was conducted in the u-space; and at termination, the retransformation, Equation 29, was used to determine the CMLE.

In our implementation of the conjugate gradient algorithm with logarithm-transformation of the stock composition space, the search is begun with a feasible guess p, this guess is transformed to u by Equation 28, a new value for u is found, and the new u is transformed back to a feasible p by Equation 29 before the next iteration of the search. The process obviates any constraints- in the u-space other than limiting magnitude of large negative u-values corresponding to stock proportions near zero.

Gradients at u in the transformed-space can be expressed in terms of p in the untransformed space as follows:

$$\tilde{s}_i(\boldsymbol{u}) = \left.\frac{\partial \log L}{\partial u_i}\right|_{\boldsymbol{u}} = -p_i m_\cdot + \sum_{h=1}^{H} m_h \frac{p_i g_{hi}}{\sum_{j=1}^{c} p_j g_{hj}} = -p_i(m_\cdot - s_i(\boldsymbol{p})) . \qquad (30)$$

The direction of movement from u equals a linear sum of current and preceding gradient vectors; however, if any component of u is less than -20, the corresponding direction of movement, if negative, is set to zero to constrain magnitude of large negative u-values (see Appendix).

The line in the transformed stock composition space passing through arbitrary point u_0 with direction Au is

$$\boldsymbol{u} = \boldsymbol{u}_0 + \lambda \Delta \boldsymbol{u} . \qquad (31)$$

Along this line the nth order derivative of the support function with respect to λ, evaluated at $\lambda = 0$, will be denoted as b_n. The line search for the optimum step size from u_0 in the direction, Au, is accomplished by an Nth ($N \leq 4$) order Taylor's series approximation for the support function at u, along the line. The approximation is

$$\log L(\boldsymbol{u}) \approx \log L(\boldsymbol{u}_0) + \sum_{n=1}^{N} b_n \lambda^n / n! . \qquad (32)$$

The first four derivatives of the support function with respect to λ, evaluated at $\lambda = 0$, are as follows:

$$b_1 = \left.\frac{\partial \log L}{\partial \gamma}\right|_{\gamma=0} = \sum_{h=1}^{H} m_h \frac{w_h^{(1)}}{v_h} - m_\cdot t^{(1)},$$

$$b_2 = \left.\frac{\partial^2 \log L}{\partial \gamma^2}\right|_{\gamma=0} = \sum_{h=1}^{H} m_h \left[\frac{w_h^{(2)}}{v_h} - \frac{w_h^{(1)^2}}{v_h^2}\right] - m_\cdot [t^{(2)} - t^{(1)^2}],$$

$$b_3 = \left.\frac{\partial^3 \log L}{\partial \gamma^3}\right|_{\gamma=0} = \sum_{h=1}^{H} m_h \left[\frac{w_h^{(3)}}{v_h} - \frac{3 w_h^{(2)} w_h^{(1)}}{v_h^2} + 2\left(\frac{w_h^{(1)}}{v_h}\right)^3\right] - m_\cdot [t^{(3)} - 3 t^{(2)} t^{(1)} + 2 t^{(1)^3}],$$

$$\text{and } b_4 = \left.\frac{\partial^4 \log L}{\partial \gamma^4}\right|_{\gamma=0} = \sum_{h=1}^{H} m_h \left[\frac{w_h^{(4)}}{v_h} - \frac{4 w_h^{(3)} w_h^{(1)}}{v_h^2} + \frac{12 w_h^{(2)} w_h^{(1)^2}}{v_h^3} - 3\left(\frac{w_h^{(2)}}{v_h}\right)^2 - 6\left(\frac{w_h^{(1)}}{v_h}\right)^4\right]$$

$$- m_\cdot [t^{(4)} - 4 t^{(3)} t^{(1)} + 12 t^{(2)} t^{(1)^2} - 3 t^{(2)^2} - 6 t^{(1)^4}],$$

where

$$w_h^{(i)} = \sum_{j=1}^{c} p_j \Delta u_j^i g_{hj}, \qquad t^{(i)} = \sum_{j=1}^{c} p_j \Delta u_j^i, \text{ and } \qquad v_h = \sum_{j=1}^{c} p_j g_{hj}. \tag{33}$$

The maximum of this approximation occurs when its first derivative with respect to λ equals zero. This derivative is a polynomial equation in λ of order N-l. The minimum, positive, real solution is chosen as the optimal step size,, λ_{opt}, and the next guess is computed as $\mathbf{u_0} + \lambda_{opt} \cdot \Delta \mathbf{u}$. The order of the approximation, N, is set equal to the smallest of 2, 3, or 4, for which an increase in the support function is found.

We also implemented the conjugate gradient algorithm using a square root transformation of the stock composition space. A square root transformation avoids the necessity of constraining the solution away from the boundaries in the transformed space. The only constraint on the square root parameters is that the sum of their squares equals 1, and this constraint is automatically satisfied by the transformation and retransformation at each iteration of the search. If we let

$$u_i = \sqrt{p_i}, \qquad i = 1, 2, \ldots, c, \tag{34}$$

the transformation is defined for all feasible values of the stock composition space including zero. The retransformation,

(35)

again assures that only feasible points of the stock composition space are included. The gradients corresponding to Equation 30 for the square root transformation are as follows:

$$\tilde{s}_i(\boldsymbol{u}) = -2\sqrt{p_i}\, m_\cdot + 2\sqrt{p_i} \sum_{h=1}^{H} m_h \frac{g_{hi}}{\sum_{j=1}^{C} p_j g_{hj}} = -2\sqrt{p_i}\,(m_\cdot - s_i(\boldsymbol{p}))\,. \qquad (36)$$

The derivatives for the line search corresponding to Equation 33 are as follows:

$$b_1 = 2\sum_{h=1}^{H} m_h \frac{w_h^*}{v_h} - 2m_\cdot t^*,$$

$$b_2 = 2\sum_{h=1}^{H} m_h \left[\frac{\sum_{i=1}^{C} \Delta u_i^2 g_{hi}}{v_h} - 2\left(\frac{w_h^*}{v_h}\right)^2 \right] - 2m_\cdot \left[\sum_{i=1}^{C} \Delta u_i^2 - 2t^{*2} \right],$$

$$b_3 = 4\sum_{h=1}^{H} m_h \left[-3w_h^* \frac{\sum_{i=1}^{C} \Delta u_i^2 g_{hi}}{v_h^2} + 4\left(\frac{w_h^*}{v_h}\right)^3 \right] + 4m_\cdot \left[3t^* \sum_{i=1}^{C} \Delta u_i^2 - 4t^{*3} \right], \text{ and}$$

$$b_4 = 12\sum_{h=1}^{H} m_h \left[8w_h^{*2} \frac{\sum_{i=1}^{C} \Delta u_i^2 g_{hi}}{v_h^3} - \left(\frac{\sum_{i=1}^{C} \Delta u_i^2 g_{hi}}{v_h} \right)^2 - 8\left(\frac{w_h^*}{v_h}\right)^4 \right]$$
$$-12m_\cdot \left[8t^{*2} \sum_{i=1}^{C} \Delta u_i^2 - \left(\sum_{i=1}^{C} \Delta u_i^2 \right)^2 - 8t^{*4} \right],$$

$$\text{where } w_h^* = \sum_{i=1}^{C} \sqrt{p_i}\, \Delta u_i g_{hi}, \quad t^* = \sum_{i=1}^{C} \sqrt{p_i}\, \Delta u_i, \text{ and } v_h = \sum_{i=1}^{C} p_i g_{hi}\,. \qquad (37)$$

The line search in the square root-transformed space is conducted-nearly the same as in the logarithm-transformed space; however, elements of the direction vector are not constrained.

A final caveat in searching for the CMLE is required when the type composition of the mixture and baseline samples requires certain of the baseline stocks be present in the

mixture. No algorithm can include points p in the search for which mixture proportions of such stocks equal zero. For example, if a type in the mixture sample is unique to one baseline stock, that stock's proportion in the mixture must never be allowed to equal exactly zero. Assuming the initial guess of $\hat{\boldsymbol{p}}$ is an interior point of the stock composition space, neither the steepest ascent, the expectation maximization, nor the conjugate gradient with logarithm transform of p will allow elements of p to equal exactly zero. Unless constrained, the search by the other algorithms does allow elements of p to equal zero, and so verification that all mixture types remain possible with each new guess of $\hat{\boldsymbol{p}}$ is required. If a point in the search results in a zero probability for one or more types of the mixture sample, the search must be drawn an arbitrarily short distance back from the boundary toward the previous permissible guess. Details are provided in the appendix for the applicable algorithms.

This concludes the description and motivation for the various algorithms. Next, methods for comparing performances of computer programs based on the algorithms are described. The computer programs, sketched in the appendix, will be denoted by abbreviations for the algorithms: STEEP for steepest ascent; EM for expectation maximization; GIRLS-B and GIRLS-D for the basic and difference versions of iteratively reweighted least squares, respectively; and CONJA-L and CONJA-S for conjugate gradient with logarithm or square root transformation, respectively.

METHODS

Two experiments were conducted to evaluate the computer programs. First, their performances were explored under hypothetical situations allowing complete control over numbers of baseline stocks and their character distributions as well as size and stock composition of mixture samples'. Second, a subset of the programs from the first experiment was applied to actual baseline and mixture samples reported in fisheries literature. The algorithms and measures of their performances were incorporated into previous stock-mixture analysis software (Masuda et al. 1991) so as to determine average performances on repeated application to any set of baseline and mixture samples perturbed by bootstrap resampling (Efron 1982). Computations were performed on a COMPAQ DESKPRO 486/50[1] under MS-DOS 6, using executable modules created by Lahey's F77L-EM/32 Fortran language and the Lahey/Ergo OS/386 Operating System.

'Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

Performance of any computer program depends on search completeness specified, or the degree to which the likelihood function is to be maximized. In theory, search completeness is easy to define by referring a found value of the likelihood function to its 'maximum, but in practice the maximum of the likelihood function is generally never known. Instead, the GPA at Equation 10 is the only absolute measure of search completeness and generally a GPA of 100% is not achievable because of rounding errors caused by limited numerical precision of the computer.

Three criteria are used to compare performances of the programs during computation of CMLEs for any set of related baseline and mixture samples within the experiments: speed, failure rate, and stability (with respect to increase in specified GPA) of CMLE distributions. Search speed is measured by processor time required to achieve the specified GPA. Search failure rate is measured by the number and percentage of design cells within an experiment or the number and percentage of trials within cells for which a program is unable to achieve the specified GPA. Stability of CMLE distributions is quantified from descriptive statistics of CMLE distributions (averages, standard deviations, and 2.5 and 97.5 percentiles) resulting from a range of specified GPAs; these descriptive statistics are commonly used to evaluate statistical bias and precision from bootstrapping in actual assessments of mixture composition. Ideally, specified GPA would not affect CMLB distributions produced by a program; but practically, the distributions change as higher GPAs are specified. Stability measures based on. the statistics differ between the two experiments and are described with the experiments below.

Total processor time used in searching comprised effective search time and stopping-rule time, Progress of the search toward the maximum of the likelihood function, as indicated by the GPA, must be checked to determine whether the specified GPA has been met and the search can be stopped. Effective search time includes only processor time required to perform computations of the search algorithm. Stopping-rule time is the additional processor time needed to evaluate the GPA and test whether the specified GPA has been achieved. In the experiments next described, the stopping rule was evaluated with each new point p along the search. The symbol T_α denotes the total processor time used in the search by a program to achieve a GPA of α% (GPAs specified were 10%, 50%, and 90% and are termed *standard* values) of the maximum of the likelihood function; the corresponding notation for effective search time required is E,. Time was recorded in units of centiseconds (cs).

Experiment I

The first experiment formed a design with five factors: 1) algorithms; 2) number of stocks in the baseline-5, 15, or 50; 3) baseline stock differences, small or large; 4) mixture sample size-50, 150, 250, or 500 individuals; and 5) specified standard GPAs-10%, 50%, and 90%. Ranges for number of stocks in the baseline and for mixture sample size include cases of practical interest for fisheries applications. A single set of original random baseline and mixture samples per cell of this design was generated by methods described below. Then, this set of original samples was bootstrapped, or resampled with replacement, 25 times to create sets of bootstrap samples with sizes equal to those of the original set. The programs were allowed to search for the CMLEs of each of the 25 bootstrap sets until either the specified GPA had been achieved or effective search time exceeded 5 minutes (5 minutes was greater than fivefold the average effective search time required by successful algorithms for the most demanding experimental conditions). The average effective (E_{10}, E_{50}, E_{90}) and average total (T_{10}, T_{50}, T_{90}) search times required by each program to find the 25 CMLEs for each bootstrap sample set were determined for corresponding specified standard GPAs if the program achieved specified GPA.

Distributions of the 25 CMLEs for any cell of the design that were obtained at standard GPAs were compared with those at GPA of 99% (if the search achieved GPA of 99%). Denote by $ave_\alpha(\hat{p})$ and $s_\alpha(\hat{p})$ the vectors of stock averages and standard deviations of estimated proportions, respectively, from the 25 CMLEs when specified GPA equaled α%. The measures of stability used are the maxima among stocks of the differences between the vectors of averages and standard deviations at specified standard GPAs and corresponding vectors at GPA of 99%, that is, maximum coordinates of 1) $\Delta ave_\alpha = ave_\alpha(\hat{p}) - ave_{99}(\hat{p})$ and 2) $\Delta s_\alpha = s_\alpha(\hat{p}) - s_{99}(\hat{p})$ for α = 10, 50, and 90. The two greatest discrepancies will be denoted as $\max(\Delta ave_\alpha)$ and $\max(\Delta s_\alpha)$, respectively.

