Bayesian and Related Approaches to Fitting Surplus Production Models

John M. Hoenig and William G. Warren

Department of Fisheries and Oceans, Science Branch, P.O. Box 5667, St. John's, NF A1C 5X1, Canada

and Max Stocker

Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, BC V9R 5K6, Canada

Hoenig, J.M., W.G. Warren, and M. Stocker. 1994. Bayesian and related approaches to fitting surplus production models. Can. J. Fish. Aquat. Sci. 51: 1823–1831.

The Schaefer surplus production model relates equilibrium yield to fishing effort and can be fitted using just information on catch and fishing effort. Sometimes, the fitted model predicts a maximum sustainable yield (height of the parabola) that is clearly unrealistic. In this case, one may wish to use prior information on maximum sustainable yield either to constrain the height of the parabola or to provide a prior distribution for Bayesian estimation. To construct a Bayes estimator, one would generally specify a noninformative prior on the residual error variance and, possibly, on the width of the parabola; the prior distribution for height could be obtained by examining fisheries for similar stocks or species on a per unit area basis. Another possibility is to use an empirical Bayes estimator when data from several fisheries (e.g., individual lakes) are available for several years. The methodology is illustrated on catch and effort data for bigeye tuna (*Thunnus obesus*) and Dungeness crab (*Cancer magister*). The approach can be extended to other fishery models, including nonequilibrium production models. The prior distribution parameters can be allowed to depend on covariates.

Le modèle de production excédentaire de Schaefer met en relation le rendement d'équilibre et l'effort de pêche, et on peut l'ajuster en utilisant uniquement des données sur les prises et l'effort de pêche. Parfois, le modèle ajusté prévoit un rendement maximal soutenu (hauteur de la parabole) nettement irréaliste. Dans ce cas, on peut alors se servir des données a priori sur le rendement maximal soutenu pour limiter la hauteur de la parabole ou pour établir une distribution a priori servant à obtenir une estimation bayesienne. Pour créer un estimateur bayesien, il faut généralement préciser des données a priori non informatives sur la variance de l'erreur résiduelle et, probablement, sur la largeur de la parabole; on peut obtenir la distribution a priori concernant la hauteur en étudiant des types de pêche avec des stocks ou des espèces semblables par unité de surface. On peut également utiliser un estimateur bayesien empirique lorsqu'on dispose de données sur plusieurs pêches (p. ex. lacs individuels) pendant plusieurs années. On illustre la méthode au moyen de données concernant les prises et l'effort de pêche du thon ventru (*Thunnus obesus*) et du Crabe dormeur (*Cancer magister*). L'approche peut être appliquée à d'autres modèles de pêche, y compris des modèles de production non à l'équilibre. Les paramètres de la distribution a priori peuvent être conçus pour dépendre de covariables.

Received November 1, 1993
Accepted February 17, 1994
(J12144)

Reçu le 1er novembre 1993
Accepté le 17 février 1994

Despite, or perhaps because of, its simplicity, the parabolic surplus production model of Schaefer (1957) is still widely used for fisheries assessment, particularly in cases where little information exists about the fishery. The requisite data are observations on (equilibrium) catches and the corresponding fishing efforts over a number of years. Equilibrium catch in year j is assumed to be given by the dome-shaped relationship

$$(1) \quad y_j = \beta_1\chi_j - \beta_2\chi_j^2 + \epsilon_j$$

where y_j is the observed catch (at equilibrium) in year j corresponding to the fishing effort of χ_j, β_1 and β_2 are regression coefficients assumed to be positive, and ϵ_j is a random error with expectation zero. The height of the parabola is known as the maximum sustainable yield (MSY). Several other production models have been devised which also have minimal data requirements (e.g., Fox 1970; Csirke and Caddy 1983; Schnute 1989). The methods developed in this paper can be modified for these models in a straightforward manner.

It sometimes happens that the predicted height of the parabola is unreasonable. In this case, one may need to use "prior" information from similar situations to improve the estimation (see Walters 1986 for an interesting discussion of the basis for this). Information will rarely be available on the likely values of β_1 and β_2 (cf. Hoenig and Hoenig 1986). However, there may be information on the height of the parabola. In many freshwater systems, the MSY can be estimated, albeit crudely, from empirically derived models such as the morphoedaphic index (cf. Ryder 1965; FAO 1980; Hanson and Leggett 1982; and references therein). Attempts have also been made to summarize yields per unit area that can be expected for various marine systems (Marshall 1980; Garcia and LeReste 1981, p.161–163). The yields of similar stocks or species can be expressed on a per area basis and plotted against effort per unit area to obtain a "composite" stock production model. The height of the composite curve provides a rough idea of the MSY per unit area (Munro and Thompson 1973; Marten and Polovina 1982; see also Hoenig et al. 1987 for a recent review).

Information on the width of the parabola for a given stock may be problematical. In some cases, however, comparisons may be made with similar stocks in other lakes or bays.

To utilize prior information on the MSY, it is convenient to reparameterize equation (1) in terms of the height and width of the parabola. Let h be the height and w the half width at $y = 0$; thus, w is the fishing effort producing the MSY, often referred to as f_{msy}. Then, equation (1) can be written as

$$(2) \qquad y_j = \frac{2h}{w} x_j - \frac{h}{w^2} x_j^2 + \epsilon_j .$$

In this paper, we begin by fitting the production model by fixing the height. This leads to consideration of a Bayes estimator. Empirical Bayes estimators are then described for multiple stocks by way of the computational procedures of Dempster et al. (1981) which make use of the expectation–maximization (EM) algorithm. Finally, the methodology is illustrated with the Schaefer model applied to a bigeye tuna (*Thunnus obesus*) stock (Bayes estimates) and to multiple stocks of Dungeness crab (*Cancer magister*) (empirical Bayes estimates).

Bayesian estimation has not been commonly used in fisheries work, there being perhaps two dozen examples in the literature. The underlying philosophy of Bayesian estimation differs fundamentally from that of classical inference and there exists a substantial degree of controversy over the meaning of the differences. The interested reader is referred to Press (1989) for an introduction to Bayesian inference and to Barnett (1982) for a thoughtful comparison of the approaches. The use of Bayesian statistics in a fisheries context is discussed in Walters (1986) and Walters and Ludwig (1994).

