PARAMETER ESTIMATION FOR THE PELLA-TOMLINSON STOCK PRODUCTION MODEL UNDER NONEQUILIBRIUM CONDITIONS

D. RIVARD AND L. J. BLEDSOE[1]

ABSTRACT

To estimate the parameters of the Pella-Tomlinson model, as restructured by Fletcher in this issue, we suggest a derivative-free version of the Levenberg-Marquardt algorithm, along with an algorithm that locates starting values for the iterative procedure. The iterative method of Levenberg-Marquardt was applied to two versions of the restructured model: five parameters were estimated in the first version and three in the second, the latter preventing degeneracy of the model to exponential form. We discuss in particular the causes of the degeneracies associated with previous applications of the model. Such faults lie, inherently, with the mathematical indeterminacy of the system equations themselves, so that all nonlinear estimation methods will tend to be inefficient in the absence of external constraints. The effectiveness of the Levenberg-Marquardt method was evaluated by Monte-Carlo simulation. As examples, we analyzed catch-effort data from the yellowfin tuna fishery of the eastern Pacific and catch-effort data from the Pacific halibut fishery (Area 2 of the International Pacific Halibut Commission).

Parameter estimation has been the greatest source of difficulty in applying the generalized stock-production model to management schemes, and the problem has attracted considerable attention. Pella and Tomlinson (1969) fitted the model to the catch-effort history of a fishery under nonequilibrium conditions by means of a search algorithm, and although good graphical fits are generally obtained by that procedure, unreasonable parameter estimates are frequently generated owing to the lack of internal constraints on parameter values (see Ricker 1975, example 13.6). Fox (1971) constructed a stochastic representation of the generalized production model and employed simulation to infer the effects of random variability in catch data. Fox suggested that variation in catch increases with the size of the catch (additive proportional error) and he gave a new formulation of the minimization criterion for the Pella-Tomlinson procedure. Walter (1975) suggested a graphical method for calculating the coefficients of the Graham-Schaefer model. Walter's procedure requires the plotting of catch per effort against effort data and then correcting for disequilibrium of the fishery. Fox (1975) also described a procedure for fitting the Pella-Tomlinson model that requires equilibrium approximations. And finally, Schnute (1977) derived linear and nonlinear methods for finding estimates of the coefficients of the Schaefer model; his method also includes a way of measuring the uncertainty of the estimates.

Fletcher (1978b) presented a reparametrization of the generalized production model and explains how the tendency to ill-determined parameter estimates arises from a conflict between variable graph curvature and its coupling with the coefficients of the system. In this paper, we take advantage of that restructuring and examine the use of a derivative-free version of the Levenberg-Marquardt numerical optimization algorithm, together with a Runge-Kutta differential equation solver, to estimate parameters in Fletcher's differential form of the Pella-Tomlinson model. Estimates of the variability in the coefficients are also provided, and the complete estimation procedure is analyzed by a Monte-Carlo simulation. The estimation problem is finally reformulated to prevent ill determination of the parameters and degeneracy of the model to exponential form.

MODEL AND NOTATION

As indicated by Fletcher (1975, 1978b), the generalized production model can be generated by the single differential equation

[1]Center for Quantitative Science in Forestry, Fisheries and Wildlife, University of Washington, Seattle, WA 98195.

Manuscript accepted March 1978.
FISHERY BULLETIN: VOL. 76, NO. 3, 1978.

$$\dot{B} = \gamma m \left[\frac{B}{B_\infty}\right] - \gamma m \left[\frac{B}{B_\infty}\right]^n - \dot{Y}, \qquad (1)$$

with the quantity γ wholly a function of n in the relationship

$$\gamma = \frac{n^{n/n-1}}{n-1}. \qquad (2)$$

$B(t)$, the solution of $\dot{B}$, represents the stock size at time t, while $Y(t)$, the solution of $\dot{Y}$, represents the cumulative catch of the stock. Parameter B_∞ is the maximum stock size of the unexploited population, while m is the maximum productivity in the productivity function or the maximum sustainable yield (MSY) in the complete exploitation model. Exponent n controls the location of the inflexion point in the latent productivity function of the stock. Therefore, parameters B_∞, m, and n are nonnegative. With this new formulation of the system equations, the sign reversals of the coefficients at the turning point $n = 1$ are now automatic. Also the parameters are expressed in more meaningful terms for the fishery scientist, and some aspects of parameter estimation are simplified thereby.

By presuming that $f(t)$ units of effort operate on the population over time increment dt, the yield rate is often put into the instantaneous form

$$\dot{Y} = q f(t) B(t) \qquad (3)$$

where q is the catchability coefficient. Equations (1) and (3) constitute a coupled system of nonlinear differential equations. The system, as it is formulated in Equations (1) and (3), represents the continuous-time model. In practice, though, for each finite time interval τ over which yield statistics are integrated, fishing effort is usually assumed to be constant. It follows that $f(t)$ must be a step function that describes the effort as being constant over each time interval τ with abrupt changes at the end of each period. Then the effort required, over one time interval, to maintain maximum productivity is given by

$$f_{MSY} = \frac{\gamma m}{q B_\infty}\left[\frac{n-1}{n}\right], \qquad (4)$$

and at MSY the yield per unit of effort U_{MSY} is obtained by dividing m by f_{MSY}.

DETERMINISTIC MODEL AND STOCHASTIC REPLICATES

The Pella-Tomlinson system has the five parameters m, B_∞, n, q, and B_0. The fifth parameter, initial population size B_0, is needed to specify a particular solution of Equation (1). By taking the following arbitrary values for the parameters,

$$m = 1.36 \times 10^6 \qquad q = 4.60 \times 10^{-6}$$
$$B_\infty = 3.48 \times 10^6 \qquad B_0 = 3.48 \times 10^6$$
$$n = 1.80,$$

we constructed an example of a fishery over the course of 20 yr with $f(t)$ increasing within the first 10 yr and stabilizing thereafter (Table 1).

We also constructed 20 stochastic replicates of the deterministic catch history. In all the stochastic versions we assume additive proportional error terms ϵ_i (with i the annual index), consisting of 20 sets of 20 values each of normally distributed, independent random variables with expected means of zero and standard deviations (σ) of 0.025, 0.050, 0.075, 0.100, 0.125, 0.150, 0.175, 0.200 (12 replicates), and 0.250. Although Fox (1975) takes a similar approach, we recognize the fact that serial correlation of errors is likely to exist in natural data. As put by J. J. Pella in a personal communication, "If yield is above average in one year because the population is above average, it will probably be above average in the following year." But at this stage of the analysis, the explicit consideration of serially correlated errors would only complicate the estimation problem unneces-

TABLE 1.—Simulation of a logistic stock under exploitation (deterministic model).