The original set of random baseline stock and mixture samples was generated by computer. Five hypothetical, independent genetic loci, each with two alleles labeled (for discussion) A and a, B and b, ... and E and e, were the basis of estimation. The relative frequencies of upper case alleles A, B, C, D, and E (or the complementary relative frequencies of lower case alleles a, b, c, d, and e) were identical for all loci of a stock, but these relative allele frequencies (RAFs) differed among stocks. Neither the number of characters nor their independence is significant to the comparison of program performances, but the number of baseline stocks and degree of character differentiation among them affect program performance.

Original baseline samples were generated for 5, 15, and 50 stocks at two levels of character differences among the stocks, as were original mixture samples of subsets of the baseline stocks. The baselines of stocks with greater character differences will be termed diverse baselines, and those with lesser character differences, similar baselines. RAFs used to generate original baseline and mixture samples for five stocks are specifically provided (Table 1). The RAFs for 15 and 50 stocks can be succinctly reported by developing a notation. The stock-specific and loci-shared RAFs for five stocks of the diverse and similar baselines (Table 1) can be condensed as follows: ranges = 0.8, 0.6; first stock RAFs = 0.2, 0.4; diverse baseline intervals = 0.2 (4); similar baseline intervals = 0.1 (2), 0.2 (2). In this notation, "ranges = x, y" provides the range of stock RAFs used for diverse (range = x) and similar (range = y) baselines. The notation, "first stock RAFs = x, y", refers to the first stock RAFs for diverse (RAF = x) and similar (RAF = y) baselines. Also "diverse (or similar) baseline intervals = x_1 (y_1), x_2 (y_2) ,..., x_n (y_n)" provides the n distinct differences between RAFs (x_is) of successive adjacent stock pairs and the numbers of adjacent stock pairs with those differences (y_is), beginning with the difference between the first and second stocks and ending with that between the penultimate and last stocks. The condensed description for 15 stocks was as follows: ranges = 0.933, 0.7; first stock RAFs = 0.067, 0.3; diverse baseline intervals = 0.067 (9), 0.066 (5); similar baseline intervals = 0.033 (5), 0.034 (2), 0.067 (5), 0.066 (2). The condensed description for 50 stocks was as follows: ranges = 0.98, 0.74; first stock RAFs = 0.02, 0.26; diverse baseline intervals = 0.02 (49); similar baseline intervals. = 0.01 (24), 0.02 (25).

An original baseline sample of 100 fish was drawn for each stock of each cell of the experimental design. The genotype of each fish was generated by randomly and independently drawing two alleles at each locus and repeating such draws independently among the five loci. The probabilities of drawing the various alleles (A, a, B, b, ..., E, e) were the RAFs in the preceding paragraph. Among the 100 fish, the random numbers of the two alleles at each of the five loci constituted the original baseline sample.

Mixtures were constructed with some baseline stocks missing because commonly the CMLEs of real-world applications indicate baseline stocks are missing from mixtures. Mixtures composed of equal contributions by 3 of 5, 8 of 15, or 25 of 50 baseline stocks underlaid the original mixture samples. Alternating stocks from the baseline tables (e.g., Table 1 and unreported analogues for 15 and 50 baseline stocks), subsequent to the first stock, were absent.

Original mixture samples of several sizes ranging from 50 to 500 individuals were drawn, depending on the cell of the experimental design. An original mixture sample was drawn in two stages: 1) a multinomial sample of numbers from each stock was drawn from the specified stock mixture, and 2) the genotype of each mixture individual, given its stock origin from the previous step, was generated from the appropriate baseline stock RAF as had been done to generate the original baseline stock samples.

The original set of random baseline stock and mixture samples of each cell were resampled with replacement 25 times to create derived sets of bootstrap samples for comparing program performances. Each bootstrap sample set was used to estimate stock composition of the mixture by each program. The correct genetic model of independence of loci and Hardy-Weinberg equilibrium was used to compute by stock the relative frequencies of genotypes observed in the mixture sample [i.e., the g_{hi} in Equation I], using the random RAFs observed in the bootstrap baseline samples.

All programs were provided the same initial guess (with all stocks present) for any maximization trial. In theory, time required by any of the algorithms to find the CMLE at specified GPA may be affected by the starting guess, but eventual success should be certain. The initial guess was far-removed from the stock composition of the original mixture sample (and from the CMLE anticipated) in order to test capability of the programs to locate the CMLE. Initial guesses attributed most of the contribution to the first stock. For five stocks, the guess was 0.6 for the first stock and 0.1 for each of the other stocks. For 15 stocks, the guess was 0.86 for the first stock and 0.01 for each of the other 14 stocks. For 50 stocks, the guess was 0.51 for the first stock and 0.01 for each of the other 49 stocks.

Experiment II

In the second experiment, three of the six programs-CONJA-S, EM, and GIRLS-B-were selected for further comparison based on their diversity of search methods and performance in Experiment I. The three programs computed CMLEs for three real mixed-fishery data sets comprising baseline and mixture samples (Table 2) : 1) Columbia River chinook salmon (0. tshawytscha); 2) Yukon River chum salmon (O. keta); and 3) West Coast sockeye salmon (0. *nerka).* Characters observed for individuals from the first two data sets comprised electrophoretic observations on allozymes, and those of the third data set included allozymes, a brain parasite, and freshwater age of individuals. Numbers of stocks in the baselines ranged from 14 to 73; numbers of characters ranged from 4 to 19; and size of mixture samples

ranged from 197 to 1,597. Further details of the data sets can be found in the original publications (Table 2).

One thousand bootstrap CMLEs were computed from resamplings of baseline and mixture samples of each data set for each of the standard GPAs specified. For each mixture analysis, all three programs were provided the same initial guess of equal proportions from the baseline stocks. Search was terminated if either the specified GPA was achieved or effective search time exceeded 5 minutes. Performances were compared with many of the same criteria used in Experiment I: failure rate of a program to achieve specified GPA, and average effective and total search times needed per bootstrap resampling. Average GPA achieved per resampling was also recorded. Several characteristics of the bootstrap distributions of estimated individual stock or stock group proportions were examined including means, standard deviations, and the upper and lower 95% confidence bounds (i.e., the 2.5 and 97.5 percentiles of the bootstrap distribution) from the percentile method (Efron 1982). Stock group proportions were estimated by summing estimates of individual baseline stock percentages over stocks comprising groups (the basis for grouping was similarity of baseline character distributions) defined in the source publications.

RESULTS AND DISCUSSION

Generally, for each algorithm and set of baseline and mixture samples, the increase of the likelihood function between successive guesses decreased as the search continued, as did the distance between successive guesses. The gradients of the support function in feasible directions all tended to zero as the maximum was approached. (Gradients in nonfeasible directions could remain large and even positive, implying greater likelihood function values could have been obtained with infeasible choices for p.) In theory, search progress of increasing likelihood function values would be uninterrupted until the CMLE was found; in practice, rounding errors intervened and the search sometimes regressed temporarily to a reduced value for the likelihood function. The search led only to a neighborhood of the CMLE beyond which numerical precision used in computations limited further progress.

Typical time series of support, function values achieved by the programs were initially rapidly increasing curves with a following asymptote (i.e., the rate of increase in value of the support function during the search decreased as the search for the CMLE advanced) (Fig. 1). In the example illustrated, EM initially achieved greater values for the support function (Fig. 1, left) than did GIRLS-B until about 50 centiseconds (cs) of total search time. Shortly thereafter, GIRLS-B surpassed support function values found by EM (Fig. 1, right).

The critical times, T_{10}, T_{50}, and T_{90}, at which 10%, 50%, and 90% of the maximum of the likelihood function were first known from the GPA (Equation 10) to be achieved by EM are illustrated as well (Fig. 1, right).

GIRLS-B achieved a final support function value of -2182.283917 after 138 cs; at this point, the GPA showed for the first time that the program was within 99% of the maximum of the likelihood function. Similarly, EM continued to find increasing values of the support function until 451 cs, at which time the support function value was slightly less than that found by GIRLS-B, -2182.284001. At this point in the search, the GPA showed for the first time that EM was within 99% of the maximum of the likelihood function. GIRLS-B required 6 iterations whereas EM required 156 iterations during the search. Stepwise increases in support function value and GPA by EM during the search were smaller than for GIRLS-B, so EM could detect GPA achievement of 99%. at a smaller value of the support function than GIRLS-B.

The maximum support value for this example was found, by STEEP (performance not illustrated) after 403 cs and equaled -2182.283773; at this point GPA showed STEEP was only within 90% of the maximum of the likelihood function. Therefore, although a higher value of the likelihood function was found by STEEP than by GIRLS-B or EM, knowledge that this value was nearer than 99% of the maximum possible value could not be obtained without comparison to results of the other programs. Subsequent rounding errors also prevented STEEP from achieving a self-determined GPA of 99%. The measures of performance illustrated by this example are next used in program evaluation by Experiment I and Experiment II.

Experiment I

Interacting effects of number of baseline stocks and their character differences, GPA, and size of the mixture sample were evident either from occurrence of failures to achieve specified GPA or from processing times required by the -programs. Failure rates by programs to achieve specified GPA differed substantially and ranged from no failures by EM and CONJA-S to numerous failures by CONJA-L (Table 3). Excluding CONJA-L, cell failures were rare at GPAs of 10% and 50%; only GIRLS-D had such a cell failure (Table 3, upper half), and that cell failure was due to a single failure among the 25 resamplings for that cell (Table 3, bottom half). Cell failures increased at GPA of 90% for GIRLS-D, GIRLS-B, and STEEP, but GIRLS-B had at most one resampling failure per failing cell. Among programs with cell failures, GIRLS-D, STEEP, and CONJA-L (but not GIRLS-B) tended to fail more often with increase in number of baseline stocks, and only CONJA-L (Table 3, 15 stocks) showed much evidence of more frequent

failure for similar baseline than diverse baseline. Failures by the programs among the 1,800 resamplings for the entire experiment were generally rare: CONJA-S, 0 (0%); EM, 0 (0%); GIRLS-B, 2 (0.1%); GIRLS-D, 12 (0.7%); and STEEP, 78 (4.3%). (CONJA-L computations, unlike those of the other programs, were interrupted in numerous cells due to arithmetic errors or overflows; so its experiment resampling failures were not determined.)

An increase in mixture sample size generally increased effective search time needed by EM, STEEP, CONJA-L, and CONJA-S to achieve specified GPA for diverse and similar baselines (Figs. 2 and 3). Regardless of baseline kind, the GIRLS programs also required increasing effective search time with increase in mixture sample size to achieve GPA if 5 stocks were present in the baseline, but time required remained roughly constant or declined with increase of mixture sample size for 15 or 50 stocks (Figs. 2 and 3). First-place finishes or ties (finish times were ties if they differed by less than 1 centisecond) for minimum effective time required to complete the 25 bootstrap resamplings for each of the 72 experimental cells show that the CONJA-S and EM programs performed best most frequently (Table 4). Numbers of first-place finishes by program were as follows: CONJA-S, 48; EM,, 23; GIRLS-B, 3; GIRLS-D, 3; STEEP, 0; and CONJA-L, 0. EM was generally best at combinations of lower number of baseline stocks, lower mixture sample size, and lower GPA. CONJA-S superseded EM as number of baseline stocks, mixture sample size, and GPA increased. The GIRLS-B and GIRLS-D programs showed in first-place only at the largest mixture sample size and then only at the higher GPAs and lower baseline stock numbers.

The outcome for total search **time** is in broad agreement with that for effective search time but with CONJA-S becoming even more dominant. Increase in mixture sample size generally increased total search time required by EM, STEEP, CONJA-L, and CONJA-S to achieve specified GPA for either baseline kind (Figs. 4 and 5). For either baseline kind, the GIRLS programs also required increasing total **time** with increase in mixture sample size to achieve GPA if 5 stocks were present in the baseline, but time required remained roughly constant or declined with increase of mixture sample size for 15 or 50 stocks (Figs. 4 and 5). First-place finishes or ties for minimum total time required to complete the 25 bootstrap resamplings for each of the 72 experimental cells show that the CONJA-S and EM programs performed best **most** frequently (Table 5). Numbers of first-place finishes by program were as follows: CONJA-S, 62; EM, 12; STEEP, 5; GIRLS-D, 5; GIRLS-B, 5; and CONJA-L, 1. CONJA-S was generally best with a higher number of baseline stocks, higher mixture sample sizes, and higher GPA. EM was displaced by CONJA-S in total search time performance from experimental combinations intermediate to the

two algorithms' strengths in effective search time performance. The GIRLS-B and GIRLS-D programs again showed in first-place only at the largest mixture sample size. STEEP performed well only at the smaller mixture sample sizes.

Discrepancies between the bootstrap distributions of stock composition estimates (average and standard deviation) at specified standard GPAs as compared to 99% differed more among programs when mixture samples were small (m = 50) than when they were large (m = 500); necessarily, maximum discrepancies decreased with increase in specified GPA. Relative performance of programs was similar for diverse (Tables 6 and 7) and similar baselines (Tables 8 and 9). Consistently, smaller discrepancies by the GIRLS programs were evident when the GPA was specified at 10% especially, but also at 50%. Among the remaining programs, CONJA-S generally had the smallest discrepancies, and EM, the largest. Differences among programs in stability can be explained largely by the differences in their stepwise rates of convergence and terminal GPAs: GIRLS-B and GIRLS-D, which satisfy the specified GPA in relatively few steps with large GPA increases between steps, usually exceed by greater amounts the specified GPA at the end of search than do programs that require higher numbers of steps having smaller GPA increases between steps. Differences in maximum discrepancies among all programs at specified GPA of 90% were small.