Methodology

Fishing Effort (f_{msy}) for Given MSY

Suppose that the MSY, h, is known. The least-squares estimate of the corresponding fishing effort, $\hat{w}$, is then obtained as the solution of the cubic equation

$$(3) \qquad \hat{w}^3 \sum_{j=1}^{n} y_j x_j - \hat{w}^2 \sum_{j=1}^{n} x_j^2 (y_j + 2h) + 3\hat{w}h \sum_{j=1}^{n} x_j^3 - h \sum_{j=1}^{n} x_j^4 = 0$$

where n is the number of years of data. This equation can be written as

$$a_1 \hat{w}^3 + 3a_2 \hat{w}^2 + 3a_3 \hat{w} + a_4 = 0.$$

Let $H = a_1a_3 - a_2^2$, $G = a_1^2a_4 - 3a_1a_2a_3 + 2a_2^3$, and $\Delta = G^2 + 4H^3$. If $\Delta > 0$, there exist one real and two complex roots. In this case, Cardan's solution of the cubic is convenient; specifically

$$(4) \qquad \hat{w} = \frac{Z - a_2}{a_1}$$

where

$$Z = Q - \frac{H}{Q} \text{ and } Q = \left(\frac{G - \sqrt{\Delta}}{2} \right)^{1/3} .$$

Let S^2 denote the residual sum of squares obtained by estimating β_1 and β_2 and, thus, h and w, by ordinary, unconstrained least squares, and let S_h^2 denote the residual sum of squares obtained by assuming h known. S_h^2 cannot, of course, be less than S^2. We may readily determine the range of values for h, $[h_1, h_2]$ say, for which $S_h^2/S^2 \le 1 + \delta$, where δ is some predefined (small) value, for example 0.1. For each chosen value of h, there is, by equation (4), a corresponding value of $\hat{w}$. We may thus demarcate a set of sustainable yields and corresponding fishing efforts for which the residual sum of squares is inflated over that obtained by the unconstrained least squares fit by at most $100\delta\%$. We may then ask whether such MSYs, which, in this sense, are consistent with the data, are believable. Conversely, we may judge whether MSYs that we deem feasible are consistent with the data.

Extending this idea, we might place a prior distribution on the MSY. Through equation (4), this will impose a distribution on the corresponding fishing effort. If the relationship between h and w turns out to be essentially linear, then the distribution imposed on w will have essentially the same form as that assumed for h (see the example below).

Bayesian Estimation

The idea of placing a prior distribution on h leads then to consideration of Bayesian estimation per se.

Assume in equation (2) that the ϵ_j are independent normally distributed random variables with mean zero and variance σ^2. The likelihood of the data is then

$$(5) \qquad \Lambda = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^{n} \left(y_j - \frac{2h}{w} x_j + \frac{h}{w^2} x_j^2 \right)^2 \right).$$

As noted above, we may well have sufficient information to place a prior distribution on h but feel unable to place prior distributions on w and σ^2. Bayesian estimation, however, requires priors on all unknown parameters. So-called noninformative priors can be placed on w and σ^2. According to Seber and Wild (1989), the criteria followed by Box and Tiao (1973) lead to the Jeffreys' noninformative prior which here, under the assumption that w and σ^2 are independent, is given by the product of $|I_1(\sigma)|^{1/2}$ and $|I_2(w)|^{1/2}$ where, with $L = \log(\Lambda)$,

$$I_1(\sigma) = E\left[-\frac{\partial^2 L}{\partial \sigma^2} \right]$$

$$I_2(w) = E\left[\frac{\partial^2 L}{\partial w^2} \right]$$

and, in each, the other parameters are treated as known. Thus

$$I_1(\sigma) = \frac{2n}{\sigma^2} \propto \sigma^{-2}$$

$$I_2(w) = \frac{4h^2}{\sigma^2 w^6} \sum_{j=1}^{n} x_j^2 (x_j - w)^2 \propto \frac{\sum x_j^2 (x_j - w)^2}{w^6}.$$

Note that if, as seems reasonable in this and many situations, the fishing gear is not saturated even under the highest catches, h may be assumed to be independent of w. Then, following Seber and Wild (1989) the posterior probability density of h, w, and σ, $P(h,w,\sigma|\mathbf{y})$, is proportional to

$$\frac{1}{\sigma^{n+1}}\exp\left(-\sum_{j=1}^{n}(y_j-\hat{y}_j)^2/2\sigma^2\right)\times\frac{\sqrt{\sum_{j=1}^{n}x_j^2\,(x_j-w)^2}}{w^3}\,p(h)$$

where $\hat{y}_j=\frac{2h}{w}\chi_j-\frac{h}{w^2}\chi_j^2$ and $p(h)$ denotes the prior probability density of h chosen by the analyst. The above can be integrated over σ to eliminate σ, whence

$$(6)\qquad P(h,w\,|\,\mathbf{y})\propto\frac{1}{w^2}\left[\frac{\sum x_j^2\,(1-x_j/w)^2}{\left[\sum(y_j-2hx_j/w+h^2x_j^2/w^2)\right]^n}\right]^{1/2}p(h)=\xi(h,w),\text{ say.}$$

Next, ξ(h,w) can be integrated numerically over h and w to determine the constant necessary to obtain a proper posterior probability density. Let $1/\kappa=\int_0^\infty\int_0^\infty\xi(h,w)dh\,dw$; then the posterior joint density for h and w is

$$(7)\qquad P(h,w\,|\,\mathbf{y})=\kappa\xi(h,w).$$

The marginal posterior density for h can be obtained by integrating equation (7) over w; thus

$$P(h\,|\,\mathbf{y})=\int_0^\infty P(h,w\,|\,\mathbf{y})\,dw$$

and the mean of the posterior marginal density is a Bayes estimator of the maximum sustainable yield, i.e.,

$$\hat{h}_B=\int_0^\infty hP(h\,|\,\mathbf{y})\,dh.$$

The 100(1 − α)% credibility interval (the Bayesian analog of the confidence interval) is (h_1,h_2) where

$$\int_0^{h_1}P(h\,|\,\mathbf{y})\,dh=\alpha/2$$
$$\int_{h_2}^{\infty}P(h\,|\,\mathbf{y})\,dh=\alpha/2.$$

The Bayes estimate for the corresponding fishing effort, w, and its credibility interval are obtained from equation (7) in a parallel manner.

Credibility intervals describe the (posterior) degree of belief of the parameter falling in the interval, given the data (Press 1989). Bayesian statisticians find this a natural interpretation for an interval estimate. They point out that, in contrast, the traditional confidence interval has an awkward, frequentist interpretation of being a random interval which, if constructed many times, will include the true value of the parameter a specified percentage of the times.

If it should happen that substantial information exists concerning the fishing effort, w, that will result in catches equal to the MSY (i.e., h), then one may wish to specify an informative prior distribution for w. The Bayesian estimation procedure described above is easily modified to accommodate an informative prior on w, say $q(w)$. The posterior probability density of h, w, and σ, $P(h,w,\sigma|\mathbf{y})$ is then proportional to

$$\frac{1}{\sigma^{n+1}}\exp\left(-\sum_{j=1}^{n}(y_j-\hat{y}_j)^2/2\sigma^2)\,p(h)q(w)\right)$$

where all symbols are as previously defined. This expression can again be integrated over σ to eliminate σ. The result is proportional to the joint posterior density of h and w and can be converted to a proper probability density by multiplying by a constant determined as for equation (6). The determination of marginal posterior densities, Bayes estimates, and credibility intervals proceeds as for the case when the prior distribution for w was noninformative.