Time (yr)	Catch	Effort	$U = \frac{\text{Catch}}{\text{Effort}}$
1	79,529	5,000	15.91
2	156,989	10,000	15.70
3	306,812	20,000	15.34
4	516,722	35,000	14.76
5	893,564	65,000	13.75
6	1,129,740	90,000	12.55
7	1,402,830	125,000	11.22
8	1,565,040	160,000	9.78
9	1,631,400	195,000	8.37
10	1,614,130	230,000	7.02
11	1,478,930	250,000	5.92
12	1,434,220	300,000	4.78
13	1,196,850	320,000	3.74
14	988,745	320,000	3.09
15	895,127	350,000	2.56
16	737,908	350,000	2.11
17	607,937	300,000	2.03
18	649,627	300,000	2.17
19	632,633	250,000	2.53
20	773,600	250,000	3.09

sarily; our immediate purposes are better served by the simpler form of the ϵ_i. We want to observe the response of the estimation procedure to realistic levels of stochastic error, but we also want to avoid wrong interpretations in those cases where parameter values might be ill determined because of some inherent fault of the estimation procedure itself and not because of some complication of the error structure. Therefore, the stochastic replicates, as well as the objective function of our estimation procedure, are constructed on the assumption of independence of errors. As we shall see in a subsequent discussion, the following estimation procedure is indeed robust with respect to that assumption.

PARAMETER ESTIMATION PROCEDURE

In its general form, the solution of the nonlinear model described by Equation (3) can be written as

$$\hat{Y}_i = g(f_1, f_2, \ldots, f_i; \Theta) \quad i = 1, 2, \ldots, r \qquad (5)$$

where[2] $\Theta = [m, B_\infty, n, q, B_0]^T$.

Quantity r represents the total number of observations over time, f_i the fishing effort during time interval i, and $\hat{Y}_i$ the predicted yield (biomass or number) over the interval i. Following Fox (1971), we also consider an error term ϵ_i proportional to population size and equivalent in terms of yield to the form

$$Y_i = \hat{Y}_i + \hat{Y}_i \epsilon_i, \qquad (6)$$

where Y_i represents the observed yield over the interval i. Then the error is described by the relationship

$$\epsilon_i = (Y_i - \hat{Y}_i) / \hat{Y}_i. \qquad (7)$$

Least squares estimation of Θ by a vector $\hat{\Theta}$ requires minimization of the function

$$S(\hat{\Theta}) = \sum_i \epsilon_i^2. \qquad (8)$$

In terms of residuals $(Y_i - \hat{Y}_i)$, this is equivalent to

$$S(\hat{\Theta}) = \sum_i W_i (Y_i - g(f_1, \ldots, f_i; \hat{\Theta}))^2, \qquad (9)$$

where the W_i are statistical weights. That is, from Equation (7),

$$W_i = \hat{Y}_i^{-2}. \qquad (10)$$

If $S(\hat{\Theta})$ were an analytic form, we would find $\hat{\Theta}$ by writing the normal equations

$$\left[\frac{\partial S(\hat{\Theta})}{\partial \Theta_i}\right] = 0.$$

Since S must be calculated via numerical methods, we will instead consider $S(\hat{\Theta})$ as a continuous function that describes a hypersurface in a five-dimensional parameter space; that space must be searched for the appropriate minimum value of $S(\hat{\Theta})$. The iterative process of successive approximations which we employ is an adaptation of the Levenberg-Marquardt technique (Levenberg 1944; Marquardt 1963). Given some initial estimate Θ_0, the method generates a sequence of estimates $\hat{\Theta}_j$ from the inductive relation

$$\hat{\Theta}_{j+1} = \hat{\Theta}_j - \left[\beta_j I_5 + J_j^T J_j\right]^{-1} J_j^T \epsilon_j. \qquad (11)$$

In Equation (11), β_j is a positive constant, I_5 the identity matrix of order 5, J_j an r by five matrix having elements $\partial\epsilon_i/\partial\hat{\theta}_k$ $(i = 1, \ldots, r; k = 1, \ldots, 5)$, and ϵ_j the vector of errors after j iterations. The method combines the best festures of both the gradient and the Taylor series methods and avoids their most serious limitations (Conway et al. 1970). We employ a FORTRAN computer program which incorporates a derivative-free version of the Levenberg-Marquardt method (Brown and Dennis 1972), and we approximate the solution of the model differential equations (1) and (3) by a fourth order Runge-Kutta algorithm for numerical integration. The general structure of the program is shown by the flow diagram of Figure 1. Since all the parameters have to be positive, we also constrain the optimization by transforming each component of $\hat{\Theta}$ by its absolute value before evaluating the model.

ACCURACY OF RESULTS

Since the solution to the least-squares estimation problem is the result of a numerical search along the $S(\hat{\Theta})$ hypersurface, we do not generate

[2]The notation [. . .] indicates that a row vector or matrix is formed of the elements enclosed by brackets.

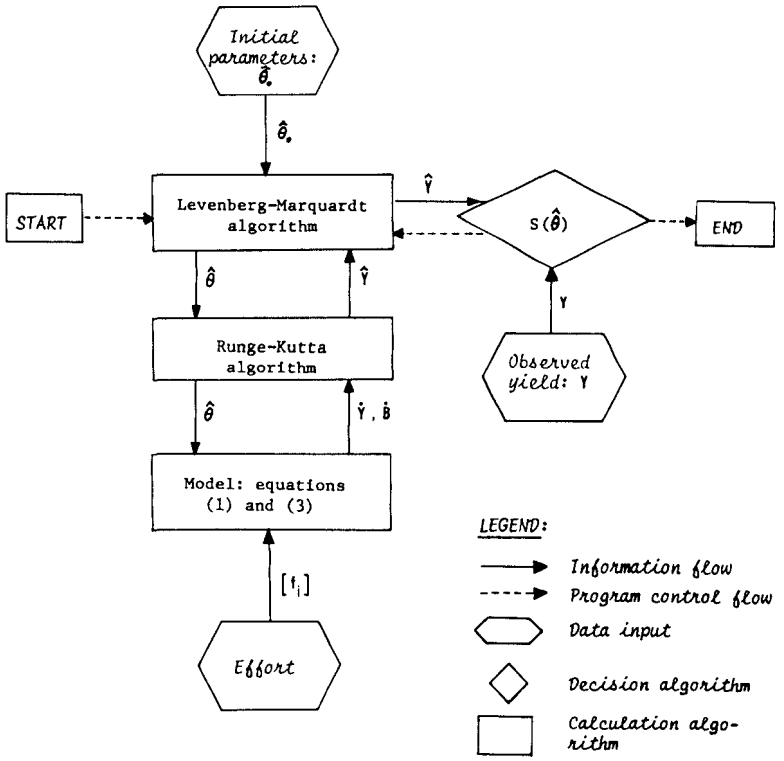


FIGURE 1.—Information flow diagram for the computer program written to estimate the coefficients of the generalized production model.