Experiment II

Effective and total search time required for the three real-world data sets by CONJA-S, EM, and GIRLS-B differed substantially (Table 10). CONJA-S was the fastest program in effective and total search time at all specified GPAs for all three data sets. Comparing speeds of the other two programs, EM was much faster than GIRLS-B for two of the data sets (Yukon River chum salmon and West Coast sockeye salmon), but slower for the other set (Columbia River chinook salmon). Relative speed of the programs depended on specified GPA and search time measure, either effective or total; and speed ranges given next reflect this dependence. For the Columbia River chinook salmon data, CONJA-S was roughly two to seven times as fast as EM (CONJA-S required between approximately one-half and one-seventh the amount of effective or total search time used by EM, depending on specified GPA) and about twice as fast as GIRLS-B. For the Yukon River chum salmon data, CONJA-S was almost two to four times as fast as EM and 50 to 85 times as fast as GIRLS-B. For West Coast sockeye salmon, CONJA-S was one to three times as fast as EM and over 40 times as fast as GIRLS-B. The speed advantage of CONJA-S generally increased with increase of specified GPA, and was

slightly greater for total search time than effective search time.

Neither CONJA-S nor EM failed on any data set at any specified GPA up to the maximum attempted, 90%. On the other hand, GIRLS-B failed at low rates (<3% of resamplings) for two of the data sets (Columbia River chinook salmon and West Coast sockeye salmon) but did not fail on the third (Yukon River chum salmon) (Table 10). Average GPA achieved per resampling by each of the programs always equaled or exceeded that specified, even when failures (GIRLS-B) occurred (Table 10). EM average GPA agreed closely with specified GPA, while CONJA-S and GIRLS-B average GPAs exceeded that specified. These differences between average GPA achieved and that specified are due to the differences in the stepwise rate of convergence among the programs. GIRLS-B and CONJA-S satisfied the specified GPA in relatively few steps with large GPA increases between steps and usually exceeded by greater amounts the specified GPA at the end of search than did EM, which required more steps having smaller GPA increases between steps.

Statistics of the bootstrap distributions for estimated stock percentages comprising the three salmon mixtures-Columbia River chinook salmon, Yukon River chum salmon, and West Coast sockeye salmon-showed that the finer the stock mixture was separated into its components, the greater were the changes in the distributions as specified GPA varied. At coarsest resolution when the mixture composition was assessed for only a few major stock groups, neither the program used nor GPA specified had meaningful effect on the bootstrap distributions. Details are presented next.

Contributions from four Columbia River chinook salmon stock groups were assessed (Table 11) from estimated mixture percentages of their component-stocks. Average estimates of stock group percentages, their standard deviations, and their upper and lower 95% confidence bounds differed only slightly (50.5%) between all GPAs and programs. Maximum discrepancy in average estimates for stock groups among GPAs and programs was only 0.1%, for lower confidence bounds was only 0.5%, for upper confidence bounds was only 0.4%, and for standard deviations was only 0.19%. Estimates for the fourteen component stocks computed for the Columbia River chinook salmon mixture by CONJA-S for standard GPAs (Table 12) show greater discrepancies than when stocks were consolidated. Maximum discrepancy in average estimates for stocks among GPAs was 0.1% (stocks RR, CA, and LW); for lower confidence bounds, 0.7% (stock KO); for upper confidence intervals, 1.8% (stock LS); and for standard deviations, 0.32% (stock CA).

Bootstrap estimates for summer run and fall run groups of the Yukon River chum salmon mixture were also computed (Table 13) from individual stock contribution estimates. Average estimates of stock group percentages, standard deviations, and

upper and lower 95% confidence bounds differed at most by 2.3% among all GPAs and programs (Table 13). Maximum discrepancy in average estimates for stock groups among GPAs and programs was only 0.3%, for lower or upper confidence bounds was 2.3%, and for standard deviations was 1.10%.

Statistics for bootstrap distributions of composition estimates for two major consolidations of stock groups (A-G and H-M) (Table 14) as well as the thirteen individual groups (A through M) (Table 15) were computed from individual stock contribution estimates for the West Coast sockeye salmon mixture. Maximum discrepancy in average estimates for consolidations of stock groups among GPAs and programs was only 0.3%, for lower or upper confidence bounds was 0.4%, and for standard deviations was 0.04% (Table 14). Estimates for the thirteen component stock groups (A through M) computed by CONJA-S for standard GPAs (Table 15) show far greater discrepancies than occurred when component groups were consolidated. Maximum discrepancy in average estimates for stock groups among GPAs was 2.5% (group J); for lower confidence bounds, 6.4% (group L); for upper confidence intervals, 8.5% (group J); and-for standard deviations, 3.96% (group J).

CONCLUSION

Several general-purpose optimization algorithms have been specialized to compute the CMLE of stock composition from characters of individuals in mixtures. This study reviews these algorithms, describes others, and compares performances of implementing computer programs. The original application to genetic information by staff of the former Northwest and Alaska Fisheries Science Center of the National Marine Fisheries Service used the expectation maximization algorithm (Milner et al. 1981). Next, scientists of the Pacific Biological Station of Canada Department of Fisheries and Oceans used the conjugate gradient algorithm with logarithm-transformed stock proportions (Fournier et al. 1984). Most recently, researchers of the Auke Bay Laboratory of the Alaska Fisheries Science Center have used the iteratively reweighted least-squares algorithm suggested by Pella (1986) and Pella and Milner (1987) as well as a variant described here. The present study introduces two other specializations of general-purpose algorithms: steepest ascent and conjugate gradient with square root transformation of stock proportions.

The numerical precision (i.e., number of decimal places) to which the stock proportions composing -the CMLE can be determined by any of the algorithms is limited by computing machinery and programming languages used; Increasing precision costs rapidly increasing computer time. A compromise between precision and number of CMLE evaluations

required in a study is necessary to control total computer time needed. In this study, search for a particular CMLE was stopped when either a time limit was reached or a criterion of proximity to the maximum of the conditional likelihood function was satisfied. The proximity criterion was the guaranteed percent achieved (GPA), which is the minimum percentage that the current likelihood function value during search represents of the maximum possible likelihood function value. Generally, achieving a GPA of 100% was not possible because of rounding errors in computations. Other criteria for stopping search could have been 1) relative change in the likelihood function between iterations or 2) sum of absolute values of components of the gradient vector (in feasible directions). None of the stopping criteria was ideal because none delimited a region within the stock composition space wherein the CMLE occurred.

Program performances were studied under hypothetical (Experiment I) and real-world (Experiment II) stock mixtures. Three performance measures were used: speed, failure rate, and stability of CMLE distributions as specified GPA was increased. These measures are important when computing numerous CMLEs to describe distributions of stock. composition estimates by simulation- and bootstrapping studies. The ideal algorithm would be fastest, have the lowest failure rates, and have most stable CMLE distributions. Although no program excelled at every performance measure, overall performance of CONJA-S recommends its use in time-consuming studies. Findings from both experiments showed that CONJA-S was either fastest by far or only slightly slower than the fastest program for all situations examined. The good speed of CONJA-S held for total search time and effective search time. Under hypothetical situations of Experiment I, the EM was usually slightly faster than CONJA-S for cases with lower numbers of baseline stocks, smaller mixture samples, lower specified GPA, and the diverse baseline. The good speed of EM was unexpected considering previous mention of the slow convergence of the expectation maximization algorithm in mixture problems (e.g., Redner and Walker 1984 and Roeder et al. 1989). CONJA-S supplanted EM for cases with larger numbers of baseline stocks, larger mixture samples, higher specified GPA, and the similar baseline. The remaining programs (GIRLS-D, GIRLS-B, STEEP, and CONJA-L) were fastest in only a few cells of Experiment I. In all three real-world applications of Experiment- II, CONJA-S was fastest of the three programs examined (CONJA-S, EM, and GIRLS-B). The CONJA-S superiority in speed was greater as the specified GPA increased.

Failures to achieve specified GPA differed substantially among programs. CONJA-S and EM had no failures to achieve specified GPA up to 90% among the total 1,800 trials of Experiment I and 3,000 trials of Experiment II. GIRLS-B had next lowest failure rates with 2 of 1,800 (0.1%) trials of

Experiment I and 40 of 3,000 (1.3%) trials for Experiment II. GIRLS-D, STEEP and CONJA-L had increasing failures rates for Experiment I and were not examined further in Experiment II. Preliminary experimentation (not reported) showed that all the programs failed for some trials during both Experiment I and Experiment II if the value of the likelihood function was not allowed to decrease (search to regress) between iterations of the search. In theory, the algorithms advance with uninterrupted increase to the maximum of the likelihood function; but in practice, rounding errors in computations cause temporary decrease (real or apparent) in likelihood values during search. Therefore, search was not stopped for such decrease in likelihood value.

Stability of CMLE distributions was measured by comparing statistics of resampling distributions obtained at a range of specified GPAs. Ideally, the stock composition distributions would be independent of the specified GPA so that a low GPA could be specified with concomitant low search time. When the required level of stock detail was to the individual stock as in Experiment I, the superior performances of the GIRLS programs were evident; differences in means and standard deviations of CMLE distributions between specified GPA of 10%, 50%, and 90% as compared with 99% were smallest among all programs. CONJA-S was best among the remaining programs, and EM was the worst. The relative performances reflected differences in numbers of steps and corresponding GPA increments between steps. The GIRLS programs required few steps to achieve specified GPA, which resulted in large GPA increments between steps; the EM program required many steps, which resulted in small GPA increments. The final GPA at termination when specified GPA was achieved tended to be greater for a program using fewer steps; and as a result, stability for such a program was higher. Differences in stability were caused by differences in GPA achieved rather than GPA specified.

Stability of stock composition distributions to the level of individual stocks may often be unnecessary. Such stock composition estimates may have too large sampling variation to be of practical use. If so, stocks should be combined into groups with similar characteristics, and the CMLE estimate for individual stocks should be summed over stocks of any group to provide the CMLE for groups. In the real-world examples for which stock detail demanded was only to the level of similarity groups, the differences in stability performance among programs was less evident and apparently of little practical consequence. Even differences in group CMLE distributions among GPAs of 10%, 50%, and 90% appeared inconsequential.

Certain caveats and considerations may be helpful in choosing among the algorithms. Two warnings are first issued.

First, this study was limited to line search (for algorithms other than the expectation maximization) based on approximating the support function by a Taylor's series along a line in the original stock composition space or in a transform of that space. The supposition for this choice was that such an approximation, if sufficiently accurate, would provide a fast line search. However, other approaches to line search are possible (e.g., Press et al. 1989), and these may well be more certain of success than the method chosen. Second, the algorithms are implemented by our FORTRAN coding, which may contain errors. The GIRLS programs have been used by numerous organizations for several years. Coding errors have been reported, corrected by us, and revisions returned to all known users. Although reports of GIRLS program errors have not occurred recently, detection of errors is difficult and some may remain. The other programs are not so well tested. Naturally, any errors remaining could affect performance.

Several final reflections to assist in selection of an algorithm conclude this discussion. First, complexity of the algorithm weighs against its use unless a trustworthy coding with required features is already available. If original coding must be developed, expectation maximization is simplest; iteratively reweighted least squares, most complex; and the conjugate gradient with square root transform, intermediate. The superior speed of the conjugate gradient with square root transform together with its reliable maximizing of the likelihood function justifies its greater complexity compared to expectation maximization. Input and output procedures and features for conducting simulations and bootstrapping constitute a significant amount of coding for any of the programs. Therefore, complexity of the algorithm becomes relatively less important if a versatile program including such procedures and features is required. Second, the iteratively reweighted least squares algorithm can easily be modified to include additional linear constraints on the stock proportions by adding to the constraint equation set. Such modification can be used to compute likelihood profiles for stock group contributions to a mixture', for example. Third, the iteratively reweighted least squares algorithm becomes more efficient as mixture sample size increases (contrary to the other algorithms). Conceivably, advances in sampling technology will permit large mixture samples favoring use of this algorithm. Fourth, and last, another criterion that could have been used for comparing algorithms is their computer memory requirement. Determining the minimal amount of memory required by each algorithm would have been

[2]Richard Gates, CFMD, Genetics, 333 Raspberry Road, Anchorage, AK 99518-1599, pers. commun., 15 November 1994.

technically difficult, although clearly the GIRLS programs require more than the other programs. More importantly, the criteria chosen are more constraining to analysis of stock composition than memory capacity. Readily-available, inexpensive computer memory chips for personal computers (386 machines or higher) with limited memory allow use of the executable modules developed for any of the algorithms during this study. However, performances of the algorithms including speed and convergence properties are not easily modified, and those properties can be limiting in practical applications even using fast mainframe computers.

ACKNOWLEDGMENTS

Data files used in Experiment II were kindly provided by Drs. Fred Utter (chinook salmon), Richard Wilmot (chum salmon), and Chris Wood (sockeye salmon).

CITATIONS

Bartley, D., B. Bentley, J. Brodziak, R. Gomulkiewicz, M. Mangel, and G.A.E. Gall. 1992. Geographic variation in population genetic structure of chinook salmon from California and Oregon. Fish. Bull., U.S. 90:77-100 (authorship amended per errata, Fish. Bull., U.S. 90(3) :iii).