Empirical Bayes Estimation

Empirical Bayes methods are useful when one has several series of observations on similar situations or similar systems. For example, there may be observations on a stock in each of several similar lakes. If information is strong for a particular stock, the empirical Bayes estimates will be little affected by consideration of information from other stocks. However, if there is little information for a particular stock, consideration of information from other stocks will cause the estimates for that stock to be pulled towards the mean estimates of all stocks. If information is very weak, these "shrunken" estimates may well be better estimates for the particular stock than the estimates based on the information from that stock only.

Let y_{ij} and x_{ij} be the catch and effort, respectively, for fishery (e.g., lake) i in year j for i = 1, 2, ..., I and j = 1, 2, ..., n_i. The analogous model to equation (1) is

$$y_{ij}=\beta_{1i}x_{ij}-\beta_{2i}x_{ij}^2+\epsilon_{ij}$$

where random error terms, ϵ_{ij}, are, again, independent normally distributed random variables with mean zero and variance σ^2.

Suppose, now, that production parameters for the various stocks arise from a bivariate normal distribution with mean (μ_1,μ_2) and covariance matrix $\mathbf{\Sigma}$. The model can be rewritten

$$(8)\qquad y_{ij}=\mu_1x_{ij}+(\beta_{1i}-\mu_1)x_{ij}-\mu_2x_{ij}^2-(\beta_{2i}-\mu_2)x_{ij}^2+\epsilon_{ij}.$$

Here, μ_1 and μ_2 are regarded as fixed effects, i.e., intrinsic to the species, while $\beta_{1i}-\mu_1$ and $\beta_{2i}-\mu_2$ are regarded as random effects, i.e., varying over the different stocks of the species, with bivariate $N(\mathbf{0},\mathbf{\Sigma})$ distribution.

Dempster et al. (1981) showed that Bayesian estimation of the β_{1i} and β_{2i} is straightforward when σ^2 and $\mathbf{\Sigma}$ are known. Indeed, estimates of the posterior distribution of β_{1i} and β_{2i} can be obtained from ordinary least squares algorithms. In the present case, σ^2 and $\mathbf{\Sigma}$ are not known.

The EM algorithm is a useful procedure for finding maximum likelihood estimates when the data are incomplete. In this case, the data are incomplete in the sense that the sufficient statistics for σ^2 and $\mathbf{\Sigma}$ involve not only the observations, y_{ij}, but also the unknown regression coefficients, β_{1i} and β_{2i}. The EM algorithm proceeds as follows. Begin with an initial guess of σ^2 and $\mathbf{\Sigma}$. Given these, compute the expected value of the sufficient statistics for σ^2 and $\mathbf{\Sigma}$ (the expectation step) — this involves computing Bayes estimates of the regression coefficients β_{1i} and β_{2i} conditional on the [current] values of σ^2 and $\mathbf{\Sigma}$. Given these estimates of the sufficient statistics, it is an easy matter to compute maximum likelihood

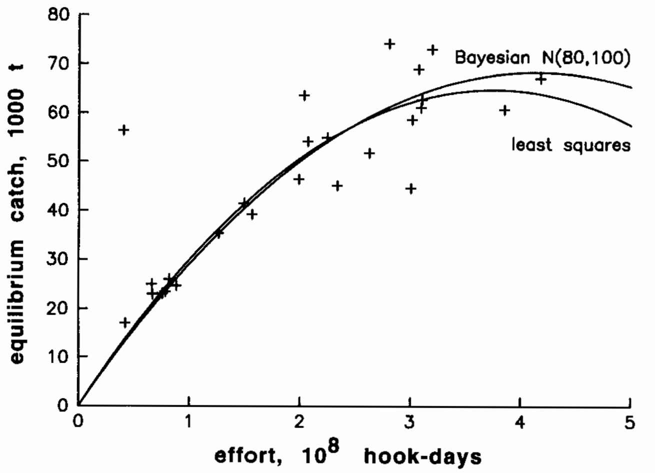

FIG. 1. Scatter plot of catch versus fishing effort for bigeye tuna from the Atlantic Ocean (from Miyabe 1989). Also shown are the Schaefer model (parabola) fitted by ordinary least squares and by the Bayesian method with an $N(80,100)$ prior of height (h) and noninformative priors on half-width (w) and the residual variance (σ^2).

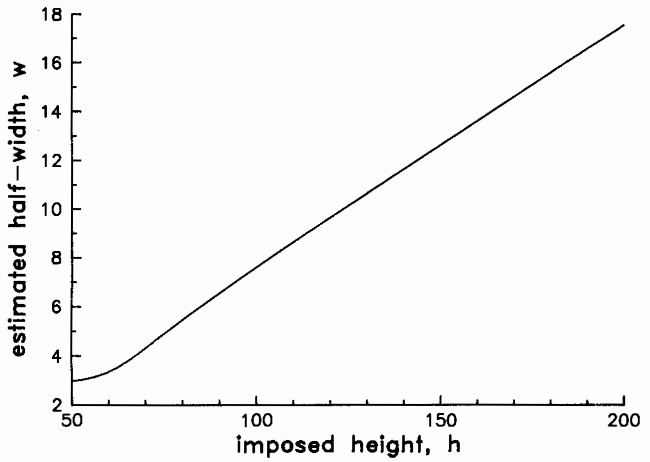

FIG. 2. Relationship between the imposed height (h) of the parabola fitted to the bigeye tuna data and the resulting estimated half-width (w). Note that the relationship is nearly linear for imposed heights greater than approximately 65 t.

estimates of σ^2 and Σ (the maximization step). The new values of σ^2 and Σ can then be used to begin the next cycle of iteration, i.e., to update estimates of the expected values of the sufficient statistics so that new maximum likelihood estimates of σ^2 and Σ can be obtained, and so forth until convergence is achieved. Note that the empirical Bayes estimates of the regression coefficients are obtained automatically as a by-product of the algorithm. Computational details are given in the Appendix.

Examples

Optimal Fishing Effort for Given MSY

Catch and effort data for bigeye tuna in the Atlantic Ocean have been tabulated by Miyabe (1989). The observations fall on the ascending (left) limb of the parabola (Fig. 1).