an analytical form for the uncertainties in the final values of the parameters (Bevington 1969). By letting $S(\hat{\Theta})$ be the weighted residual sum of squares for the final parameter estimates, however, the variance-covariance matrix of the estimates (Bard 1974) can be approximated by

$$V_{\Theta} = \left[\mathbf{J}^{\mathrm{T}}\,\mathbf{J}\right]^{-1} S(\hat{\Theta})/(r-5). \qquad (12)$$

Some idea of the joint variability of the parameters can be obtained by evaluating the ellipsoidal confidence region, based on the assumption that the linearized form has validity around $\hat{\Theta}$ (Draper and Smith 1966). The confidence region is then given by

$$[\Theta-\hat{\Theta}]\;\mathbf{J}^{\mathrm{T}}\;\mathbf{J}[\Theta-\hat{\Theta}]^{\mathrm{T}} \leq \frac{5S(\hat{\Theta})}{r-5}\;F(5, r-5, 1-\alpha), \qquad (13)$$

where $F(5, r\text{-}5, 1\text{-}\alpha)$ is the standard tabulated F-statistic. The ellipsoid is not a true confidence region, of course, since the dependent variable, $\dot{Y}$, is a nonlinear function of $\hat{\Theta}$. The intervals obtained are valid to the extent that a linear approximation in the neighborhood of $\hat{\Theta}$ is appropriate. A necessary and sufficient condition for the F-distribution to be appropriate here is that differences in true and estimated parameter values are independent and approximately normally distributed with zero mean and equal variance.

DETERMINATION OF STARTING VALUES

In order to reduce the number of iterations required to minimize Equation (8), reasonably accurate starting values should be employed. Starting values can be calculated from a linearization and simplification of the basic model.

STEP 1. By using $Y_1, Y_2, \ldots, Y_r$ and $f_1, f_2, \ldots f_r$, find an estimate of q from the Delury technique. Note that this procedure generally underestimates q (see Ricker 1975). Correction for q will be provided in step 4.

STEP 2. Find estimates of B_{t+1} from the equation

$$B_{t+1} = (\overline{B}_{t,t+1} + \overline{B}_{t+1,t+2})/2, \qquad (14)$$

where (by assuming $f_{t,t+1}$ constant over the interval $t,t+1$)

$$\bar{B}_{t,t+1} = Y_{t,t+1}/q\, f_{t,t+1}. \tag{15}$$

Note that $Y_{t,t+1}$ and $f_{t,t+1}$ correspond to Y_i and f_i of Equations (6) and (5).

STEP 3. Let $n = 2$, as in the Graham-Schaefer model, and estimate m and B_∞ by fitting the linear model

$$y_t = \alpha_0 + \alpha_1 x_t, \tag{16}$$

where $y_t = \dfrac{dB_t}{B_t\, dt} + qf_t, \quad x_t = B_t^{n-1},$

$$\alpha_0 = \gamma m/B_\infty, \quad \alpha_1 = \gamma m/B_\infty^n.$$

Equation (16) is derived from Equations (1) and (3). However, Equation (16) requires an estimate of the relative growth rate $dB_t/B_t dt$, say R_t. As suggested by Causton (1969), the mean value of R between t and $t+2$ is given by

$$\bar{R}_{t,t+2} = (\ln B_{t+2} - \ln B_t)/2. \tag{17}$$

For the purpose of fitting Equation (16), quantity $\bar{R}_{t,t+2}$ may be considered an estimate of R_{t+1}, which corresponds to B_{t+1}. Whence. Equation (16) provides estimates of m and B_∞ as

$$m = \frac{\left[\alpha_1\, \alpha_0{}^n\right]^{(1/1-n)}}{\gamma}, \tag{18}$$

$$B_\infty = \left[\frac{\alpha_1}{\alpha_0}\right]^{1/1-n}. \tag{19}$$

STEP 4. Steps 2 and 3 are repeated iteratively for increasing values of q. The value of q which provides the minimum residual sum of squares $[\sum_i (Y_i - \hat{Y}_i)^2]$ is accepted as the appropriate starting value for q.

STEP 5. Step 3 is repeated iteratively for increasing values of n, parameter q being kept constant. The value of n which provides the minimum residual sum of squares $[\sum_i (Y_i - \hat{Y}_i)^2]$ is accepted as the appropriate starting value for n. In the last iteration, Equations (18) and (19) provide estimates of m and B_∞. Finally, B_0 is approximated by

$$B_0 = 2\bar{B}_{0,1} - B_1 \tag{20}$$

where B_1 and $\bar{B}_{0,1}$ are estimated by Equations (14) and (15), respectively.

Steps 1 through 5 provide a set of starting values for the optimization algorithm (11). Usually the starting values are near the solution and few iterations will be needed. Of course, it would be possible to derive algorithms for more accurate starting values, but our purpose here is to find a rough estimate for each coefficient and to let the iterative procedure (11) converge to the minimum. Sometimes, by experience or by prior information, it is possible to provide starting values as satisfactory as those provided by the algorithm given above.

MONTE-CARLO SIMULATIONS

The parameter values that we chose to generate the data of Table 1 (deterministic model) were recovered exactly by the estimation procedure. Results of fitting 18 stochastic versions of the deterministic model are also included in Table 2. Based on our simulation results, there do not appear to be any serious problems with bias of parameter estimates. The bottom line of Table 2, which gives the coefficients of variation of the parameter estimates, reveals that estimates of the three parameters of principal interest to the manager have the smallest variability. Those parameters are maximum sustainable yield ($\hat{m}$, C.V. = 14%), optimal effort level ($\hat{f}_{MSY}$, C.V. = 6%), and yield per unit of effort at optimum effort ($\hat{U}_{MSY}$, C.V. = 9%). Our results confirm the observations of Fox (1971) and Pella and Tomlinson (1969) on the robustness of m and f_{MSY} with respect to error in the measurement of the yield data. From Table 2, we can also compare variance estimates from Equation (12) with variance of estimates for 10 replicates at $\sigma = 0.200$. Equation (12) appears to give (approximately) unbiased estimates of the variance of the sampling distribution of $\hat{\boldsymbol{\Theta}}$. Also, out of the 19 cases considered, the true parameter value lay outside the arbitrary ±2 (SD) confidence interval twice for $\hat{m}$ and only once each for $\hat{B}_\infty$, $\hat{n}$, and $\hat{B}_0$. Although we did not employ an extensive Monte-Carlo simulation, our results suggest that the normal approximation to the sampling distribution of $\hat{\boldsymbol{\Theta}}$ is an acceptable approximation, at least for management purposes.