Bowen, B. W., F. A. Abreu-Grobois, G. H. Balazs, N. Kamezaki, C. J. Limpus, and R. J. Ferl. 1995. Trans-Pacific migrations of the loggerhead turtle *(Caretta* caretta) demonstrated with mitochondrial DNA markers. Proc. Nat. Acad. Sci. 92:3731-3734.

Cook, R., and G. Lord. 1978. Identification of stocks of Bristol Bay sockeye salmon by evaluating scale patterns with a polynomial discriminant method. Fish. Bull,, U.S. 76:415-423.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion). J. Royal Stat. Soc., series B, 39:1-38.

Edwards, A. W. F. 1972. Likelihood.. Cambridge Univ. Press. 235 p.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. Soc. Ind. Appl. Math., Philadelphia, 92 p.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science l(l) :54-77.

Fletcher, R., and C. M. Reeves. 1964. Function minimization by conjugate gradients. Computer Journal 7(2):149-154.

Fournier, D. A., T. D. Beacham, B. E. Riddell, and C. A. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Can. J. Fish. Aquat. Sci. 41:400-408.

Fukuhara, F. M., S. Murai, J. J. LaLanne, and A. Sribhibhadh. 1962. Continental origin of red salmon as determined from morphological characters. Int. North Pac. Fish. Comm., Bull. 8:15-109.

Grant, W. S., G. B. Milner, P. Krasnowski, and F. M. Utter. 1980. Use of biochemical genetic variants for identification of sockeye salmon *(Oncorhynchus nerka)* stocks in Cook Inlet, Alaska. Can. J. Fish. Aquat. Sci. 37:1236-1247.

Hillier, F. S., and G. J. Lieberman. 1967. Introduction to operations research. Holden-Day, Inc., San Francisco. 639 p.

Jennrich, R. I., and M. L. Ralston. 1978. Fitting nonlinear models to data. Tech. Rep. 46. Health Sciences Computing Facility, Univ. of California, Los Angeles, 71 p.

Kuhn, H. W., and A. W. Tucker. 1951. Nonlinear programming. *In* J. Neyman (editor), Proceedings of the second Berkeley symposium on mathematical statistics and probability, p. 481-492. Univ. of California Press, Berkeley.

Lehmann, E. L. 1983. Theory of point estimation. Wiley & Sons, New York, 506 p.

Lindsay, B. G. 1983. The geometry of mixture likelihoods: a general theory. Annals of Statistics 11:86-94.

Marshall, S. L., G. T. Oliver, D. R. Bernard, and S. A. McPherson. 1984. Accuracy of scale pattern analysis in separating major stocks of sockeye salmon *(Oncorhynchus nerka)* from southern southeastern Alaska and northern British Columbia. Alaska Dep. Fish and Game, Informational Leaflet 230, 29 p.

Masuda, M., S. Nelson, and J. Pella. 1991. User's manual for GIRLSEM, GIRLSYM, and CONSQRT. Personal computer version. USA-DOC-NOAA-NMFS, Auke Bay Lab., US-Canada Salmon Program, 11305 Glacier Hwy., Juneau, AK 99801-8626. 70 p.

Millar, R. B. 1987. Maximum likelihood estimation of mixed stock fishery composition. Can. J. Fish. Aquat. Sci. 44:583-590.

Milner, G. B., D. J. Teel, F. M. Utter, and C. L. Burley. 1981. Columbia River stock identification study: validation of genetic method. Final report of research (FY80). Coastal Zone and Estuarine Studies Div., NWAFC, NOAA, 2725 Montlake Blvd. East, Seattle, WA 98112-2097. 35 p.

Moles, A., P. Rounds, and C. Kondzela. 1990. Use of the brain parasite Myxobolus *neurobius* in separating mixed stocks of sockeye salmon. Am. Fish. Soc. Symp. 7:224-231.

Mulligan, T. J., S. McKinnell, and C. C. Wood. 1988. Uncertainty in stock composition estimates of oceanic steelhead trout using electrophoretic characteristics: comments on a recent study. Can. J. Fish. Aquat. Sci. 45:432-442.

Pella, J. J. 1986. The method of fitting expectations applied to computation of conditional maximum likelihood

estimates of stock composition from genetic marks. Manuscr. Rep. File MR-F 171, 17 p. Auke Bay Lab., 11305 Glacier Hwy., Juneau, AK 99801-8626.

Pella, J. J., and G. B. Milner. 1987. Use of genetic marks in stock composition analysis. In N. Ryman and F. Utter (editors), Population genetics & fishery management. Univ. of Washington Press, Seattle, 420 p.

Pella, J. J., M. Masuda, C. Guthrie, C. Kondzela, A. Gharrett, A. Moles, and G. Winans. In press. Stock composition of some sockeye salmon *(Oncorhynchus* nerka) catches in Southeast Alaska, based on allozyme, scale, and brain parasite characters of sampled individuals. NOAA Tech. Rep.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1989. Numerical recipes. The art of scientific computing (FORTRAN version). Cambridge Univ. Press, New York, 702 p.

Rao, S. S. 1984. Optimization: Theory and applications. 2nd ed. Wiley & Sons, New York, 747 p.

Redner, R. A., and H. F. Walker. 1984. Mixture densities, maximum likelihood and the EM algorithm. SIAM Rev. 26:195-239.

Roeder, K., B. Devlin, and B. G. Lindsay. 1989. Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities. Biometrics 45:363-379.

Shaklee, J. B., and S. R. Phelps. 1990. Operation of a large-scale, multiagency program for genetic stock identification. Am. Fish. Soc. Symp. 7:817-830.

Shaklee, J. B., C. Busack, A. Marshall, M. Miller, and S. R. Phelps. 1990. The electrophoretic analysis of mixed-stock fisheries of Pacific salmon. *In* Z-I. Ogita and C. L. Markert (editors), Isozymes: structure, function, and use in biology and medicine, p. 235-265. Wiley-Liss, Inc., New York.

Smouse, P. E., R. S. Waples, and J. A. Tworek. 1990. A genetic mixture analysis for use with incomplete source population data. Can. J. Fish. Aquat. Sci. 47:620-634.

Utter, F., D. Teel, G. Milner, and D. McIsaac. 1987. Genetic estimates of stock compositions of 1983 chinook salmon, *Oncorhynchus tshawytscha,* harvests off the Washington coast and the Columbia River. Fish. Bull., U.S. 85:13-23.

Wilmot, R. L., R. J. Everett, W. J. Spearman, R. Baccus, N. V. Varnavskaya, and S. V. Putivkin. 1994. Genetic stock structure of western Alaska chum salmon and a comparison

with Russian Far East stocks. Can. J. Fish. Aquat. Sci. 51 (Suppl. 1):84-94.

Wirgin, I., L. Maceda, J. R. Waldman, and R. N. Crittenden. 1993. Use of mitochondrial DNA polymorphisms to estimate the relative contributions of the Hudson River and Chesapeake Bay striped bass stocks to the mixed fishery on the Atlantic coast. Trans. Am. Fish. Soc. 122:669-684.

Wolfe, P. 1959. The simplex method for quadratic programming. Econometrica 27:382-398.

Wood, C. C., D. T. Rutherford, and S. McKinnell. 1989. Identification of sockeye salmon *(Oncorhynchus nerka)* stocks in mixed-stock fisheries in British Columbia and southeast Alaska using biological markers. Can. J. Fish. Aquat. Sci. 46:2108-2120.

Xu, s., C. J. Kobak, and P. E. Smouse. 1994. Constrained least squares estimation of mixed population stock composition from mtDNA haplotype frequency data. Can. J. Fish. Aquat. Sci. 51:417-425.

TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

Table 1.-Stock-specific and loci-shared relative allele frequencies (RAFs) of alleles A, B, C, D, and E at their respective loci for the five-stock diverse and similar baselines of Experiment I.[a]

	RAF of alleles A, B, C, D, and E				
	Stock				
Baseline	1	2	3	4	5
Diverse	0.2	0.4	0.6	0.8	1.0
Similar	0.4	0.5	0.6	0.8	1.0

[a]The range and intervals of the proportions for diverse and similar baselines are as follows: diverse (range = 0.8; intervals = 0.2) and similar (range = 0.6; intervals 1 and 2 = 0.1 and intervals 3 and 4 = 0.2).

Table 2.-Features of real-world data sets used in Experiment II: number of baseline stocks, number of characters observed, range of baseline sample sizes, mixture sample size, and literature source of information.

Data set	Baseline stocks	Characters	Baseline sizes	Mixture size	Reference*
Columbia chinook	14	9	151 - 400	1,597	a
Yukon chum	26	19	24 - 297	197	b
West Coast sockeye	73	4	10 - 3,010	299	c

*a = Pella and Milner (1987); b = Wilmot et al. (1994); c = Pella et al. (in press).

Table 3.-Number of the four mixture sample size cells (50, 150, 250, or 500) of Experiment I for which search by programs (GIRLS-D, GIRLS-B, STEEP, EM, CONJA-L, and CONJA-S) for stock composition estimates failed to achieve specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%) for one or more of the 25 bootstrap resamplings, by number of baseline stocks (5, 15, and 50) and baseline (Diverse or Similar) (upper table); and the corresponding total number of failures per 100 total resamplings (lower table).

Number of cells with failures

	5 stocks						15 stocks						50 stocks					
	Diverse			Similar			Diverse			Similar			Diverse			Similar		
	GPA (%)			GPA (%)			GPA (%)			GPA (%)			GPA (%)			GPA (%)		
Program	10	50	90	10	50	90	10	50	90	10	50	90	10	50	90	10	50	90
GIRLS-D	0	0	0	0	0	0	0	0	1	0	0	1	0	1	2	0	0	2
GIRLS-B	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
STEEP	0	0	1	0	0	1	0	0	1	0	0	2	0	0	2	0	0	2
EM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CONJA-L	0	0	0	0	0	0	1	2	2	4	4	4	3	4	4	4	4	4
CONJA-S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Number of resampling failures among cells

	5 stocks						15 stocks						50 stocks					
	Diverse			Similar			Diverse			Similar			Diverse			Similar		
	GPA (%)			GPA (%)			GPA (%)			GPA (%)			GPA (%)			GPA (%)		
Program	10	50	90	10	50	90	10	50	90	10	50	90	10	50	90	10	50	90
GIRLS-D	0	0	0	0	0	0	0	0	1	0	0	1	0	1	4	0	0	5
GIRLS-B	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
STEEP	0	0	7	0	0	3	0	0	9	0	0	14	0	0	23	0	0	22
EM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CONJA-L	0	0	0	0	0	0	1	2	3	[a]	[a]	[a]	[a]	[a]	[a]	[b]	[b]	[b]
CONJA-S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

[a]In one or more resamplings, the search was stopped due to an invalid number, a division by zero, or an arithmetic overflow.

[b]In one or more resamplings, the search was stopped due to an invalid argument for the log function.

Table 4 .-Fastest programs* in effective search time for finding the 25 bootstrap stock composition estimates with specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%) given number of baseline stocks (5, 15, and 50), mixture sample size (50, 150, 250, and 500) and baseline (Diverse or Similar).

		Fastest program in effective search time					
		GPA 10%		GPA 50%		GPA 90%	
Stocks	Mixture	Diverse	Similar	Diverse	Similar	Diverse	Similar
5	50	EM	EM	EM	EM	EM,S	S
	150	EM	EM	EM	EM	EM,S	S
	250	EM	EM,S	EM	S	EM	S
	500	EM	EM	D	B	S,D	D
15	50	EM	EM	EM	S	S	S
	150	EM	S	S	S	S	S
	250	S	S	S	S	S	S
	500	S	S	S	S	B	B
50	50	EM	EM,S	EM	S	S	S
	150	S	S	S	S	S	S
	250	S	S	S	S	S	S
	500	S	S	S	S	S	S

*Nonstandard notation for programs is as follows: S = CONJA-S, D = GIRLS-D, and B = GIRLS-B.

Table 5 .-Fastest programs* in total search time for finding the 25 bootstrap stock composition estimates with specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%) given number of baseline stocks (5, 15, and 50), mixture sample size (50, 150, 250, and 500), and baseline (Diverse or Similar).

		Fastest program in total time					
		GPA 10%		GPA 50%		GPA 90%	
Stocks	Mixture	Diverse	Similar	Diverse	Similar	Diverse	Similar
5	50	EM,S,L, STEEP	EM,S, STEEP	EM	EM,S, STEEP	EM,S, STEEP	S
	150	EM	EM,S	STEEP	S	S	S
	250	EM	S	EM,S	S	S	S
	500	EM,S, B,D	B,D	D,S	D,B	S	D
15	50	EM	EM,S	S	S	S	S
	150	S	S	S	S	S	S
	250	S	S	S	S	S	S
	500	S	S	S	S	B	B
50	50	S	S	S	S	S	S
	150	S	S	S	S	S	S
	250	S	S	S	S	S	S
	500	S	S	S	S	S	S

*Nonstandard notation for programs is as follows: S = CONJA-S, D = GIRLS-D, B = GIRLS-B, and L = CONJA-L.

Table 6 .-Maximum discrepancies" among stocks of bootstrap mean estimated mixture percentages [max (Δave_{α})] in relation to number of stocks in the diverse baselines (5, 15, or 50), specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%), and mixture sample size (50 and 500).