The ordinary least squares (unconstrained) estimates of the parameters are MSY, $h = 64.7 \times 10^3$ t, and corresponding effort, $w = 3.75 \times 10^8$ hook-days, with a residual sum of

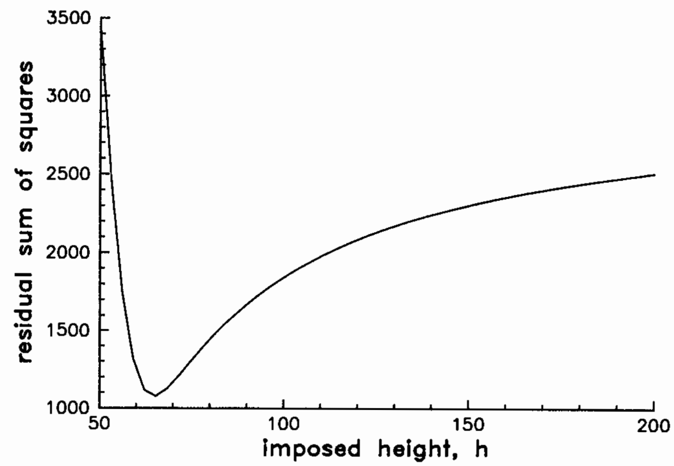

FIG. 3. Relationship between the imposed height (h) of the parabola fitted to bigeye tuna data and the resulting residual sum of squares.

squares of 1075.9 (Fig. 1). Computation of constrained estimates, i.e., values of the effort for selected values of the MSY (from equation (4)), shows that w is a nearly linear function of the imposed height, h, for h greater than approximately 65×10^3 t (Fig. 2). Accordingly, it is relatively easy to translate one's feelings about the probable value of MSY into feelings about the associated effort. In particular, one could place a probability distribution on h and determine the consequent distribution on w.

Suppose now that we wish to consider estimates that inflate the (unconstrained) sum of squares by at most 10%, i.e., find constrained estimates so that the residual sum of squares ≤ 1183.5. This is achieved with any h in the range $(60.5 \times 10^3, 70.0 \times 10^3)$ (Fig. 3). By equation (4), the corresponding efforts fall in the range $(3.40 \times 10^8, 4.65 \times 10^8)$.

Bayesian Estimation

The same data are used to illustrate Bayesian estimation per se. Noninformative (Jeffreys') priors are assumed for w and σ^2 whereas, for illustrative purposes, a normal prior, $N(80,100)$, is assumed for h (note that the prior distribution should be specified without reference to the data). From equation (7) the Bayes estimate of the MSY is then 68.3×10^3 t with a 95% credibility interval of $(62.2 \times 10^3, 77.4 \times 10^3)$ t. The Bayes estimate of the needed effort is 4.14×10^8 hook-days with a 95% credibility interval of $(3.42 \times 10^8, 5.21 \times 10^8)$ hook-days. Observe that the Bayes estimate of MSY is moved from the unconstrained least squares estimate towards the mean of the prior distribution. The stronger the degree of prior belief, as reflected by a smaller variance for the prior distribution, the greater the shift towards the prior mean.

Empirical Bayes Estimates

Computation of empirical Bayes estimates is illustrated with catch and effort data from 12 Dungeness crab fisheries around British Columbia, Canada. A previous analysis of these data has been made by Stocker and Butler (1990). Observations are available for periods ranging from 27 to 43 yr. Following Stocker and Butler, it is logical to consider surplus production models for these fisheries. Equilibrium models should be appropriate for this species for the following reasons: maturity is attained at age 2–3 yr

TABLE 1. Areas inhabited by 12 stocks of Dungeness crab in British Columbia, Canada, and estimates of the production parameters expressed on an areal basis. LS = least squares; EB = empirical Bayes; EBH = empirical Bayes without data from Hecate Strait.

Stock	Area (km^2)	Method	$\hat{\beta}_1$	$\hat{\beta}_2(\times 10^{-2})$	$\hat{w}$ (d/km^2)	$\hat{h}$ (t/km^2)
Boundary Bay	45	LS	0.1463	0.0792	92.3	6.75
		EB	0.1463	0.0794	92.1	6.74
		EBH	0.1457	0.0784	92.9	6.77
Burrard Inlet	39	LS	0.0856	0.0738	58.0	2.48
		EB	0.0857	0.0740	57.9	2.48
		EBH	0.0861	0.0747	57.6	2.48
Chatham Sound	82	LS	0.2141	0.7665	14.0	1.49
		EB	0.1692	0.2749	30.8	2.60
		EBH	0.1640	0.2723	30.1	2.47
English Bay	52	LS	0.0496	−0.5160	—	—
		EB	0.0844	−0.2551	—	—
		EBH	0.0846	−0.2496	—	—
Fraser River Mouth	107	LS	0.1928	0.4065	23.7	2.29
		EB	0.1844	0.3590	25.7	2.27
		EBH	0.1807	0.3435	26.3	2.38
Gulf Islands	62	LS	0.0512	0.0704	36.4	0.93
		EB	0.0475	0.0480	49.4	1.17
		EBH	0.0529	0.0717	36.9	0.98
Hecate Strait	645	LS	1.8958	0.2289	4.1	3.92
		EB	1.6256	2.0132	40.4	32.8
Nanaimo	52	LS	0.0479	0.0527	45.4	1.09
		EB	0.0464	0.0439	52.9	1.23
		EBH	0.0510	0.0614	41.6	1.06
Queen Charlotte Sound	168	LS	0.0851	0.3926	10.8	0.46
		EB	0.0612	0.0572	53.5	1.64
		EBH	0.0647	0.0478	67.6	2.19
Sidney–Esquimalt	78	LS	0.0618	0.0829	37.3	1.15
		EB	0.0609	0.0790	38.6	1.17
		EBH	0.0640	0.0888	36.1	1.15
Sooke Harbour	91	LS	0.0506	0.0314	80.5	2.04
		EB	0.0508	0.0319	79.6	2.02
		EBH	0.0542	0.0412	65.8	1.78
Tofino	104	LS	0.1789	0.2234	40.0	3.58
		EB	0.1786	0.2224	40.2	3.59
		EBH	0.1750	0.2159	40.7	3.58

and males recruit to the fishery at 4–5 yr so that long lags would not be expected in the population's response to changing fishing effort, and the fishery is mainly seasonal so that the population has some time between seasons to respond to such changes.

These crab data have been chosen for the purpose of illustrating the application of empirical Bayes methods to estimate the parabolic form of the Schaefer model (equation (1)). Stocker and Butler (1990) considered several possible error structures and selected a different form of the production model than we used. Consequently, their estimates differ somewhat from those obtained here. The interested reader is referred to their paper for a careful discussion of these data.