In a few additional simulations (replicates 13 and 15), parameters obtained by the five-parameter procedure were ill determined. A parameter is considered ill determined if its estimated value responds unreasonably to seemingly insignificant variations in the data (Bard 1974). The basic difficulty is that the model is extremely general and capable of several types of behavior over the space of Θ. In the Pella-Tomlinson system, ill determination often occurs whenever an iteration of the algorithm (11) gives an estimate of Θ such that the point $(\hat{m}, \hat{f}_{MSY})$ of the yield-effort plane lies outside the concentration of data. In such a circumstance the exponent $\hat{n}$ takes on smaller and smaller values in the successive iterations and the solution of system (1) and (3) degenerates to an exponential form for which only four parameters are required for uniqueness. That is, as $n \to 0$, in Equation (1), then $(B/B_\infty)^n \to 1$ and $\gamma \to -1$. The five-parameter procedure then overprescribes the system, which in turn predisposes the coefficient estimates to extremely large variances. The ultimate irony here is the fact that wholly unrealistic parameter estimates still generate good fits to the catch-effort history (i.e. small residuals). For example, in Figure 2 the fitted five-parameter curve predicts f_{MSY} near infinity while in the true model f_{MSY} actually corresponds to 174,000 units of effort. However different the equilibrium curves are, the five-parameter procedure still generates a good fit to the catch history ($S(\hat{\Theta}) = 1.10$). Incompleteness of information over a wide range of effort values, as well as excessive noise in the catch-effort data, will tend to bring about such pathological conditions.

To overcome these difficulties, reformulation of the estimation problem is necessary. By the following considerations, the five-dimensional parameter space can be reduced to three dimensions. First, we will approximate B_0 by Equation (20). Furthermore, if the data contain information on the yields under low exploitation, we may define B_∞ as

$$B_\infty = \text{MAX}(Y_i/q f_i) \quad i = 1, \ldots, r. \tag{21}$$

By using Equations (20) and (21), B_0 and B_∞ can be deleted from Θ, leaving only m, q, and n as the independent parameters requiring estimation. It is important to understand at this point that B_0 and B_∞ are not fixed; they are reevaluated by Equations (20) and (21) at each iteration, along with the parameters m, q, and n. In fact, the solution of Equations (1) and (3), as well as Equations (20) and (21), specify a new model with unknowns $\Theta = [m, q, n]^T$. By this restructuring, much of the degeneracy associated with the model can be eliminated. As shown in Figure 2, this procedure also provides a closer correspondence between the "estimated" and the "true" equilibrium model. Furthermore, the three-parameter procedure still generates an adequate nonequilibrium catch history ($S(\hat{\Theta}) = 1.40$). In a Monte-Carlo simulation study, parameter estimates obtained by using these transformations fell within reasonable

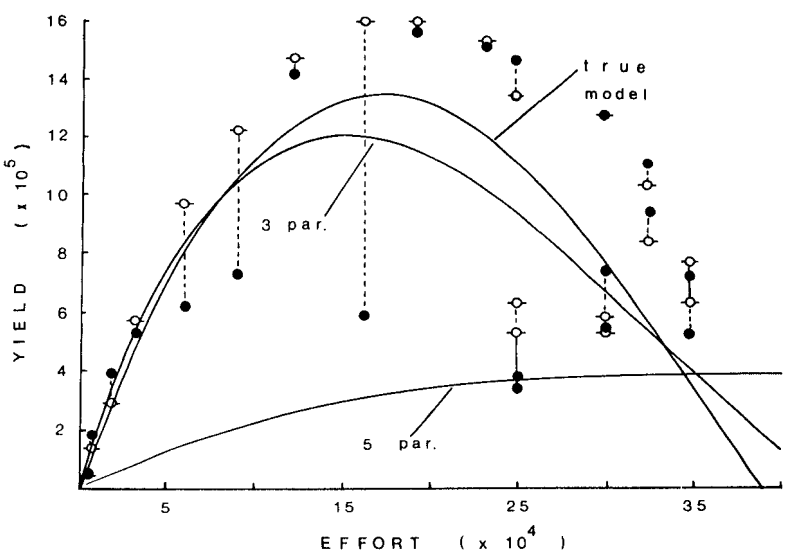


FIGURE 2.—Comparison of the "true" model with the models obtained by using the estimation procedure on three and five parameters respectively. Solid lines show equilibrium yield curves; data points show nonequilibrium simulated (dots) yields and predicted (circles) yield values from the three-parameter approach. Dashed vertical lines indicate the magnitude of residuals.

bounds (Table 3). Out of the 20 cases considered, the true parameter value lay outside the arbitrary ±2 (SD) confidence interval only once for $\hat{m}$ and $\hat{n}$. Also, variance estimates were comparable with the variance estimates of the five-parameter procedure (compare Tables 2 and 3).

TABLE 2.—Estimated parameters for the deterministic model and for 18 stochastic replicates. The Levenberg-Marquardt algorithm is employed in a five-dimensional parameter space (m, B_∞, n, q, B_0). For each parameter and replicate, the parameter estimate ± its estimated standard deviation from Equation (12) are tabulated. Replicates 13 and 15 have been excluded due to degeneracy of the model, as discussed in the text.