	Maximum discrepancy [max(Δave_{α})]								
	Stocks in baseline								
	5			15			50		
	GPA to stop			GPA to stop			GPA to stop		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
Mixture sample size = 50									
GIRLS-D	0.7	0.2	0.0	0.6	0.2	0.0	0.8	0.2	0.1
GIRLS-B	0.8	0.2	0.0	0.6	0.2	0.0	0.8	0.4	0.0
STEEP	0.9	0.2	0.1	4.1	2.2	0.5	2.6	1.4	0.6
EM	2.9	1.2	0.3	6.1	3.4	0.7	3.3	1.4	0.4
CONJA-L	1.8[b]	1.1[b]	0.2[b]	3.9[b]	[c]	[c]	2.3[b]	[c]	[c]
CONJA-S	1.0	0.4	0.1	3.6	1.3	0.2	2.2[b]	1.2[b]	0.4[b]
Mixture sample size = 500									
GIRLS-D	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]
GIRLS-B	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.1[d]	0.0[d]
STEEP	0.2[d]	0.1[d]	0.0[d]	0.9[d]	0.4[d]	[c]	1.1[d]	0.4[d]	[c]
EM	1.5[d]	0.7[d]	0.2[d]	0.6[d]	0.3[d]	0.1[d]	1.5[d]	0.6[d]	0.2[d]
CONJA-L	0.9[d]	0.5[d]	0.1[d]	0.9[d]	0.9[d]	0.0[d]	[c]	[c]	[c]
CONJA-S	0.2[d]	0.1[d]	0.0[d]	0.7[d]	0.2[d]	0.1[d]	0.5[d]	0.2[d]	0.1[d]

[a]Discrepancies were computed as absolute differences between bootstrap means resulting from specified standard GPAs and that when GPA of 99% was specified. Bootstrap means were based on 25 resamplings of original baseline stock and mixture samples of Experiment I.

[b]Because some of the programs did not attain the GPA of 99%, the means from GIRLS-B with GPA of 99% were used to compute differences.

[c]GPA was not achieved.

[d]Because some of the programs did not attain the GPA of 99%, the means for GIRLS-B with GPAs of 94, 98, and 98% for 5, 15, and 50 stocks, respectively, were used to compute differences.

Table 7 .-Maximum discrepancies" among stocks of bootstrap standard errors of estimated mixture percentages [max (AS,)] in relation to number of stocks in the diverse baselines (5, 15, or 50), specified standard values of the stopping criterion (GPA of 10%, 50%, or 90%), and mixture sample size (50 and 500).

	Maximum discrepancy [max(Δs_α)]								
	Stocks in baseline								
	5			15			50		
	GPA to stop			GPA to stop			GPA to stop		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
Mixture sample size = 50									
GIRLS-D	0.5	0.2	0.0	0.6	0.2	0.0	1.0	0.5	0.1
GIRLS-B	0.3	0.2	0.0	0.5	0.2	0.0	1.3	0.5	0.1
STEEP	2.2	0.5	0.1	7.7	4.7	0.7	6.0	3.5	2.9
EM	3.6	1.6	0.3	9.0	5.0	1.5	5.9	4.1	1.5
CONJA-L	2.1[b]	1.4[b]	0.3[b]	3.4[b]	[c]	[c]	4.7[b]	[c]	[c]
CONJA-S	1.7	0.5	0.1	4.5	2.4	0.5	4.5[b]	4.3[b]	2.0[b]
Mixture sample size = 500									
GIRLS-D	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.2[d]	0.1[d]	0.0[d]
GIRLS-B	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.2[d]	0.1[d]	0.0[d]
STEEP	0.9[d]	0.3[d]	0.1[d]	1.4[d]	0.6[d]	[c]	3.1[d]	1.3[d]	[c]
EM	1.7[d]	0.7[d]	0.2[d]	2.3[d]	1.2[d]	0.4[d]	2.4[d]	1.1[d]	0.5[d]
CONJA-L	0.6[d]	0.3[d]	0.1[d]	1.2[d]	1.0[d]	0.1[d]	[c]	[c]	[c]
CONJA-S	0.4[d]	0.2[d]	0.0[d]	0.8[d]	0.3[d]	0.1[d]	0.9[d]	0.6[d]	0.1[d]

[a]Discrepancies were computed as absolute differences between bootstrap standard errors resulting from specified standard GPA and that when GPA of 99% was specified. Bootstrap standard errors were based on 25 resamplings of original baseline stock and mixture samples of Experiment I.

[b]Because some of the programs did not attain the GPA of 99%, the standard errors from GIRLS-B with GPA of 99% were used to compute differences.

[c]GPA was not achieved.

[d]Because some of the programs did not attain the GPA of 99%, the standard errors from GIRLS-B with GPAs of 94, 98, and 98% for 5, 15, and 50 stocks, respectively, were used to compute differences.

Table 8.--Maximum discrepancies" among stocks of bootstrap mean estimated mixture percentages [max (Δave_{α}) l in relation to number of stocks in the similar baselines (5, 15, or 50), specified standard values of the stopping criterion (GPA of 10%, 50%, or 90%), and mixture sample size (50 and 500).

	Maximum discrepancy [max(Δave_{α})]								
	Stocks in baseline								
	5			15			50		
	GPA to stop			GPA to stop			GPA to stop		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
Mixture sample size = 50									
GIRLS-D	0.8	0.4	0.0	1.2	0.3	0.1	1.9	0.3	0.1
GIRLS-B	1.4	0.3	0.0	1.1	0.3	0.0	1.8	0.3	0.1
STEEP	0.8	0.3	0.3	5.3	1.4	0.6	7.1[b]	2.0[b]	0.5[b]
EM	2.2	1.6	0.4	5.7	2.2	0.7	5.1	1.9	0.3
CONJA-L	8.3[b]	6.3[b]	0.3[b]	2.5[b]	1.6[b]	[c]	1.8[b]	[c]	[c]
CONJA-S	2.5	0.4	0.1	2.5[b]	1.0[b]	0.2[b]	3.1[b]	1.2[b]	0.4[b]
Mixture sample size = 500									
GIRLS-D	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.1[b]	0.1[b]	0.0[b]
GIRLS-B	0.1[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.1	0.0	0.0
STEEP	0.7[d]	0.2[d]	0.0[d]	1.8[d]	0.5[d]	[c]	1.9[b]	0.6[b]	[c]
EM	0.5[d]	0.2[d]	0.1[d]	2.3[d]	0.8[d]	0.2[d]	1.7[b]	0.6[b]	0.2[b]
CONJA-L	0.6[d]	0.3[d]	0.1[d]	[c]	[c]	[c]	1.0[b]	[c]	[c]
CONJA-S	0.1[d]	0.1[d]	0.0[d]	0.6[d]	0.2[d]	0.0[d]	0.6[b]	0.3[b]	0.1[b]

[a]Discrepancies were computed as absolute differences between bootstrap means resulting from specified standard GPA and that when GPA of 99% was specified. Bootstrap means were based on 25 resamplings of original baseline stock and mixture samples of Experiment I.

[b]Because some of the programs did not attain the GPA of 99%, the means from GIRLS-B with GPA of 99% were used to compute differences.

[c]GPA was not achieved

[d]Because some of the programs did not attain the GPA of 99%, the means from GIRLS-B with GPAs of 96 and 98% for 5 and 15 stocks, respectively, were used to compute differences.

Table 9.--Maximum discrepancies" among stocks of bootstrap standard errors of estimated mixture percentages [max (As,)] in relation to number of stocks in the similar baselines (5, 15, or 50), specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%), and mixture sample size (50 and 500).

	Maximum discrepancy [max(Δs_α)]								
	Stocks in baseline								
	5			15			50		
	GPA to stop			GPA to stop			GPA to stop		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
Mixture sample size = 50									
GIRLS-D	0.2	0.4	0.0	1.1	0.5	0.2	1.8	1.2	0.1
GIRLS-B	0.7	0.4	0.0	1.6	0.2	0.1	1.9	1.1	0.1
STEEP	5.0	1.5	0.6	3.9	2.4	0.5	6.9[b]	3.5[b]	1.4[b]
EM	7.8	3.8	0.9	5.4	3.0	0.8	7.2	3.7	0.9
CONJA-L	6.8[b]	6.9[b]	0.4[b]	3.9[b]	2.4[b]	[c]	4.0[b]	[c]	[c]
CONJA-S	4.6	1.6	0.2	3.6[b]	1.3[b]	0.4[b]	4.5[b]	2.2[b]	1.2[b]
Mixture sample size = 500									
GIRLS-D	0.0[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.2[b]	0.1[b]	0.0[b]
GIRLS-B	0.0[d]	0.0[d]	0.0[d]	0.1[d]	0.0[d]	0.0[d]	0.2	0.1	0.0
STEEP	1.1[d]	0.3[d]	0.1[d]	2.3[d]	0.6[d]	[c]	3.2[b]	2.1[b]	[c]
EM	2.3[d]	1.0[d]	0.2[d]	3.8[d]	1.9[d]	0.4[d]	3.5[b]	1.7[b]	0.6[b]
CONJA-L	0.7[d]	0.3[d]	0.0[d]	[c]	[c]	[c]	2.1[b]	[c]	[c]
CONJA-S	0.6[d]	0.3[d]	0.0[d]	0.7[d]	0.4[d]	0.1[d]	1.4[b]	0.8[b]	0.2[b]

[a]Discrepancies were computed as absolute differences between bootstrap standard errors resulting from specified standard GPA and that when GPA of 99% was specified. Bootstrap standard errors were based on 25 resamplings of original baseline stock and mixture samples of Experiment I.

[b]Because some of the programs did not attain the GPA of 99%, the standard errors from GIRLS-B with GPA of 99% were used to compute differences.

[c]GPA was not achieved.

[d]Because some of the programs did not attain the GPA of 99%, the standard errors from GIRLS-B with GPAs of 96 and 98% for 5 and 15 stocks, respectively, were used to compute differences.

Table 10.-Performances of programs (CONJA-S, EM, and GIRLS-B) on real-world data sets used in Experiment II: average effective search time, average total search time, average guaranteed percent achieved, and failure rate for 1000 resamplings at specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%). All times are in centiseconds.

	Effective time			Total time			% achieved			Failure rate (%)		
	GPA to stop											
	GPA			GPA			GPA			GPA		
Data set	10%	50%	90%	10%	50%	90%	10%	50%	90%	10%	50%	90%
Columbia chinook												
CONJA-S	48	64	100	49	65	102	22	62	93	0.0	0.0	0.0
EM	114	222	508	171	328	748	10	50	90	0.0	0.0	0.0
GIRLS-B	97	120	*	102	126	*	36	71	95	0.0	0.0	2.2
Yukon chum												
CONJA-S	26	43	88	27	43	89	21	59	92	0.0	0.0	0.0
EM	46	101	273	66	143	380	11	50	90	0.0	0.0	0.0
GIRLS-B	1308	2683	7391	1340	2747	7572	13	54	91	0.0	0.0	0.0
West Coast sockeye												
CONJA-S	95	226	527	97	230	535	14	53	91	0.0	0.0	0.0
EM	115	335	1248	157	462	1744	10	50	90	0.0	0.0	0.0
GIRLS-B	4155	*	*	4159	*	*	41	70	94	0.0	0.4	1.8

'Search for the CMLE was stopped on one or more resamplings because the effective search time exceeded 5 min.

Table 11.--Averages (ave), standard deviations (s), and upper ($\hat{p}_{0.975}$) and lower ($\hat{p}_{0.025}$) 95% confidence bounds from the bootstrap distributions (1000 resamplings) of estimated stock group (groups* are denoted as B, A+B, A+B+C, and D) percentages for the Columbia River chinook salmon mixture by programs CONJA-S, EM, and GIRLS-B for specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%).

	GPA to stop											
	10%				50%				90%			
Program	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s
CONJA-S												
B	42.2	60.9	51.9	4.75	42.3	60.9	51.9	4.76	42.3	60.9	51.9	4.77
A+B	44.8	61.9	53.5	4.26	44.7	61.9	53.5	4.28	44.7	61.9	53.5	4.29
A+B+C	76.2	88.1	82.1	3.03	76.2	88.1	82.1	3.05	76.2	88.1	82.1	3.07
D	11.9	23.8	17.9	3.03	11.9	23.8	17.9	3.05	11.9	23.9	17.9	3.07
EM												
B	42.6	60.6	51.9	4.59	42.3	60.9	51.9	4.71	42.3	60.9	51.9	4.76
A+B	45.0	61.9	53.5	4.15	44.8	61.8	53.5	4.24	44.7	61.9	53.5	4.28
A+B+C	76.5	87.8	82.0	2.88	76.3	88.1	82.1	3.00	76.2	88.1	82.1	3.06
D	12.2	23.5	18.0	2.88	11.9	23.7	17.9	3.00	11.9	23.8	17.9	3.06
GIRLS-B												
B	42.2	61.0	51.9	4.77	42.2	60.9	51.9	4.78	42.2	60.9	51.9	4.78
A+B	44.7	61.8	53.5	4.30	44.8	61.9	53.5	4.30	44.7	61.9	53.5	4.30
A+B+C	76.0	88.2	82.1	3.07	76.2	88.1	82.1	3.07	76.1	88.1	82.1	3.07
D	11.8	23.9	17.9	3.07	11.9	23.9	17.9	3.07	11.9	23.9	17.9	3.07

'Groups are defined in Table 10.1 of Pella and Milner (1987).