Since the production available from a stock is determined, in part, by the size of the stock, it is reasonable to standardize the data by expressing catches and efforts on a per area basis. Accordingly, the catches and efforts have been divided by the fishing area, i.e., the area occupied by, and fished for, each stock (Table 1; T. Butler, 2630 Lynburn Crescent, Nanaimo, BC V9S 3T6, personal communication). This has no effect on the relative form of the fitted models but rescales the data so that the assumption that the parameters, $\beta_{1i} - \mu_1$, $\beta_{2i} - \mu_2$, can be regarded as random variables from a common distribution should be better approximated.

Scatter plots (Fig. 4) reveal much variability in the catches for a given effort, as is typical of production data. Schaefer models (equation (1)) fitted by ordinary least squares appear to give reasonable fits in some cases but unreasonable fits in others (Fig. 4; Table 1). In particular, the fitted parabola for the English Bay data is cup- rather than dome-shaped. This is inconsistent with the assumptions of the Schaefer model. We have elected, however, not to reject the English Bay data in order to see whether the empirical Bayes estimate, which incorporates information from the other stocks, will be feasible. It is noted that Hecate Strait is atypical in that the second highest MSY per square kilometre is attained with the smallest (standardized) effort.

With the model taken as in equation (8), empirical Bayes estimates were obtained by following Appendix equation (A.1) through (A.7) for all 12 stocks. The procedure converged to $\hat{\sigma}^2 = 0.1935$,

$$\hat{\Sigma} = \begin{bmatrix} 0.1980 & -0.0025 \\ -0.0025 & 0.000034 \end{bmatrix},$$

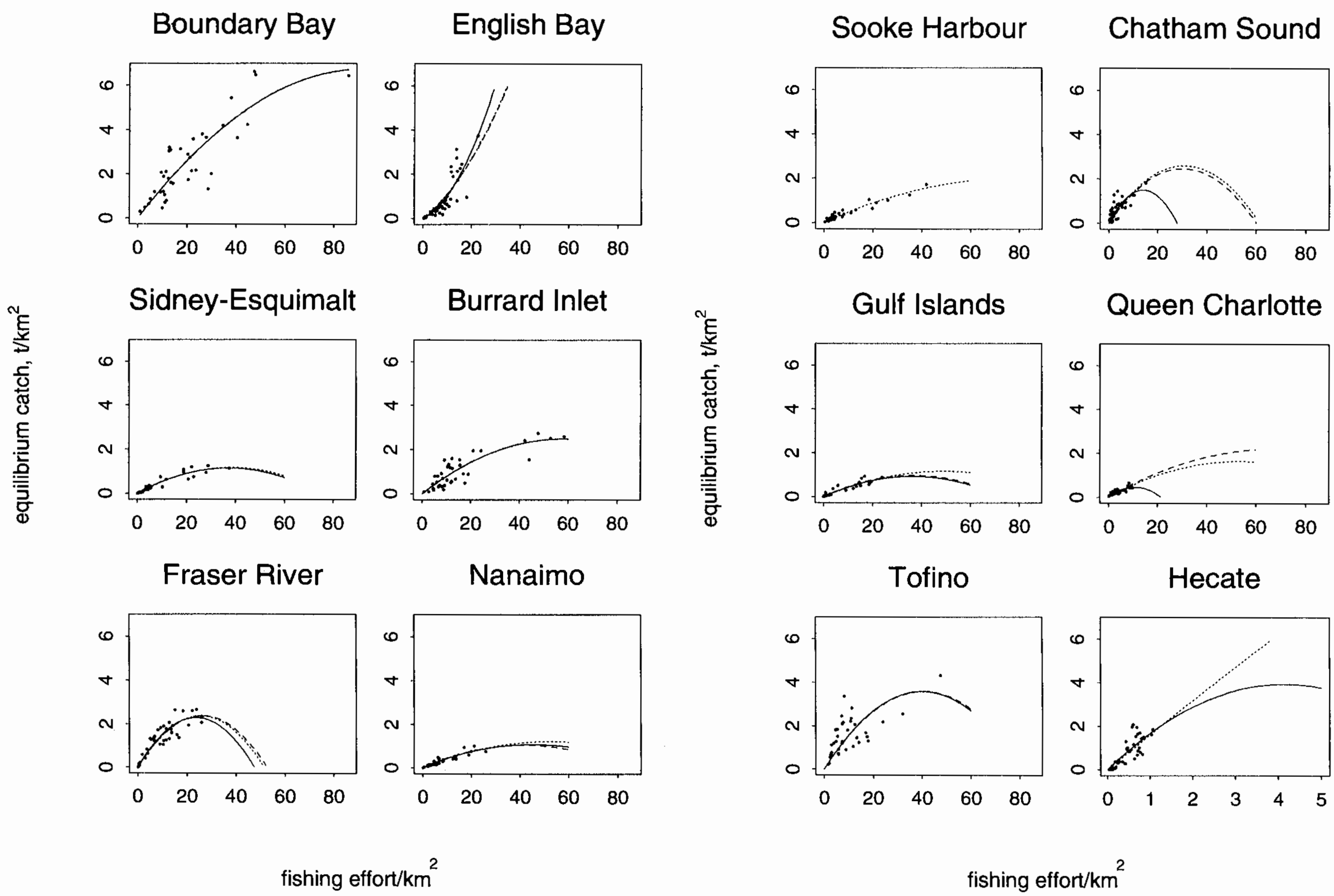

FIG. 4. Scatter plots of catch versus fishing effort for 12 stocks of Dungeness crab from British Columbia, Canada. Also shown are the least squares (solid line) estimates of the Schaefer production model and empirical Bayes estimates with (dotted line) and without (dashed line) the data from Hecate Strait. Note that the scale for the Hecate Strait data differs from the other scales.

TABLE 2. Range of standardized fishing efforts (d/km^2).

Stock	Minimum	Maximum
Boundary Bay	1.0	86.0
Burrard Inlet	0.9	58.3
Chatham Sound	0.3	15.1
English Bay	0.4	22.7
Fraser River Mouth	0.1	25.9
Gulf Islands	<0.1	23.8
Hecate Strait	<0.1	1.1
Nanaimo	0.2	26.0
Queen Charlotte Sound	0.2	9.8
Sidney–Esquimalt	0.1	37.4
Sooke Harbour	0.4	41.9
Tofino	1.6	47.6

and $\hat{\boldsymbol{\mu}}' = [0.228, 0.00252]$. The estimates of the production parameters are given in Table 1.