Replicate	σ	$\hat{\sigma}$[1]	$\hat{m}$ (10^6)	$\hat{B}_\infty$ (10^6)	$\hat{n}$	$\hat{q}$ (10^{-6})	$\hat{B}_0$ (10^6)	$\hat{f}_{MSY}$ (10^5)	$\hat{U}_{MSY}$
1	0.000	0.000	1.34±0.00	3.49±0.00	1.80±0.00	4.60±0.00	3.48±0.00	1.74	7.70
2	0.025	0.014	1.34±0.01	3.56±0.10	1.79±0.03	4.54±0.12	3.44±0.12	1.72	7.75
3	0.050	0.046	1.37±0.04	3.34±0.32	1.88±0.10	4.80±0.42	3.24±0.41	1.76	7.82
4	0.075	0.071	1.49±0.06	2.78±0.36	2.14±0.15	5.68±0.67	2.89±0.63	1.84	8.10
5	0.100	0.111	1.41±0.09	2.62±0.50	2.36±0.28	5.28±0.92	2.86±0.99	1.91	7.36
6	0.125	0.096	1.36±0.09	3.34±0.62	1.81±0.19	4.96±0.87	3.84±0.95	1.72	7.95
7	0.150	0.109	1.55±0.10	2.94±0.59	2.33±0.27	5.23±0.96	5.97±2.60	1.90	8.15
8	0.175	0.159	1.26±0.17	3.92±1.37	1.61±0.35	4.26±1.41	4.13±1.74	1.64	7.66
9	0.200	0.189	1.15±0.24	5.16±2.25	1.56±0.49	2.90±1.36	7.75±3.53	1.69	6.79
10	0.200	0.216	1.75±0.16	1.90±0.53	2.61±0.47	8.34±2.02	0.85±0.66	2.00	8.73
11	0.200	0.159	1.19±0.12	2.91±0.84	1.71±0.28	5.10±1.45	4.89±1.93	1.71	6.97
12	0.200	0.211	1.50±0.23	3.63±1.57	2.06±0.55	4.59±1.79	3.08±1.85	1.78	8.41
14	0.200	0.136	1.24±0.17	4.33±1.55	1.52±0.39	3.85±1.24	3.47±1.43	1.67	7.46
16	0.200	0.187	1.46±0.14	2.45±0.77	2.03±0.38	6.40±1.87	4.34±2.85	1.85	7.88
17	0.200	0.261	1.62±0.20	2.33±0.94	2.28±0.51	7.08±2.55	2.24±2.05	1.87	8.66
18	0.200	0.113	1.37±0.12	3.88±0.91	1.94±0.26	4.10±0.92	5.68±1.59	1.73	7.88
19	0.200	0.185	1.24±0.24	4.50±2.19	1.86±0.57	3.27±1.53	5.32±2.61	1.73	7.15
20	0.200	0.185	1.48±0.16	2.94±1.04	1.98±0.42	5.46±1.77	2.03±1.15	1.85	7.99
21	0.250	0.269	1.38±0.31	4.04±2.46	1.38±0.54	5.25±2.86	2.89±2.25	1.52	9.08
Mean[2]			1.40±0.18	3.40±1.38	1.96±0.44	5.11±1.71	3.97±2.12	1.79	7.79
SD[3]			0.20	1.06	0.33	1.74	2.04	0.10	0.69
Coeff. of var.			14%	31%	17%	34%	51%	6%	9%

[1] $\hat{\sigma} = S(\hat{\mathbf{\Theta}})/(r\text{-}5)$.
[2] For 10 replicates with $\sigma = 0.200$.
[3] Overall standard deviation of parameter estimates for 10 replicates with $\sigma = 0.200$.

TABLE 3.—Estimated parameters for 20 stochastic replicates of the deterministic model. The Levenberg-Marquardt algorithm is employed here in a three-dimensional parameter space (m, q, n). For each parameter and replicate, the parameter estimate ± its estimated standard deviation from Equation (12) are tabulated.

Replicate	σ	$\hat{\sigma}$[1]	$\hat{m}$ (10^6)	$\hat{q}$ (10^{-6})	$\hat{n}$	$\hat{f}_{MSY}$ (10^5)	$\hat{U}_{MSY}$
2	0.025	0.017	1.35±0.02	4.62±0.15	1.86±0.03	1.75	7.68
3	0.050	0.045	1.38±0.04	4.86±0.40	1.92±0.08	1.78	7.77
4	0.075	0.081	1.46±0.07	5.31±0.74	1.89±0.12	1.73	8.45
5	0.100	0.106	1.40±0.08	5.17±0.86	2.24±0.18	1.87	7.47
6	0.125	0.094	1.35±0.09	4.86±0.85	1.69±0.15	1.66	8.15
7	0.150	0.152	1.48±0.18	4.58±1.34	1.67±0.26	1.63	9.06
8	0.175	0.169	1.25±0.16	4.36±1.26	1.33±0.23	1.49	8.42
9	0.200	0.212	1.19±0.32	3.61±1.96	1.20±0.49	1.51	7.87
10	0.200	0.228	1.72±0.17	7.18±1.92	1.98±0.25	1.75	9.78
11	0.200	0.188	1.17±0.18	5.14±2.01	1.27±0.28	1.48	7.88
12	0.200	0.203	1.44±0.22	4.17±1.48	1.74±0.34	1.66	8.68
13	0.200	0.216	1.39±0.18	3.47±1.06	1.70±0.31	1.66	8.38
14	0.200	0.140	1.06±0.23	2.94±1.18	1.05±0.41	1.53	6.97
15	0.200	0.287	1.19±0.19	4.10±1.50	1.37±0.29	1.48	8.08
16	0.200	0.198	1.42±0.17	5.99±2.07	1.56±0.26	1.64	8.64
17	0.200	0.251	1.60±0.21	6.54±2.38	1.99±0.32	1.74	9.17
18	0.200	0.140	1.31±0.19	3.87±1.22	1.49±0.27	1.53	8.58
19	0.200	0.194	1.24±0.25	3.67±1.72	1.51±0.41	1.60	7.74
20	0.200	0.176	1.46±0.15	5.27±1.63	1.88±0.29	1.82	8.05
21	0.250	0.254	1.38±0.28	5.24±2.65	1.39±0.39	1.53	9.04
Mean[2]			1.35±0.21	4.66±1.72	1.56±0.33	1.62	8.32
SD[3]			0.19	1.34	0.31	0.11	0.73
Coeff. of var.			14%	29%	20%	7%	9%

[1] $\hat{\sigma} = S(\hat{\mathbf{\Theta}})/(r\text{-}3)$.
[2] For 12 replicates with $\sigma = 0.200$.
[3] Overall standard deviation of parameter estimates for 12 replicates with $\sigma = 0.200$.

CASE STUDIES

We applied the three-parameter method to the catch-effort data of the yellowfin tuna fishery of the eastern tropical Pacific, 1934 through 1967 [the same data that were analyzed by Pella and Tomlinson (1969) and by Fox (1971)]. Table 4 gives a comparison of results, and our final equilibrium model is shown by Figure 3. As indicated by Table 4, the parameter estimates of the Levenberg-Marquardt method are comparable with the estimates that Fox obtained with his search algorithm. Pella and Tomlinson also employed a searching algorithm but their minimization criterion was an unweighted least-squares function. Our standard deviation estimate is very small for $\hat{m}$(MSY) but relatively large for $\hat{B}_\infty$, $\hat{n}$, $\hat{q}$, and $\hat{B}_0$, which is a consequence of insufficient information in the yellowfin tuna data on yield at high fishing rates. With such limited information, one can anticipate that neither the shape nor the location of the descending portion of the equilibrium curve (dashed in Figure 3) could be determined with much accuracy, and the large variance estimates on the system coefficients reflect this situation. Of course, the variance estimates for $\hat{f}_{MSY}$ and $\hat{U}_{MSY}$ can always be calculated by the delta method, and to avoid the complex derivations that accompany the presence of covariance terms, an alternative would be to define a new parameter space so as to estimate f_{MSY} or U_{MSY} directly. The variance-covariance matrix for the coefficients would then provide the desired information on the variability of those parameters.