Table l2.--Averages (ave), standard deviations (s), and upper ($\hat{p}_{0.975}$) and lower ($\hat{p}_{0.025}$) confidence bounds from the bootstrap distributions (1000 resamplings) of estimated individual stock* percentages for the Columbia River chinook salmon mixture by the program CONJA-S for specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%).

	GPA to stop											
	10%				50%				90%			
Stock	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s
WS	0.0	5.3	1.6	1.62	0.0	5.4	1.6	1.63	0.0	5.4	1.6	1.64
RR	22.6	34.8	28.5	3.14	22.6	34.8	28.6	3.15	22.6	34.8	28.6	3.16
KO	0.8	40.8	21.9	10.30	0.3	41.0	21.9	10.49	0.1	41.1	21.9	10.61
RB	0.0	5.3	0.9	1.60	0.0	5.4	0.9	1.66	0.0	5.5	0.9	1.68
CA	0.6	37.2	16.8	9.13	0.3	37.3	16.9	9.33	0.0	37.5	16.9	9.45
EC	0.0	0.0	0.0	0.02	0.0	0.0	0.0	0.02	0.0	0.0	0.0	0.02
LS	0.0	3.0	0.2	1.11	0.0	1.8	0.2	1.08	0.0	1.2	0.2	1.07
SS	0.0	10.2	4.4	2.70	0.0	10.2	4.4	2.71	0.0	10.2	4.4	2.72
OR	2.5	12.5	7.6	2.52	2.6	12.4	7.6	2.52	2.6	12.4	7.6	2.52
KA	0.0	4.6	0.9	1.38	0.0	4.6	0.9	1.38	0.0	4.6	0.9	1.39
CO	0.0	11.0	5.0	2.82	0.0	11.1	5.0	2.84	0.0	11.1	5.0	2.86
MK	0.0	0.2	0.0	0.16	0.0	0.1	0.0	0.18	0.0	0.0	0.0	0.18
LW	0.0	9.9	2.0	2.82	0.0	10.0	2.0	2.89	0.0	10.0	1.9	2.92
KT	0.1	21.9	10.1	5.65	0.0	21.9	10.1	5.74	0.0	22.0	10.1	5.79

'Stocks are those in Table 10.1 of Pella and Milner (1987).

Table 13.-Averages (ave), standard deviations (s), and upper ($\hat{p}_{0.975}$) and lower ($\hat{p}_{0.025}$) 95% confidence bounds from the bootstrap distributions (1000 resamplings) of estimated stock group (groups are summer and fall runs) percentages for the Yukon River chum salmon mixture by programs CONJA-S, EM, and GIRLS-B for specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%).

	GPA to stop											
	10%				50%				90%			
Program	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s
CONJA-S												
Summer run	54.3	89.0	73.0	9.08	54.1	89.5	73.2	9.34	53.8	89.7	73.2	9.48
Fall run	11.0	45.7	27.0	9.08	10.5	45.9	26.8	9.34	10.3	46.2	26.8	9.48
EM												
Summer run	55.2	87.4	72.9	8.41	54.5	88.9	73.2	9.10	53.8	89.5	73.2	9.45
Fall run	12.6	44.8	27.1	8.41	11.1	45.5	26.8	9.10	10.6	46.2	26.8	9.45
GIRLS-B												
Summer run	54.4	88.4	72.9	9.01	53.9	89.3	73.2	9.36	53.7	89.6	73.2	9.51
Fall run	11.6	45.5	27.1	9.01	10.7	46.1	26.8	9.37	10.4	46.4	26.8	9.51

Table 14.--Averages (ave), standard deviations (s), and upper ($\hat{p}_{0.975}$) and lower ($\hat{p}_{0.025}$) 95% confidence bounds from the bootstrap distributions (1000 resamplings) of estimated stock group (groups* are denoted as A-G and H-M) percentages for the West Coast sockeye salmon mixture by programs CONJA-S, EM, and GIRLS-B for specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%).

	GPA to stop											
	10%				50%				90%			
Program	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s
CONJA-S												
A-G	10.4	23.7	16.7	3.50	10.5	23.7	16.8	3.52	10.5	23.8	16.8	3.53
H-M	76.3	89.6	83.3	3.50	76.3	89.6	83.2	3.52	76.2	89.5	83.2	3.53
EM												
A-G	10.2	23.6	16.6	3.49	10.4	23.8	16.7	3.51	10.4	23.8	16.8	3.52
H-M	76.4	89.8	83.4	3.49	76.2	89.6	83.3	3.51	76.2	89.6	83.2	3.52
GIRLS-B												
A-G	10.6	24.0	16.9	3.53	10.5	23.8	16.8	3.53	10.5	23.8	16.8	3.52
H-M	76.0	89.4	83.1	3.53	76.2	89.5	83.2	3.53	76.2	89.5	83.2	3.52

'A-G = groups A through G; H-M = groups H through M (Pella et al. in press).

Table 15.--Averages (ave), standard deviations (s), and upper ($\hat{p}_{0.975}$) and lower ($\hat{p}_{0.025}$) 95% confidence bounds from the bootstrap distributions (1000 resamplings) of estimated stock group (groups* are denoted by letters A through M) percentages for the West Coast sockeye salmon mixture by the program CONJA-S for specified standard values of the stopping criterion (GPA of 10%, 50%, and 90%).

	GPA to stop											
	10%				50%				90%			
Group	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s	$\hat{p}_{0.025}$	$\hat{p}_{0.975}$	ave	s
A	0.0	8.1	3.7	2.05	0.0	8.2	3.7	2.05	0.0	8.1	3.7	2.04
B	0.0	1.0	0.1	0.57	0.0	1.2	0.1	0.65	0.0	0.8	0.1	0.68
C	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.00
D	0.0	0.5	0.0	0.30	0.0	0.1	0.0	0.34	0.0	0.0	0.0	0.33
E	0.0	14.0	6.5	3.42	0.0	14.0	6.5	3.47	0.0	14.1	6.5	3.51
F	0.0	12.1	5.5	3.34	0.0	12.5	5.5	3.45	0.0	12.6	5.5	3.49
G	0.0	6.4	0.9	1.70	0.0	6.8	0.9	1.83	0.0	6.8	0.9	1.86
H	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.00
I	0.1	19.4	8.8	5.12	0.0	21.5	9.6	6.09	0.0	23.0	10.4	6.60
J	5.0	37.3	19.8	8.34	1.1	41.4	18.9	10.94	0.0	45.8	17.3	12.30
K	0.0	2.4	0.3	0.80	0.0	0.3	0.0	0.41	0.0	0.0	0.0	0.36
L	32.9	62.1	48.0	7.45	29.4	62.5	47.5	8.51	26.5	64.3	47.7	9.37
M	0.0	16.9	6.4	4.66	0.0	18.1	7.3	5.09	0.0	18.5	7.9	5.32
A-G	10.4	23.7	16.7	3.50	10.5	23.7	16.8	3.52	10.5	23.8	16.8	3.53
H-M	76.3	89.6	83.3	3.50	76.3	89.6	83.2	3.52	76.2	89.5	83.2	3.53

* see Pella et al. (in press).

FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

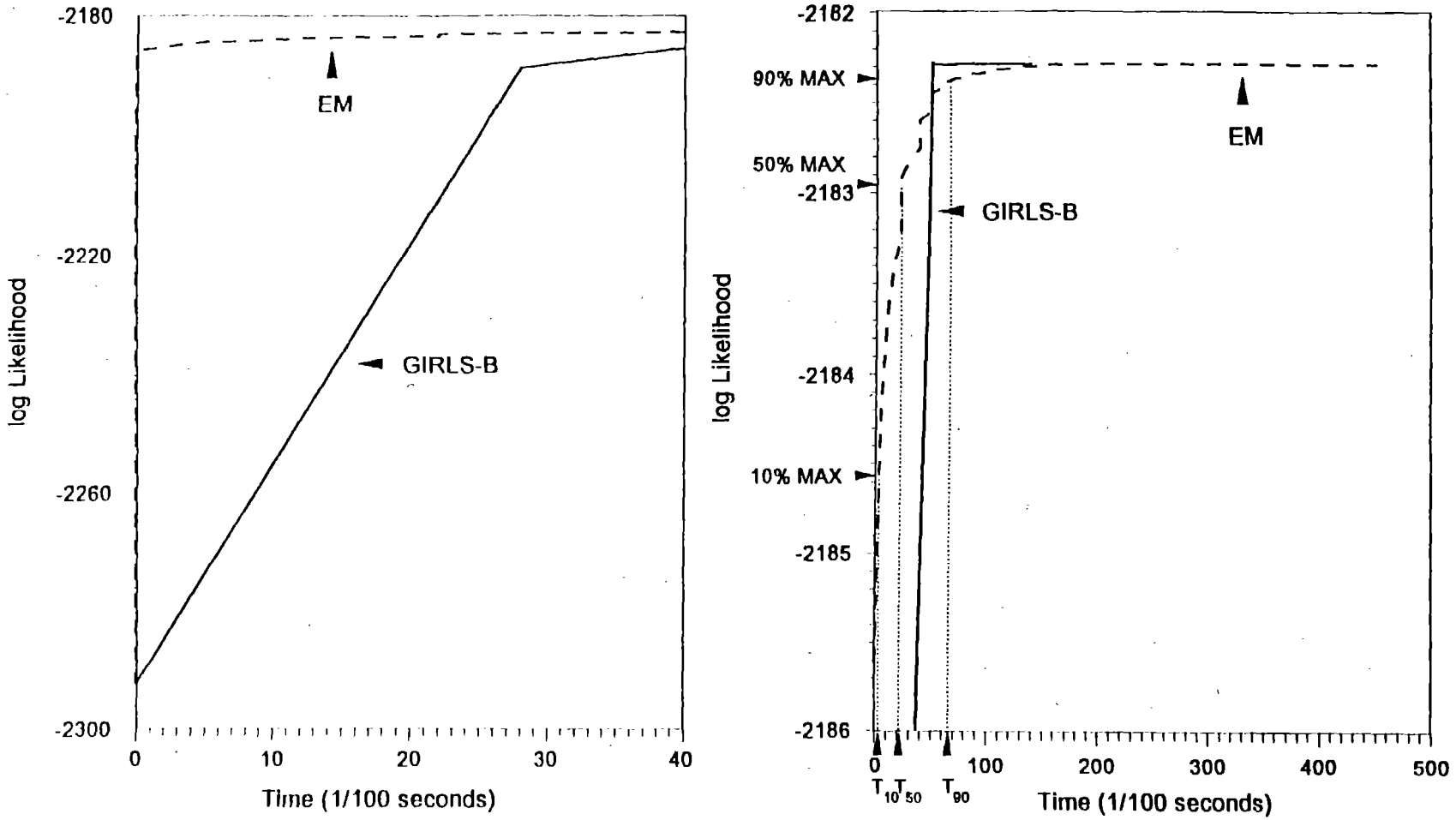

Figure 1.-Support function values achieved by two of the programs (GIRLS-B and EM) and total search time required of the processor. Data are from a cell of Experiment I. Early (left) and complete search (right) are illustrated using different scales for the support function. Levels of the support function guaranteed to equal at least 10%, 50%, and 90% of the maximum value (MAX) for the likelihood function correspond to critical search times, T_{10}, T_{50}, and T_{90}, illustrated here for EM (right).

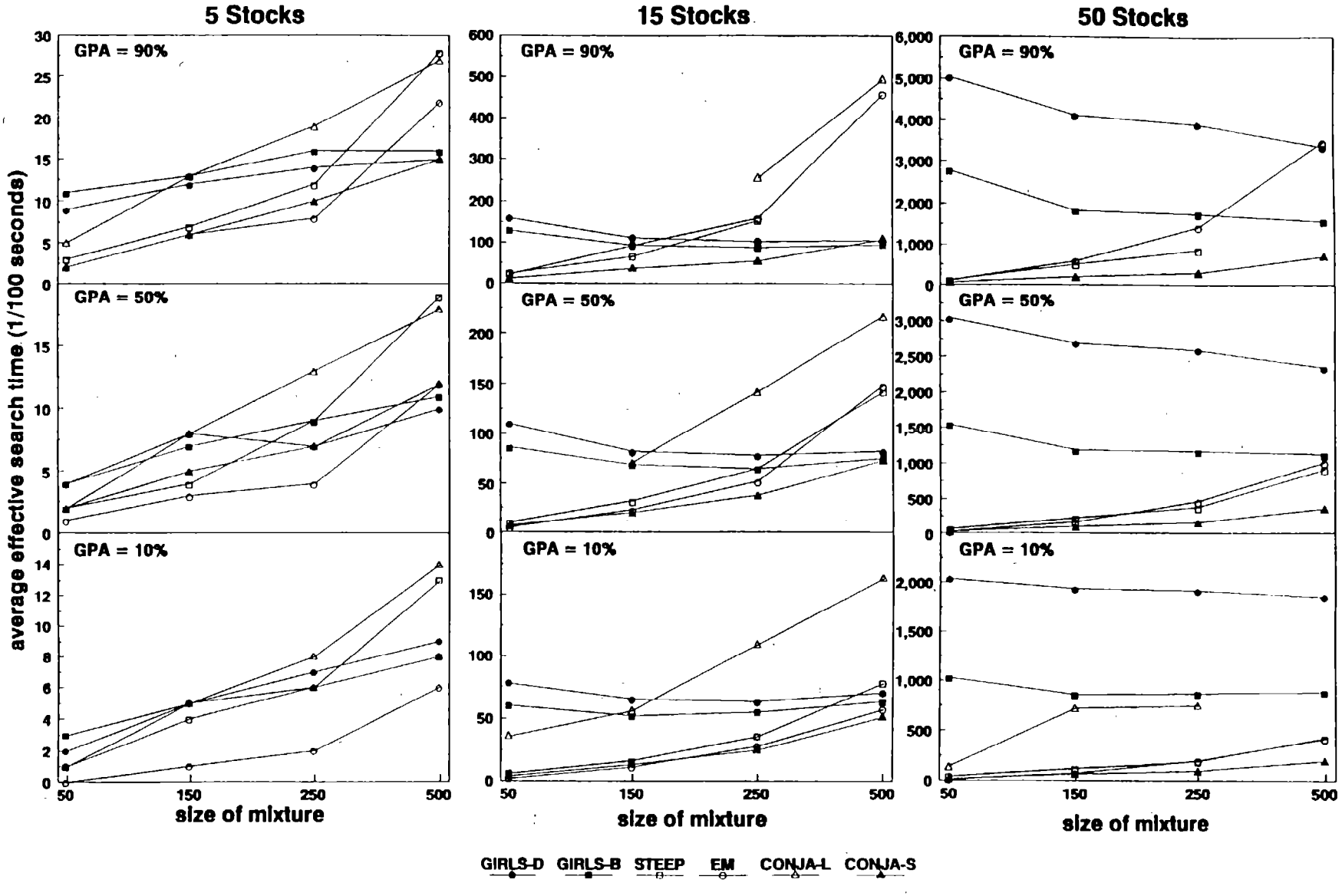

Figure 2.-For diverse baseline mixtures, average effective search times required by the six programs-CONJA-S, EM, GIRLS-B, GIRLS-D, STEEP, and CONJA-L-to achieve standard stopping criterion (GPA) values for varying numbers of baseline stocks and mixture sample sizes. Cells are plotted if average GPA equaled or exceeded that specified.