For some stocks the differences between the empirical Bayes estimates and the least squares estimates are virtually imperceptible, for example, Boundary Bay, Burrard Inlet, Sooke Harbour, and Tofino. On the other hand, there are stocks, namely Chatham Sound, Queen Charlotte Sound, and Hecate Strait, where the differences are substantial. The similarity, or lack thereof, between the empirical Bayes and least squares estimates is related to the range of fishing efforts (Table 2). It would appear that, for Chatham Sound and Queen Charlotte Sound, the fishing effort employed does not approach that corresponding to the MSY; forcing the Schaefer model on these data leads to seemingly unrealistic least squares estimates whereas the empirical Bayes estimates are much more in accordance with what would be expected in relation to the production of other stocks. The standardized efforts (effort per unit area) for Hecate Strait are so small as to contain virtually no information on the catch–effort relationship. Accordingly, the empirical Bayes procedure takes as estimate of w a value close to the average of the other stocks. Because the catches are relatively high for these small efforts, the empirical Bayes estimate of h is then, seemingly, unrealistically high.

Although moved in the right direction, the catch–effort curve for English Bay remains cup-shaped. The cup-shaped curve for English Bay arises from the English Bay data appearing to fall into two groups. This was noticed by Stocker and Butler (1990) who fitted separate lines to what they judged to be separate phases of the fishery, namely (i) 1963–72 and (ii) the remaining years 1951–62 and 1973–88.

It has been assumed that the $\beta_{1i} - \mu_1$ and $\beta_{2i} - \mu_2$ are bivariate normally distributed. However, one does not need any formal test to conclude that it is inconceivable that the empirical Bayes estimates (Table 3) could arise from a normal distribution. Hecate Strait is clearly an outlier. This

TABLE 3. Ranked empirical bayes estimates of regression parameters.

Stock	$\hat{\beta}_{1i} - \hat{\mu}_1$	Stock	$\hat{\beta}_{2i} - \hat{\mu}_2$ ($\times 10^{-2}$)
Nanaimo	−0.182	English Bay	−0.507
Gulf Islands	−0.181	Sooke Harbour	−0.220
Sooke Harbour	−0.178	Nanaimo	−0.208
Sidney–Esquimalt	−0.168	Gulf Islands	−0.204
Queen Charlotte Sound	−0.167	Queen Charlotte Sound	−0.195
English Bay	−0.144	Burrard Inlet	−0.178
Burrard Inlet	−0.143	Sidney–Esquimalt	−0.173
Boundary Bay	−0.082	Boundary Bay	−0.173
Chatham Sound	−0.059	Tofino	−0.030
Tofino	−0.050	Chatham Sound	0.023
Fraser River Mouth	−0.044	Fraser River Mouth	0.107
Hecate Strait	1.397	Hecate Strait	1.760

raises the question of how much influence the Hecate Strait data may have had on the empirical Bayes estimates of the other stocks. Accordingly, empirical Bayes estimates have been computed with the Hecate Strait data excluded (Table 1, with $\hat{\mu}' = [0.102, -0.00311]$). The only notable changes in the estimates occur with Queen Charlotte Sound, Sooke, and, in particular, Gulf Islands and Nanaimo, with the last two moving towards the least squares estimates; in all cases, difference in the fits over the range of the data is virtually imperceptible. For completeness, the estimates of $\beta_{1i} - \mu_1$ and $\beta_{2i} - \mu_2$ are given in Table 4.

TABLE 4. Empirical bayes estimates of regression parameters (Hecate Strait omitted).

Stock	$\hat{\beta}_{1i} - \hat{\mu}_1$	$\hat{\beta}_{2i} - \hat{\mu}_2$ ($\times 10^{-2}$)
Boundary Bay	0.044	−0.017
Burrard Inlet	−0.016	−0.020
Chatham Sound	0.062	0.177
English Bay	−0.018	−0.345
Fraser River Mouth	0.079	0.248
Gulf Islands	−0.049	−0.023
Nanaimo	−0.051	−0.034
Queen Charlotte Sound	−0.038	−0.047
Sidney–Esquimalt	−0.038	−0.006
Sooke Harbour	−0.048	−0.054
Tofino	0.074	0.121

Discussion

A suite of methods for fitting surplus production models has been presented. These methods vary in their data requirements and their ease of computation. The estimator based on fixing the height (equation (4)) is a convenient tool when stock assessment must be done quickly, such as during meetings. The convenience lies in the availability of an analytic solution and in the near linearity of the relationship between the assumed height (or MSY) and the estimated width (or corresponding effort). The other methods require considerably more computation. The Bayes estimates are useful when limited prior information is available to aid in parameter estimation. However, when observations are available on several similar stocks, empirical Bayes estimates become attractive.

A classical alternative, used for example by Polovina (1989), is to fit production models simultaneously to a set of stocks under the assumption that at least one of the parameters is common to all stocks. A special case of this is Munro's composite production model (Munro and Thompson 1973) in which a single parabola is fitted to all of the data after normalizing catches and efforts to a per area basis. This involves fitting fewer parameters than the empirical Bayes approach and may be useful when only a few years of data are available. However, it is not nearly as flexible. One drawback with empirical Bayes estimators is that, as noted by Press (1989), because there is no natural standard error, it is difficult to find credibility intervals or to test hypotheses.

It is well known that, when the necessary assumptions are met, empirical Bayes estimates will be as good as, or better than, the separate least squares estimates for each system. Indeed, Press (1989) observed that "Often, the risk of the empirical Bayes estimator is less than half that of the MLE [maximum likelihood estimator]". Of course, the ordinary least squares estimates are "best" under a simple least squares criterion. In the surplus production example the differences between least squares and empirical Bayes estimates occur primarily outside the range of the data (i.e., in the region where there is little or no direct information). The residual sums of squares from the empirical Bayes estimates are only marginally inflated over those from the least squares estimates. In those cases where least squares estimates of the height and width are poorly defined, they are made more compatible with the estimates for the other stocks; otherwise, they are little, if at all, changed.

The assumptions behind empirical Bayes estimation of the stock production model that might fail are as follows: (i) the ϵ_{ij} may not be normally distributed, their variance may vary from stock to stock and with effort within a stock, and they may not be mutually independent among years within a stock, (ii) the regression parameters may not be independent random observations from a normal prior distribution, (iii) equilibrium production may be an asymmetric function of effort rather than the parabolic function used, (iv) production may not be at equilibrium in all years, and (v) catchability may be changing (increasing) over time.

The crab example suggests that the results are reasonably robust against violation of (ii) and this assumption can be assessed by looking at the distribution of the regression parameter estimates. Likewise the estimates, although not necessarily their variance, should be reasonably robust against violations of (i). Violation of (iii) would not appear to be a problem if the bulk of the data fall on, and fully cover, the ascending arm of the curve and the second-degree equation is a reasonable approximation for that arm. The English Bay data suggest that violation of (iv) or (v) can, however, be a problem. Nonequilibrium models (not considered here) may be useful for situations where the effort has changed rapidly over time. In rapidly developing fisheries, MSY tends to be overestimated if equilibrium models are used. If the catchability increases over time, then MSY will also tend to be overestimated; this is difficult to handle.