Our final example is based on the data from the Pacific halibut fishery in International Pacific Halibut Commission Area 2, as given in Ricker (1975, table 13.1). To analyze these data, Ricker derived an estimate of q from the age composition of the catch. Then he obtained parameter estimates for a Graham-Schaefer model by regressing $Y_E/\bar{B}$ against $\bar{B}$ and Y_E/f against f (Ricker 1975, examples 13.5 and 13.6). In both cases, Ricker employed GM and Nair-Bartlett regression. The results Ricker obtained by fitting the Graham-Schaefer model were compared with the results we obtained from fitting the generalized stock production model by our three-parameter version of the Levenberg-Marquardt method (Table 5). The latter provided estimates of m, q, and n with relatively small variance estimates. Furthermore, the estimate of n appears to be significantly different from 2.00, which validates the use of the Pella-Tomlinson model. Nevertheless, estimates of m are not significantly affected by the choice of the wrong model, while estimates of f_{MSY} are slightly

TABLE 4.—Comparison of parameter estimates obtained by Pella and Tomlinson (1969), by Fox (1971) and by the Levenberg-Marquardt algorithm for the yellowfin tuna in eastern Pacific Ocean. Values that follow the ± signs are the standard-deviation estimates for each parameter.

Reference	p	B_∞ (10^8)	n	q (10^{-5})	B_0 (10^8)	m (10^8)	f_{MSY}	U_{MSY}	Residuals
Pella and Tomlinson (1969, table 5)	—	—	1.40	45.0	—	1.826	35,300	5,173	1.78×10^{16}
Fox (1971, table 4)	—	1.427	2.10	8.10	1.206	1.926	32,700	5,890	0.736
Levenberg-Marquardt algorithm	—	1.448	2.08	8.01	1.192	1.924	32,700	5,884	0.735
		±0.890	±0.75	±4.9	±1.24	±0.90			
Levenberg-Marquardt algorithm (correlated error)	0.27	1.274	2.30	9.08	1.079	1.962	32,170	6,097	0.641
	±0.25	±0.653	±0.55	±4.7	±0.553	±0.106			

TABLE 5.—Comparison between the estimates of Ricker (1975) for the Pacific halibut (International Pacific Halibut Commission Area 2) and those obtained by the Levenberg-Marquardt algorithm. Values that follow the ± signs are the standard-deviation estimates for each parameter.

Reference	p	B_∞ (10^6)	n	q (10^{-7})	m (10^5)	f_{MSY}	U_{MSY}
Ricker (1975, example 13.5)							
GM regression	—	204	2.00	9.07	31.2	3.37	92.6
Nair-Bartlett regression	—	195	2.00	9.07	31.0	3.50	88.6
Ricker (1975, example 13.6)							
GM regression	—	256	2.00	9.07	33.0	2.84	116.2
Nair-Bartlett regression	—	239	2.00	9.07	31.8	2.94	108.2
Levenberg-Marquardt algorithm	—	187	1.28	14.45	31.6	2.83	111.7
		±18	±0.09	±1.36	±0.83		
Levenberg-Marquardt algorithm (correlated error)	0.33	188	1.28	14.33	31.8	2.84	112.0
	±0.16	±22	±0.12	±1.67	±1.0		

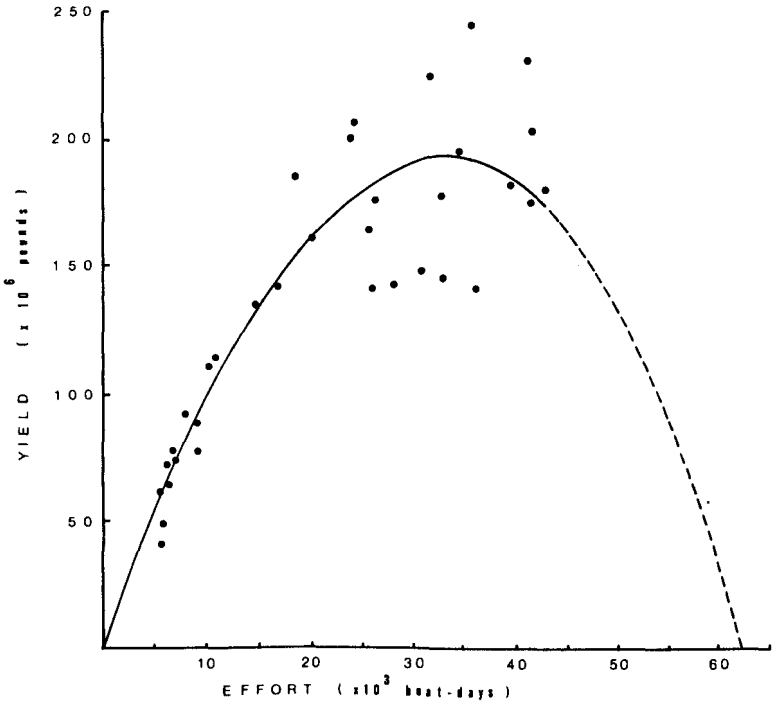
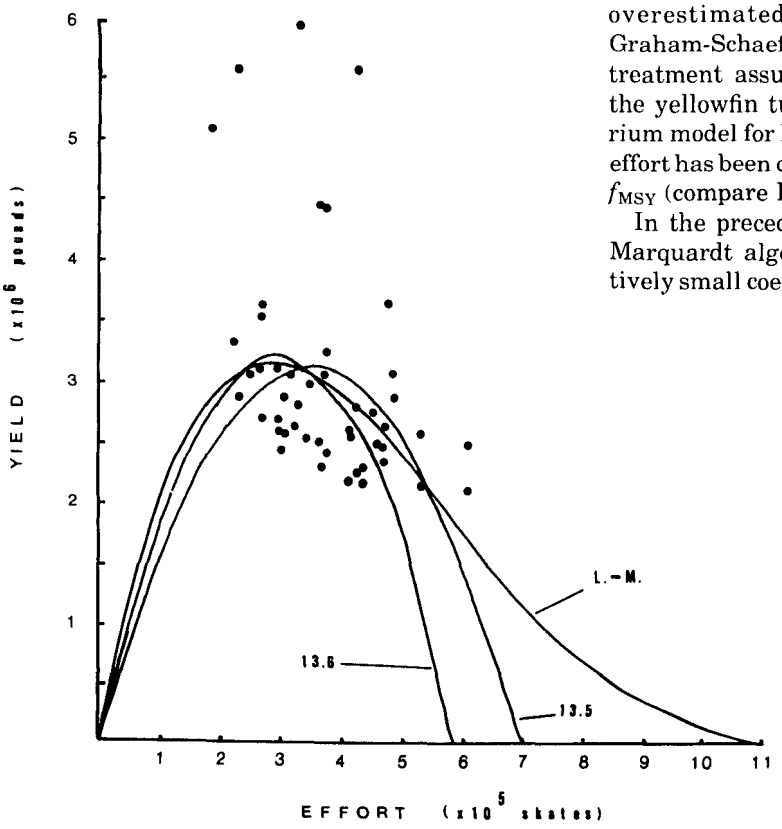


FIGURE 3.—Equilibrium stock production model for yellowfin tuna data, from 1934 through 1967, as determined by the Levenberg-Marquardt algorithm.



overestimated in most applications of the Graham-Schaefer model (note also that Ricker's treatment assumes equilibrium). In contrast to the yellowfin tuna data, analysis of the equilibrium model for halibut data indicates that fishing effort has been concentrated slightly to the right of f_{MSY} (compare Figures 3 and 4).