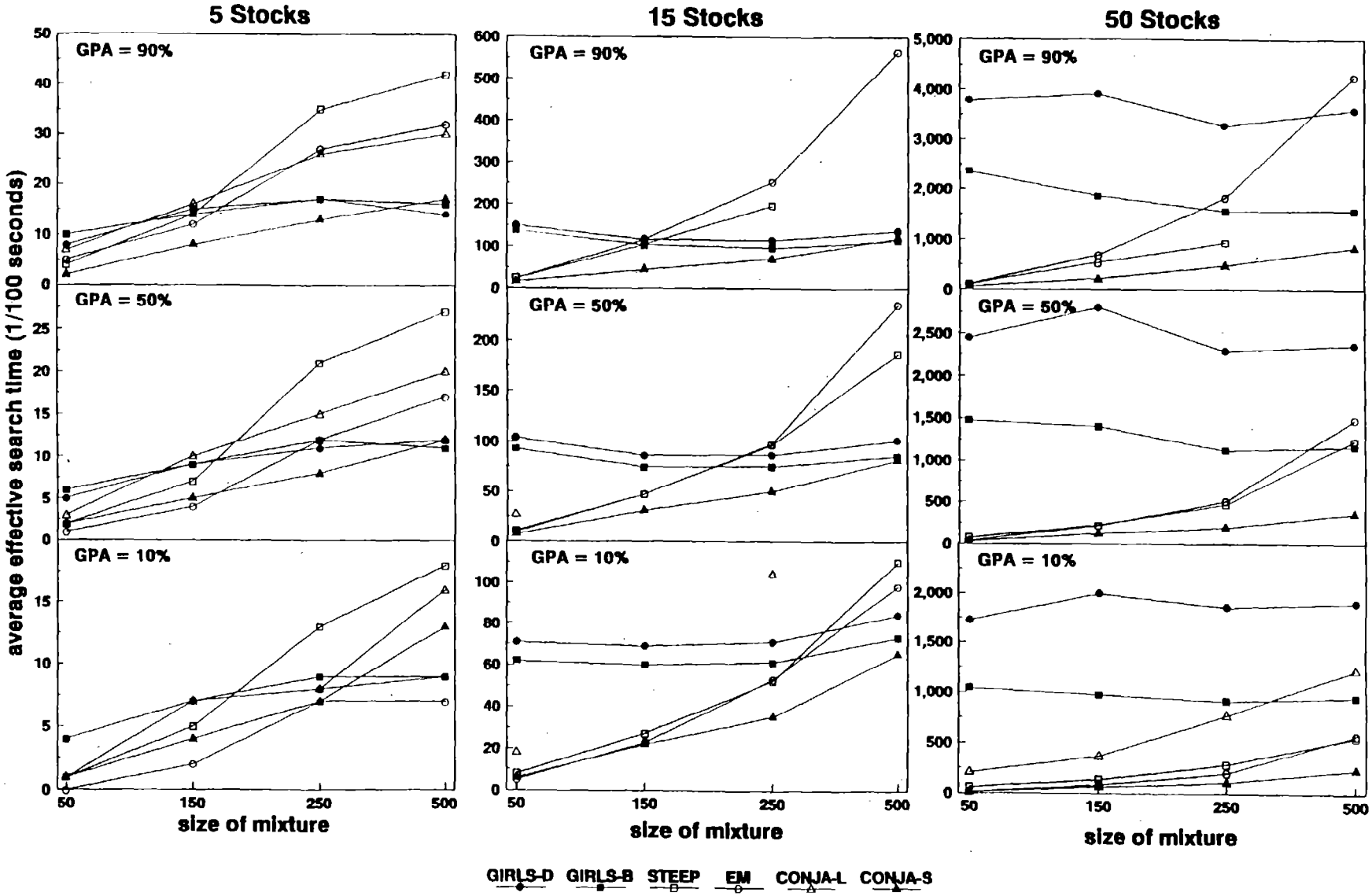

Figure 3.--For similar baseline mixtures, average effective search **times** required by the six programs-CONJA-S, EM, GIRLS-B, GIRLS-D, STEEP, and CONJA-L-to achieve standard stopping criterion (GPA) values for varying numbers of baseline stocks and mixture sample sizes. Cells are plotted if average GPA equaled or exceeded that specified.

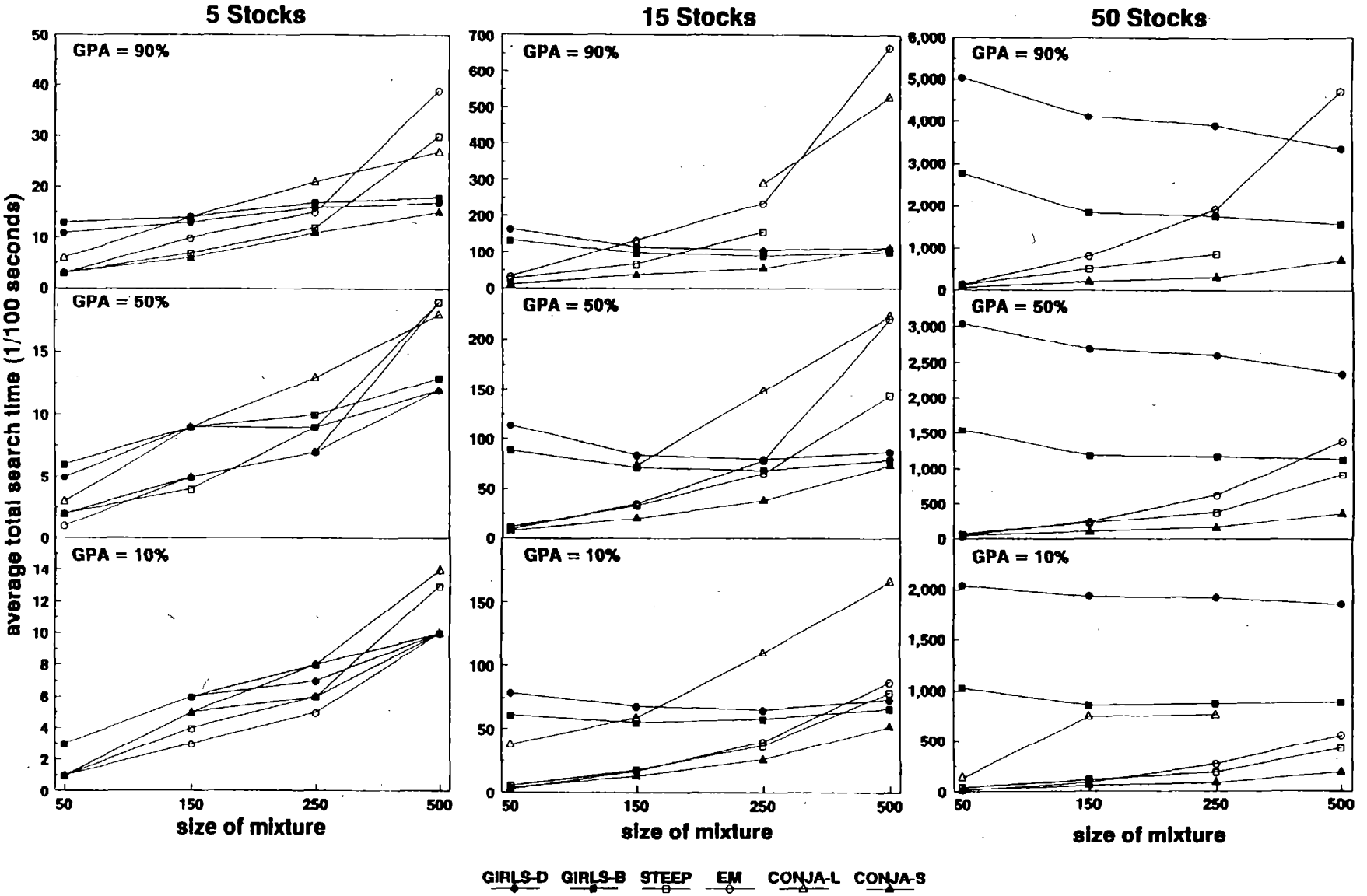

Figure 4.--For diverse baseline mixtures, average total search times required by the six programs-CONJA-S, EM, GIRLS-B, GIRLS-D, STEEP, and CONJA-L-to achieve standard stopping criterion (GPA) values for varying numbers of baseline stocks and mixture sample sizes. Cells are plotted if average GPA equaled or exceeded that specified.

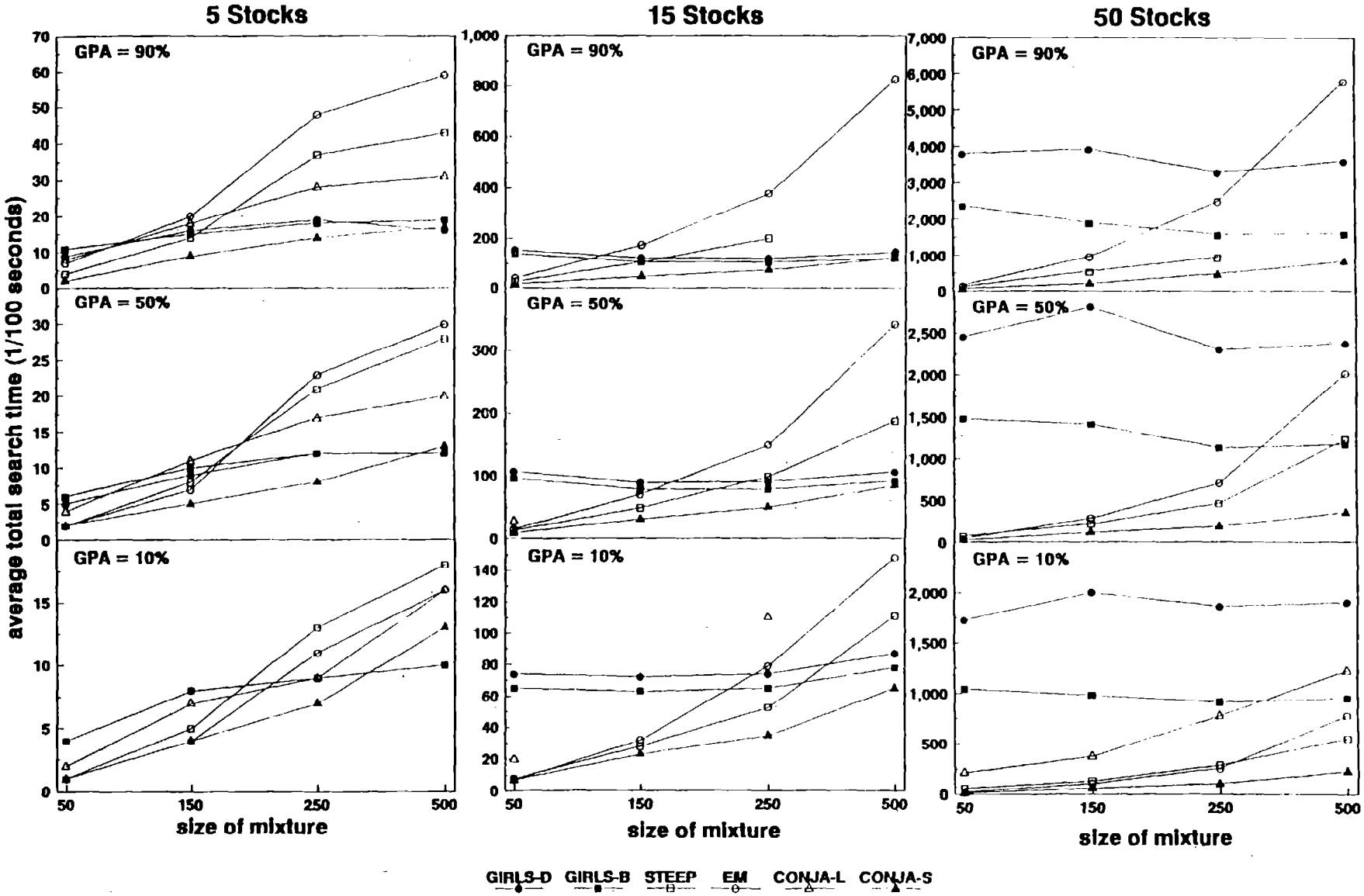

Figure 5.--For similar baseline mixtures, average total search times required by the six programs-CONJA-S, EM, GIRLS-B, GIRLS-D, STEEP, and CONJA-L-to achieve standard stopping criterion (GPA) values for varying numbers of baseline stocks and mixture sample sizes. Cells are plotted if average GPA equaled or exceeded that specified.'