Although we have chosen to illustrate Bayesian approaches to surplus production modelling using the simple equilibrium Schaefer model, the modifications required to handle other production models are in many cases minor. Consider, for example, the nonequilibrium model described by Schnute (1989), namely

$$y_j = (1-\delta)\frac{2h}{w}x_j - (1-\delta)\frac{h}{w^2}x_j^2 + \frac{\delta y_{j-1}x_j}{x_{j-1}} + \epsilon_j$$

where $|\delta| < 1$ and the other symbols are as defined above. This model describes the catch in year j as a weighted mean of the equilibrium production in year j and a residual effect dependent on the catch in the previous year ($j - 1$). It is thus a first-order autoregressive model for which Schnute (1987) shows how to construct the likelihood.

To cast this model in a Bayesian framework, one simply multiplies the likelihood for the data by the prior distributions for each parameter (h, w, δ, σ) to obtain a function which is proportional to the joint posterior probability density function. This function is converted to a proper density function by multiplying by a suitable constant (determined by numerical integration). The posterior density is then treated in the same manner as above.

The methods considered in this paper can also be extended to allow the prior distribution of the parameters to depend on covariates. Strenio et al. (1983) described a generalization of the procedures of Dempster et al. (1981) to allow the expected value (μ_1,μ_2) of the production parameters to be linear functions of covariates. Thus, for example, the production parameters might depend on water depth, water temperature, area encompassed by the stock, etc. It seems that lack of data, rather than lack of suitable methodology, is the impediment to future analysis.

Finally, it is noted that the same approach used here for stock production models can be used for other fishery models. In particular, the fit of individual stock–recruitment models is generally quite poor but there may be data available for several stocks. For example, instead of fitting a separate Ricker model to the salmon stock in each stream, one could use an empirical Bayes procedure to fit models to all the stocks simultaneously.

Acknowledgements

We thank Terry Butler for providing the Dungeness crab fishery data and the areas inhabited by the crab stocks and for his comments. Jon Schnute, Victor Restrepo, and Nadine Hoenig brought to our attention various aspects of the problem, and Russell Millar and Carl Walters provided critical reviews of the manuscript. Nick Payton and Noel Cadigan constructed the figures.

References

BARNETT, V. 1982. Comparative statistical inference. 2nd ed. John Wiley & Sons, New York, N.Y. 325 p.

BOX, G.E.P., AND G.C. TIAO. 1973. Bayesian inference in statistical analysis. Addison-Wesley, Reading, Mass. 588 p.

CSIRKE, J., AND J.F. CADDY. 1983. Production modeling using mortality estimates. Can. J. Fish. Aquat. Sci. 40: 43–51.

DEMPSTER, A.P., D.B. RUBIN, AND R.K. TSUTAKAWA. 1981. Estimation in covariance components models. J. Am. Stat. Assoc. 76: 341–353.

FAO. 1980. Comparative studies on freshwater fisheries, report of a workshop held at Istituto Italiano di Idrobiologia, Pallanza, Italy, 4–8 Sept. 1978. FAO Fish. Tech. Pap. 198: 46 p.

FOX, W.W. 1970. An exponential yield model for optimizing exploited fish populations. Trans. Am. Fish. Soc. 99: 80–88.

GARCIA, S., AND L. LERESTE. 1981. Life cycles, dynamics, exploitation, and management of coastal penaeid shrimp stocks. FAO Fish. Tech. Pap. 203: 215 p.

HANSON, J., AND W. LEGGETT. 1982. Empirical prediction of fish biomass and yield. Can. J. Fish. Aquat. Sci. 39: 257–263.

HOENIG, J.M., D.M. HEISEY, W.D. LAWING, AND D.H. SCHUPP. 1987. An indirect rapid methods approach to assessment. Can. J. Fish. Aquat. Sci. 44(Suppl. 2): 324–338.

HOENIG, J.M., AND N. HOENIG. 1986. Fitting a stock production model subject to a constraint. Univ. R.I. Dep. Stat. Tech. Rep. TR86–186: 10 p.

MARSHALL, N. 1980. Fishery yields of coral reefs and adjacent shallow-water environments, p. 103–109. *In* S.B. Saila and P.M. Roedel [ed.] Stock assessment for tropical small-scale fisheries. International Center for Marine Resource Development, University of Rhode Island, Kingston, R.I.

MARTEN, G.G., AND J.J. POLOVINA. 1982. A comparative study of fish yields from various tropical ecosystems, p. 255–289. *In* D. Pauly and G. Murphy [ed.] Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, International Center for Living Aquatic Resources Management, Manila, The Philippines.

MIYABE, N. 1989. An updated production model analysis on Atlantic bigeye tuna as of 1986. Int. Comm. Cons. Atl. Tunas Coll. Vol. Sci. Pap. 30: 1–5.

MUNRO, J.L., AND R. THOMPSON. 1973. The biology, ecology, exploitation and management of Carribean reef fishes. Part II. The Jamaican fishing industry, the area investigated and the objectives and methodology of the ODA/UWI Fisheries Ecology Research Project, Research Report 3, University of the West Indies Zoology Department. 44 p.

POLOVINA, J.J. 1989. A system of simultaneous dynamic production and forecast models for multispecies or multiarea applications. Can. J. Fish. Aquat. Sci. 46: 961–963.

PRESS, S.J. 1989. Bayesian statistics: principles, models, and applications. John Wiley & Sons, New York, N.Y. 237 p.

RYDER, R.A. 1965. A method for estimating the potential fish production of north-temperate lakes. Trans. Am. Fish. Soc. 94: 214–218.

SCHAEFER, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Bull. Inter-Am. Trop. Tuna Comm. 2: 247–268.

SCHNUTE, J. 1987. Data uncertainty, model ambiguity, and model identification. Nat. Resour. Model. 2: 159–212.

SCHNUTE, J. 1989. The influence of statistical error in stock assessments: illustrations from Schaefer's model, p. 101–109. *In* R.J. Beamish and G.A. McFarlane [ed.] Effect of ocean variability on recruitment and an evaluation of parameters used in stock assessment models. Can. Spec. Publ. Fish. Aquat. Sci. 108.

SEBER, G.A.F., AND C.J. WILD. 1989. Nonlinear regression. John Wiley & Sons, New York, N.Y. 768 p.

STOCKER, M., AND T.H. BUTLER. 1990. Surplus production model analysis of crab (*Cancer magister* Dana) fisheries of British Columbia, Canada. Fish. Res. 9: 231–254.

STRENIO, J., H.L. WEISBERG, AND A.S. BRYK. 1983. Empirical Bayes estimation of individual growth-curve parameters and their relationship to covariates. Biometrics 39: 71–86.