In the preceding case studies, the Levenberg-Marquardt algorithm gave estimates with relatively small coefficients of variation. In both cases,

FIGURE 4.—Equilibrium stock production models for Pacific halibut (International Pacific Halibit Commission Area 2), from 1910 through 1957, as determined by Ricker (1975, examples 13.5 and 13.6, Nair-Bartlett regressions) and by the Levenberg-Marquardt algorithm (three-parameter version).

e.g., the coefficients of variation for the estimates of maximum sustainable yield ($\hat{m}$) were below 6%. It is questionable, however, whether the data can justify such precision. Variability of the exploited population due to migration, to changes in fishery regulations over time, and to expansion of the fishing areas, as well as variability of q due to learning by fishermen and to technological developments, are important factors underlying the complexity of events influencing the serial catch-effort information. In future research, alternative forms of the model in which q is a variable parameterized with respect to time will be explored. Furthermore, in a randomly fluctuating environment, equilibrium population levels (and MSY, by extension) are not constant and the equilibrium points are instead described by a probabilistic cloud representing the equilibrium probability distribution (May 1974). The knowledge of this equilibrium probability distribution would give us some idea of the probability of achieving the desired management goal (MSY, for instance).

DISCUSSION ON ERROR STRUCTURE

In the preceding examples, we found runs in the time sequence plot of residuals. Those runs indicate correlations among the residuals. Serial correlation, as we usually find in applying production models to catch data, indicates that the real system is working differently than the presupposed model and that some minor effects have been neglected (such as age composition or environmental factors). But as indicated by Draper and Smith (1966), the effects of correlation can be ignored when the ratio $(r-p)/r$ tends to unity (p being the number of estimated parameters). In certain situations, of course, this ratio is likely to become small (tending to zero as r approaches p) and we may want to consider weights (W_i) which account for both the inequality of variance and the correlations. In our estimation procedure, the assumption of uncorrelated error can be relaxed by writing Equation (9) in the more general form (J. J. Pella, pers. commun.)

$$S(\hat{\Theta},\hat{\rho}) = [\mathbf{Y}-\hat{\mathbf{Y}}]\,\mathbf{W}^{-1}\,[\mathbf{Y}-\hat{\mathbf{Y}}]^{\mathrm{T}}, \qquad (22)$$

where $\mathbf{Y}$ is the row vector of observed yields, $\hat{\mathbf{Y}}$ is the row vector of predicted yields, and $\mathbf{W}$ is the symmetric, positive definite matrix

$$\mathbf{W} = \begin{bmatrix} \hat{Y}_1^2 & \hat{Y}_1\hat{Y}_2\hat{\rho} & \cdots & \hat{Y}_1\hat{Y}_r\hat{\rho}^{(r-1)} \\ \hat{Y}_1\hat{Y}_2\hat{\rho} & \hat{Y}_2^2 & \cdots & \hat{Y}_2\hat{Y}_r\hat{\rho}^{(r-2)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{Y}_1\hat{Y}_r\hat{\rho}^{(r-1)} & \hat{Y}_2\hat{Y}_r\hat{\rho}^{(r-2)} & \cdots & \hat{Y}_r^2 \end{bmatrix}.$$

Parameter ρ, constrained between 0 and 1, is a measure of the importance of lags and can be estimated along with the parameters of the differential equations (1) and (3). It can be seen that Equation (9) is a particular case of Equation (22), where the off-diagonal elements of $\mathbf{W}$ are null.

The Levenberg-Marquardt algorithm, as formulated in Equation (11), is designed to minimize directly a sum of squares of residuals as given by Equation (9). In order to minimize Equation (22) by using Equation (11), we must scale $\mathbf{W}$ by the transformation

$$\mathbf{X} = \mathbf{D}^{-1}\,\mathbf{W}\,\mathbf{D}^{-1}, \qquad (23)$$

where $\mathbf{D}$ is a diagonal matrix having elements $D_{ii} = \hat{Y}_i$ $(i = 1, \ldots, r)$, and write $\mathbf{W}$ as

$$\mathbf{W} = \mathbf{D}\,\mathbf{U}\,\Lambda\,\mathbf{U}^{\mathrm{T}}\,\mathbf{D}, \qquad (24)$$

where $\mathbf{U}\Lambda\mathbf{U}^{\mathrm{T}}$ is the eigenvalue and eigenvector decomposition of $\mathbf{X}$. Note that $\mathbf{X}$ is actually the correlation matrix of errors. Therefore Equation (22) becomes

$$S(\hat{\Theta},\hat{\rho}) = [\mathbf{Y}-\hat{\mathbf{Y}}]\,\mathbf{D}^{-1}\,\mathbf{U}\,\Lambda^{-1}\,\mathbf{U}^{\mathrm{T}}\,\mathbf{D}^{-1}\,[\mathbf{Y}-\hat{\mathbf{Y}}]^{\mathrm{T}}. \qquad (25)$$

Then Equation (25) has the same form as Equation (9), where the weights (W_i) are the square roots of the eigenvalues of $\mathbf{X}$ and where the residuals are given by $[\mathbf{Y}-\hat{\mathbf{Y}}]\,\mathbf{D}^{-1}\,\mathbf{U}$. Such a procedure requires, however, diagnoalization of an r by r matrix. Moreover, diagonalization must be repeated at least p times for each iteration. This procedure produces a 10-fold increase in computing time.

Although an exhaustive study of all possible stochastic effects on the model was not attempted, some simulations were done to determine the magnitude of error in parameter estimates due to serial correlations of the ϵ_i. Results are given on Tables 4 and 5. For the yellowfin tuna data, $\hat{\rho}$ =

0.27. For the Pacific halibut data, $\hat{\rho} = 0.33$. In either case, $\hat{\rho}$ exhibits a relatively large coefficient of variation when compared with the elements of $\hat{\boldsymbol{\Theta}}$. One could anticipate such results since ρ reflects the "persistence" of fluctuations in population size, and the estimation of ρ would therefore require a longer catch history in order to achieve a greater precision. But more importantly, the values of $\hat{\boldsymbol{\Theta}}$ and $\mathrm{Var}[\hat{\boldsymbol{\Theta}}]$ were not significantly altered by the inclusion of the additional parameter. And while the errors of any particular catch history might indeed by correlated, the minimization criterion (9) will provide satisfactory estimates of $\boldsymbol{\Theta}$ despite the fact that correlations do not enter into its formulation. The limited results contained herein suggest that serial correlation can be safely ignored when the ratio $(r - p)/r$ is near unity. Under such a condition the estimation procedure is robust with respect to the assumption of independence of errors in actual data.