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX

THIS PAGE INTENTIONALLY LEFT BLANK

Essential steps included in the programs STEEP, GIRLS-B, GIRLS-D, EM, CONJA-L, and CONJA-S are outlined next. initial point in the search is denoted as $p_{(0)}$, with typical element j denoted as $P_{j(0)}$.

STEEP (Steepest Ascent)

We define two sets of stock indices, boundary indices, and interior indices at any point p of the search to allow efficient continuation of searching after boundaries are encountered or closely approached. The boundary set B = $\{i_1, i_2, \ldots, i_b\}$ is composed of stock indices for which $0 \leq p_i < d$ if i belongs to B. The number of indices contained in B is n(B) = $b \geq 0$. The value of δ is arbitrary but meant to be small, and we use 10^{-7}. The interior set B* = $\{i_{b+1}, i_{b+2}, \ldots, i_c\}$ is composed of the remaining stock indices for which $\delta \leq p_i \leq 1$ if i belongs to B*. The number of indices contained in B* is $n(B^*) = (c-b) \geq 1$.

The program performs the following steps:

1) Computes the gradient vector $S = (s_1, s_2, \ldots, s_c)^T$, where the gradients are given by Equation 8 evaluated at $p = p_{(0)}$.

2) Finds the average gradient ($\bar{s}$) for interior stocks and, if applicable (b > 0), the largest gradient (s*) among boundary stocks together with the corresponding stock index, i^*, or

$$\bar{s} = \frac{1}{n(B^*)} \sum_{i \in B^*} s_i,$$

$$s^* = \begin{cases} 0 & n(B) = 0 \\ MAX\{s_{i_1}, s_{i_2}, \cdots, s_{i_b}\} = s_i. & n(B) > 0. \end{cases}$$

3) If $s^* \leq \bar{s}$, goes to step 4; if $s^* > \bar{s}$, adds the index i^* to B^* and deletes the index i^* from B. Then recomputes, $\bar{s}$ and continues to the next step.

4) Computes the corrected gradient vector, $S^0 = (s^\circ_1, s^\circ_2, \ldots, s^\circ_c)^T$, where

$$s_i^\circ = \begin{cases} s_i - \bar{s} & i \in B^* \\ 0 & i \in B \end{cases}.$$

5) Finds the maximum permissible step size, λ_{max}, together with the limiting interior stock index, i^+, as follows:

$$\lambda_i = -(p_{i(0)} - \delta) / s_i^\circ \quad i \in B^*, \; s_i^\circ < 0$$

$$\lambda_{max} = MIN\{\lambda_i | i \in B^*, \; s_i^\circ < 0\} = \lambda_{i^+}.$$

6) Computes the optimum step size , λ_{opt}, from the line search based on the Taylor's series approximation to the support function.

7) Chooses a step size $\lambda = \lambda_{opt}$ if $\lambda_{opt} \leq \lambda_{max}$. Otherwise, chooses $\lambda = \lambda_{max}$, removes the index i^+ from B*, and adds the index i^+ to B.

8) Computes the new point in the search path $p_{(1)} = p_{(0)} + \lambda s^{\circ}$.

9) Sets $p_{(0)} = p_{(1)}$. If specified GPA is achieved at $p_{(0)}$, stops the search; otherwise, returns to step 1.

GIRLS-B or GIRLS-D (Iteratively Reweighted Least Squares Basic and Difference Versions)

1) Computes S from Equation 8 and R (see Equation 15) using $\mathbf{p} = p_{(0)}$, and either G^TRG and G^TRy -(see Equation 15 of basic version) or D^TRD and D^TRy (see Equation 22 of difference version).

2) Solves equation system, Equation 20 or 24, for $p = p_{(1)}$.

3) Uses the line search based on the Taylor's series to find the optimum step size (λ_{opt}) along the line passing through $p_{(0)}$ and $p_{(1)}$. Sets a counter K = 0.

4) Computes $p_{(2)} = p_{(0)} + (0.99)^K \lambda_{opt} \Delta p$, where $\Delta p = p_{(1)} - p_{(0)}$. Verifies $p_{(2)}$ is feasible and that the probabilities of all observed types in the mixture sample remain positive, i.e., $\Sigma p_{j(2)} g_{h,j} > 0$ for each $h = 1, \ldots, H$. If both conditions are satisfied, goes to step 5. Otherwise, sets $K = K + 1$ and returns to the beginning of this step.

5) Sets $p_{(0)} = p_{(2)}$. If specified GPA is achieved at $p_{(0)}$, stops the search; otherwise, returns to step 1.

EM (Expectation Maximization)

1) Computes $S = (s_1, s_2, \ldots, s_c)^T$ where s_i is given by Equation 8 evaluated at $p = p_{(0)}$.

2) Computes $p_{(1)}$, using Equation 27.

3) Sets $p_{(0)} = p_{(1)}$. If specified GPA is achieved at $p_{(0)}$, stops the search; otherwise, returns to step 1.

CONJA-L or CONJA-S (Conjugate Gradient with Logarithm or Square Root Transform)

1) Computes $\tilde{S}^{(0)} = (\tilde{s}_1^{(0)}, \tilde{s}_2^{(0)}, \ldots, \tilde{s}_c^{(0)})^T$ equal to the gradient $\tilde{S}$ using Equation **30 (CONJA-L)** or Equation 36 **(CONJA-S)** evaluated at $p = p_{(0)}$. Sets the first search direction in the transformed stock composition space, $D^{(0)} = (d_1^{(0)}, d_2^{(0)}, \ldots, d_c^{(0)})$= equal to $\tilde{S}^{(0)}$ **(CONJA-S)**; or **(CONJA-L)** sets $d_i^{(0)} = \tilde{s}_i^{(0)}$ unless both $\log P_{i(0)} < -20$ and $\tilde{s}_i^{(0)} < 0$ (in which case, sets $d_i^{(0)} = 0$) $i = 1, 2, \ldots, c$.

2) Computes the optimum step size, $\lambda_{opt}^{(0)}$, from the line search in the transformed stock composition space based on the Taylor's series approximation to the support function at $u_{(0)}$ with direction $D^{(0)}$ (see Equation 32). $u_{(0)}$ is computed from $p_{(0)}$ using Equation 28 **(CONJA-L)** or Equation 34 (CONJA-S). Sets a counter J = 0.

3) Computes the new point in the search path in the transformed stock composition space $u_{(1)} = u_{(0)} + (0.99)^J \lambda_{opt}^{(0)} D^{(0)}$.

4) Sets a counter k = 1. Computes $p_{(1)}$, corresponding to $u_{(1)}$ by Equation 29, and goes to step 5 **(CONJA-L)** ; or **(CONJA-S)** computes $p_{(1)}$ corresponding to $u_{(1)}$ by Equation 35 and verifies that probabilities of all observed types in the mixture sample remain positive, i.e., $\Sigma p_{j(1)} g_{h,j} > 0$ for each $h = 1, \ldots, H$. If the condition is satisfied, goes to step 5. Otherwise, sets J = J + 1 and returns to step 3.

5) Computes the gradient by Equation **30 (CONJA-L)** or Equation 36 **(CONJA-S)**, evaluated at $p = p_{(k)}$, which is denoted as $\tilde{S}^{(k)}$, and sets the kth search direction, $D^{(k)}$, to

$$D^{(k)} = \tilde{S}^{(k)} + \left[\frac{|\tilde{S}^{(k)}|^2}{|\tilde{S}^{(k-1)}|^2} \right] \tilde{S}^{(k-1)},$$

$$\text{where} \quad |\tilde{S}^{(k)}|^2 = \sum_{i=1}^{c} \tilde{s}_i^{(k).2} .$$

Sets the counter $J = 0$. Goes to the next step **(CONJA-S)** ; or **(CONJA-L)** modifies the direction vector $D^{(k)}$ above by setting $d_i^{(k)} = 0$ if $\log (P_{i(k)}) < -20$ and $di^{(k)} < 0$, $i = 1, 2, \ldots, c$.

6) Finds the next point in the transformed space, $u_{(k+1)} = u_{(k)} + (0.99)^J \lambda_{opt}^{(k)} D^{(k)}$, by line search in the transformed stock composition space based on the Taylor's series approximation to the support function at $u_{(k)}$ with direction $D^{(k)}$ (see Equation 32). Computes $p_{(k+1)}$ from $u_{(k+1)}$ by Equation 29 and goes to step 7 (CONJA-L); or (CONJA-S) computes $p_{(k+1)}$ from $u_{(k+1)}$ by Equation 35 and verifies that

probabilities of all observed types in the mixture sample remain positive, i.e., $\Sigma p_{j(k+1)} g_{h,j} > 0$ for each $h = 1, \ldots H$. If the condition is satisfied, goes to step 7. Otherwise, sets J = J + 1 and returns to the beginning of this step.

7) If specified GPA is achieved, stops; otherwise, continues on to the next step 8.

8) If $k < c$, sets $k = k + 1$ and repeats steps 5, 6, and 7. Otherwise, restarts the search at step 1 using the last point $p_{(c+1)}$ found in the search as $p_{(0)}$.

RECENT TECHNICAL MEMORANDUMS

Copies of this and other NOAA Technical Memorandums are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22167 (web site: *www.ntis.gov)*. Paper and microfiche copies vary in price.

AFSC-

60 YANG, M-S. 1996. Diets of the important groundfishes in the Aleutian Islands in summer 1991, 105 p. NTIS No. PB96-147582.

59 MARTIN, M. H., and D. M. CLAUSEN. 1995. Data report: 1993 Gulf of Alaska bottom trawl survey, 217 p. NTIS No. PB96-135561.

58 QUEIROLO, L. E., L. W. FRITZ, P. A. LIVINGSTON, M. R. LOEFFLAD, D. A. COLPO, and Y. L. DEREYNIER. 1995. Bycatch, utilization, and discards in the commercial groundfish fisheries of the Gulf of Alaska, eastern Bering Sea, and Aleutian Islands, 148 p. NTIS No. PB96-125547.

57 SMALL, R. J., and D. P. DEMASTER. 1995. Alaska marine mammal stock assessments 1995-57, 93p. NTIS No. PB95-274734.

56 DORN, M. W., S. M. FITZGERALD, M. A. GUTTORMSEN, and M. R. LOEFFLAD. 1995. An evaluation of North Pacific groundfish observer program methods of haul weight estimation, 31 p. NTIS No. PB95-271151.

55 PELLA, J., R. RUMBAUGH, and M. DAHLBERG. 1995. Incidental catches of salmonids in the 1991 North Pacific squid driftnet fisheries, 33 p. NTIS No. PB95-252722.

54 KINOSHITA, R. K., A. GRIEG, and J. M. TERRY. 1995. Economic status of the groundfish fisheries off Alaska, 1993, 108 p. NTIS No. PB95-252714.

53 WING, B. L., and D. J. KAMIKAWA. 1995. Distribution of neustonic sablefish larvae and associated ichthyoplankton in the eastern Gulf of Alaska, May 1990, 48 p. NTIS No. PB95-241519.

52 HONKALEHTO, T., and N. WILLIAMSON. 1995. Echo integration-trawl survey of walleye pollock (Theragra chalcogramma) in the southeast Aleutian Basin during February and March, 1994, 39 p. NTIS No. PB95-219424.

51 CLARY, J. C. (editor). 1995. Poster abstracts and manuscripts from the Third International Conference on Marine Debris, May 8-13, 1994, Miami, Florida, 108 p. NTIS No. PB95-210837.

50 MUNRO, P. T., and R. Z. HOFF. 1995. Two demersal trawl surveys in the Gulf of Alaska: Implications of survey design and methods, 139 p. NTIS No. PB95-195061.

49 STARK, J. W., and D. M. CLAUSEN. 1995. Data Report: 1990 Gulf of Alaska bottom trawl survey, 221 p. NTIS No. PB95-194825.

48 NARITA, R., M. GUTTORMSEN, J. GHARRETT, G. TROMBLE, and J. BERGER. 1994. Summary of observer sampling of domestic groundfish fisheries in the northeast Pacific Ocean and eastern Bering Sea, 1991, 540 p. NTIS No. PB95-190963

47 DORN, M. W., E. P. NUNNALLEE, C. D. WILSON, and M. E. WILKINS. 1994. Status of the coastal Pacific whiting resource in 1993, 101 p. NTIS No. PB95-176467.

46 SINCLAIR, E. H. (editor). 1994. Fur seal investigations, 1993, 93 p. NTIS No. PB95-178943.

45 SINCLAIR, E. H. (editor). 1994. Fur seal investigations, 1992, 190 p. NTIS No. PB95-173472.