WALTERS, C.J. 1986. Adaptive management of renewable resources. Macmillan, New York, N.Y. 374 p.

WALTERS, C., AND D. LUDWIG. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci. 51: 713–722.

Appendix

Computation of the Empirical Bayes Estimates

Algebraically, the specifics for the situation represented by equation (8) of the main text are as follows. Let equation (8) be expressed in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

where $\mathbf{y}$ is an $n \times 1$ vector of catches ($n = \Sigma n_i$) obtained by concatenating the catches from all I stocks into a single vector and $\mathbf{X} = [\mathbf{B}_x, \mathbf{D}_x]$ where $\mathbf{B}_x = [\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_I']'$, $\mathbf{D}_x = \text{diag}\,(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_I)$, and

$$\mathbf{X}_i = \begin{bmatrix} x_{i1} & -x_{i1}^2 \\ x_{i2} & -x_{i2}^2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_{in_i} & -x_{in_i}^2 \end{bmatrix}.$$

Also, $\boldsymbol{\gamma}' = [\boldsymbol{\mu}', \mathbf{v}_\beta']$ where $\boldsymbol{\mu}' = [\mu_1, \mu_2]$, $\mathbf{v}_\beta' = [\mathbf{v}_1', \mathbf{v}_2', \ldots, \mathbf{v}_I']$ and $\mathbf{v}_i' = [\beta_{1i} - \mu_1, \beta_{2i} - \mu_2]$. Note that $\boldsymbol{\gamma}'$ is a $1 \times 2(I + 1)$ row vector. Finally $\boldsymbol{\epsilon}$ is the $n \times 1$ vector of normally distributed residuals assumed to be independent with mean zero and variance σ^2. In addition, we define the $2I \times 2I$ block diagonal matrix $\mathbf{V}$ as diag $(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}, \ldots, \boldsymbol{\Sigma})$.

To begin the EM algorithm, we need initial guesses for $\mathbf{V}$ (or, equivalently, $\boldsymbol{\Sigma}$) and σ^2 which we denote by $\mathbf{V}^{(0)}$ (or $\boldsymbol{\Sigma}^{(0)}$) and $\sigma^{2(0)}$, respectively (superscripts in parentheses indicate the number of iterations performed).

The sufficient statistics for σ^2 and $\mathbf{V}$ are $\boldsymbol{\epsilon}'\boldsymbol{\epsilon}$ and $\mathbf{v}_\beta \mathbf{v}_\beta'$, respectively (Dempster et al. 1981). In the E (expectation) step the expected values of these statistics are computed as follows. Let

$$\text{(A.1)} \quad \mathbf{P}^{(k)} = \mathbf{X}'\mathbf{X} + \sigma^{2(k)} \begin{bmatrix} \mathbf{O}_{2\times 2} & \mathbf{O}_{2\times 2I} \\ \mathbf{O}_{2I\times 2} & \mathbf{V}^{(k)-1} \end{bmatrix}.$$

The regression coefficients are then estimated as

$$\text{(A.2)} \quad \boldsymbol{\gamma}^{(k)} = [\boldsymbol{\mu}^{(k)\prime}, \mathbf{v}_\beta^{(k\prime)}]' = \mathbf{P}^{(k)-1}\mathbf{X}'\mathbf{y}$$

and the estimate of the variance–covariance matrix of the estimated regression coefficients, $\mathbf{V}_\gamma$, is given by

$$\text{(A.3)} \quad \mathbf{V}_\gamma^{(k)} = \begin{bmatrix} \mathbf{V}_\mu^{(k)} & \mathbf{C}_{\mu,\beta}^{(k)} \\ \mathbf{C}_{\mu,\beta}^{(k)\prime} & \mathbf{V}_\beta^{(k)} \end{bmatrix} = \sigma^{2(k)}\mathbf{P}^{(k)-1}$$

where $\mathbf{V}_\mu$ and $\mathbf{V}_\beta$ denote the variance matrices of $\boldsymbol{\mu}$ and $\mathbf{v}_\beta$, respectively, and $\mathbf{C}_{\mu,\beta}$ denotes the covariance matrix between $\boldsymbol{\mu}$ and $\mathbf{v}_\beta$.

Conditional on these estimates of $\boldsymbol{\gamma}$ and $\mathbf{V}_\gamma$, the expected values of the sufficient statistics for σ^2 and $\mathbf{V}$ are

$$\text{(A.4)} \quad S_{\sigma^2} = E(\boldsymbol{\epsilon}'\boldsymbol{\epsilon} \mid \mathbf{y}, \mathbf{V}^{(k)}, \sigma^{2(k)})$$

$$= \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\boldsymbol{\gamma}^{(k)} + \boldsymbol{\gamma}^{(k)\prime}\mathbf{X}'\mathbf{X}\boldsymbol{\gamma}^{(k)} + \text{tr}\,(\mathbf{X}'\mathbf{X}\mathbf{V}_\gamma^{(k)})$$

$$\text{(A.5)} \quad \mathbf{S}_V = E(\mathbf{v}_\beta \mathbf{v}_\beta' \mid \mathbf{y}, \mathbf{V}^{(k)}, \sigma^{2(k)}) = \mathbf{v}_\beta^{(k)}\mathbf{v}_\beta^{(k)\prime} + \mathbf{V}_\beta^{(k)}.$$

This completes the E step.

For the M (maximization) step, we compute the maximum likelihood estimates of σ^2 and $\mathbf{V}$. For σ^2 this is simply

$$\text{(A.6)} \quad \sigma^{2(k+1)} = S_{\sigma^2}/n.$$

The estimator for $\boldsymbol{\Sigma}$ (and hence $\mathbf{V}$) can be obtained from $\mathbf{S}_v$. Specifically

$$\text{(A.7)} \quad \boldsymbol{\Sigma}_{pq}^{(k+1)} = \sum_{i=1}^{I} (\mathbf{V}_{\gamma\,2i+p,2i+q}^{(k)} + \gamma_{2i+p}^{(k)}\gamma_{2i+q}^{(k)})/I$$

for $p,q = 1,2$. Here, $\mathbf{V}_{\gamma 2i+p,2i+q}^{(k)}$ is the $2i + p, 2i + q$ element of the matrix $\mathbf{V}_\gamma$ in the kth iteration.

Given $\sigma^{2(k+1)}$ and $\boldsymbol{\Sigma}_{pq}^{(k+1)}$ for $k = 0$, one can compute (A.1) through (A.7) and then use the results to obtain values $\sigma^{2(k+1)}$ and $\boldsymbol{\Sigma}_{pq}^{(k+1)}$ for $k = 1$, recompute (A.1) through (A.7), etc., until adequate convergence is achieved.