CONCLUSION

The purpose of this paper has been to examine a version of the Levenberg-Marquardt algorithm as an alternative method for estimating the coefficients of the generalized stock production model. The parameter values obtained by this procedure are close to those obtained by previous studies on yellowfin tuna and Pacific halibut. Obviously, data requirements are such that a full range of effort values (ranging over low and high exploitation rates) are necessary to insure convergence in the estimation procedure and to produce estimates with small variability. Our simulations reveal that with the Levenberg-Marquardt method both the estimates of coefficients and the estimates of variances remain approximately unbiased when white noise is considered. If present, such bias is sufficiently small as to be obscured in the variability associated with catch error. The simulations also showed the range of variability in parameter estimates that might be expected for given levels of normally distributed error in catch data.

Because the parameters of interest appear explicitly in the system equations, the estimation procedure for the parameters also produces the variance estimates directly. Moreover, the method has a reliability and an efficiency of computation somewhat greater than previous methods. And since the estimation procedure relies on a numerically integrated system of differential equations, modifications of the model to incorporate such hypothetical effects as migration or stock interactions can be made easily. Of course, to the extent that the estimation procedure must rely strictly on catch-effort data, it will be subject to the same information uncertainties as any other method. But within the basic estimation procedure, we can combine the catch-effort data with prior information and thereby reduce the uncertainties in our estimates. The prior information can be any information on a state variable, such as $B(t)$, or even any prior knowledge of the coefficients as expressed by $\hat{\boldsymbol{\Theta}} \pm \mathrm{Var}(\hat{\boldsymbol{\Theta}})$. Suppose, for example, that we have information from independent surveys on stock density (acoustic surveys, indirect estimation from knowledge of larval densities, or even virtual population analysis from catch records). Such surveys would then provide us with estimates $\hat{B}_t$ each having a variance $V(\hat{B}_t)$, let us say, at various times t. We can easily introduce such information into the estimation procedure by defining the new objective function

$$S(\hat{\boldsymbol{\Theta}}) = \sum_i W_i(Y_i - \hat{Y}_i)^2 + \sum_j V_{B_j}^{-1}(B_j - \hat{B}_j)^2. \quad (26)$$

Introduction of the second term in the objective function constrains the optimization and thereby improves convergence. If the prior information has extremely large variance, then this information is of no value; the second term of Equation (26) will tend to zero and the objective function then reduces to Equation (9). In general, the alteration permits the simultaneous employment of the two state variables. Therefore, the final coefficients are no longer based solely on catch and effort data; their determination includes our knowledge of previous stock densities.

As observed here in a statistical setting, and by Fletcher (1978a, b) in the exact analysis, the Pella-Tomlinson system exhibits internal instability in its parametric relationships. That behavior arises from the variable nature of the system's nonlinearity, which would not be particularly detrimental if our problems were limited strictly to the geometric syntheses of data by curve fitting. But for the purposes of management and preservation of stocks, the subject is elevated partly at least to the status of parameter estimation "where we look for procedures to obtain values of the parameters that not only fit the data well, but also come on the average fairly close to the true value" (Bard 1974). Although the Pella-

Tomlinson system exhibits a convenient flexibility with a minimum number of coefficients, the peculiar coupling of the coefficients to the nonlinearity of the system often provides more flexibility than we care to have, and a conventional least-squares statistic may not be sufficient to control the system in the estimation procedure. In consequence, many constraints have to be imposed on the system in order to obtain convergence in the estimation procedure and to insure reliability in the coefficient values thus estimated.

ACKNOWLEDGMENTS

We would like to thank R. I. Fletcher, J. J. Pella, and C. G. Walters, each of whom reviewed the manuscript and offered many helpful suggestions. This research was supported by NORFISH, a marine research project of the University of Washington Sea Grant Office and the National Marine Fisheries Service (Grant 04-7-158-44021, Office of Sea Grant, National Oceanic and Atmospheric Administration, U.S. Department of Commerce). Financial support was also provided by the National Research Council of Canada and by the Minister of Education of Québec. Contribution No. 492 of the College of Fisheries, University of Washington.

LITERATURE CITED

BARD, Y.
1974. Nonlinear parameter estimation. Academic Press, N.Y., 341 p.

BEVINGTON, P. R.
1969. Data reduction and error analysis for the physical sciences. McGraw-Hill Book Co., N.Y., 336 p.

BROWN, K. M., AND J. E. DENNIS.
1972. Derivative free analogues of the Levenberg-Marquardt and Gauss algorithms for nonlinear least squares approximations. Numer. Math. 18:289-297.

CAUSTON, D. R.
1969. A computer program for fitting the Richards function. Biometrics 25:401-409.

CONWAY, G. R., N. R. GLASS, AND J. C. WILCOX.
1970. Fitting nonlinear models to biological data by Marquardt's algorithm. Ecology 51:503-507.

DRAPER, N. R., AND H. SMITH.
1966. Applied regression analysis. John Wiley & Sons Inc., N.Y., 407 p.

FLETCHER, R. I.
1975. A general solution for the complete Richards function. Math. Biosci. 27:349-360.
1978a. Time-dependent solutions and efficient parameters for stock-production models. Fish Bull., U.S. 76:377-388.
1978b. On the restructuring of the Pella-Tomlinson system. Fish. Bull., U.S. 76:515-521.

FOX, W. W., JR.
1971. Random variability and parameter estimation for the generalized production model. Fish. Bull., U.S. 69:569-580.
1975. Fitting the generalized stock production model by least-squares and equilibrium approximation. Fish. Bull., U.S. 73:23-37.

LEVENBERG, K.
1944. A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. 2:164-168.

MAY, R. M.
1974. Stability and complexity in model ecosystems. 2d ed. Princeton Univ. Press, Princeton, N.J., 265 p.

MARQUARDT, D. W.
1963. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11:431-441.

PELLA, J. J., AND P. K. TOMLINSON.
1969. A generalized stock production model. Inter-Am. Trop. Tuna Comm., Bull. 13:419-496.

RICKER, W. E.
1975. Computation and interpretation of biological statistics of fish populations. Fish. Res. Board Can., Bull. 191, 382 p.

SCHNUTE, J.
1977. Improved estimates from the Schaefer production model: theoretical considerations. J. Fish. Res. Board Can. 34:583-603.

WALTER, G. G.
1975. Graphical methods for estimating parameters in simple models of fisheries. J. Fish. Res. Board Can. 32:2163-2168.