



A Kalman Filter Approach to Catch-at-Length Analysis

Author(s): Patrick J. Sullivan

Source: *Biometrics*, Vol. 48, No. 1 (Mar., 1992), pp. 237-257

Published by: International Biometric Society

Stable URL: <https://www.jstor.org/stable/2532752>

Accessed: 24-10-2018 14:31 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

International Biometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Biometrics*

A Kalman Filter Approach to Catch-at-Length Analysis

Patrick J. Sullivan

International Pacific Halibut Commission, P.O. Box 95009,
Seattle, Washington 98145, U.S.A.

SUMMARY

A state-space representation of a length-structured population under commercial harvest is described and a Kalman filter is used to develop the conditional likelihood equation needed for estimating the underlying system parameters. The state of the system is characterized using conventional fisheries theory with commercial harvest representing the observations taken on the population. The conditional likelihood framework embedded in the Kalman filter facilitates the incorporation of both system stochasticity as well as observation error in the development of the overall likelihood equation. Within this framework a maximum likelihood approach is used to estimate population parameters while taking into account both sources of error.

1. Introduction

Mathematical models used in the study of plant and animal populations are frequently structured by individual length, weight, or age (Schnute, 1987). Structured population models are useful in that they stratify heterogeneous data into more homogeneous subunits, thus differentiating within-strata from between-strata variation in population abundance over time. A Kalman filter formulation (Kalman, 1960; Kalman and Bucy, 1961) of a length-structured population model is used here to obtain the maximum likelihood estimates of population parameters for fishing mortality, selectivity, initial population abundance, recruitment, and growth from data taken from the catch and categorized by length class. Collie and Sissenwine (1983) develop a Kalman filter approach, using a modified DeLury (1947) method, for estimating population size using catch or catch-at-age data and Mendelssohn (1988) develops this idea further, treating the observed population sizes as missing values and arriving at population estimates using the EM algorithm. Gudmundsson (Report RH-02-87, Raunvisindastofnun Haskolans, University of Iceland, 1987) also addresses population parameter estimates using catch-at-age data in the context of state-space modeling, whereas Schnute, Richards, and Cass (1989a, 1989b) examine survivorship, recruitment, and growth using a moving average process and a size-structured model.

A set of two equations describes the process modeled here. The first depicts the change in population abundance at length over time using a length- and time-specific survivorship transition matrix and a length-specific time-invariant growth transition matrix. The growth transition matrix makes use of a stochastic version of the von Bertalanffy growth model (Sainsbury, 1982; Gutreuter and Anderson, 1985; Sullivan, Lai, and Gallucci, 1989) that represents growth as a Markov chain. The second equation depicts the relationship between population abundance and catch. The two equations combined can be used to give the

Key words: Catch-at-age analysis; Catch-at-length analysis; Fisheries; Kalman filter; Maximum likelihood estimation.

predicted catch conditioned on the set of previously observed catches and the given model. The conditional structure of the likelihood allows the analysis of more complex error structures than are allowed under traditional least squares formulations.

In the application of this approach, specific models describing the system and observation processes are used and contrasted under several postulated error structures. More generally, however, the Kalman filter can be used to develop likelihoods that represent a variety of error structures, adding another dimension to the already rich literature on stage-structured modeling. Getz and Haight (1989) provide for a recent discussion of stage-structured approaches to population assessment.

The present paper proceeds as follows. First, a model is developed using a state-space representation of a system or state equation describing the change in population numbers with time, followed by the development of an observation equation relating catch observations to the current level of the state. Next, a Kalman filter likelihood is developed by first examining a conditional representation of the means and variances of the state and observation processes. Under a Gaussian assumption, this conditional representation is used to construct a likelihood, which is formed as the product of the conditional probabilities of the process. The parameters used in the general model are then respecified using several standard fisheries models. An example error structure is then presented to indicate the form that the variance-covariance matrices may take for the state or observation processes. Following this development of the likelihood, the estimation procedure is applied to catch-at-length data on a Gulf of Alaska fisheries stock. The implications of explicitly modeling both state and observation error are considered. Finally, a discussion of assumptions and alternatives is presented with the concluding section summarizing the results.

2. Process Description

Let \mathbf{N}_t be a vector representation of the length (alternatively, weight or age) frequency distribution of the population observed at discrete time increments t . A change in the length frequency distribution of a population results from three basic processes: recruitment, survivorship, and growth. Recruitment is represented by a length-structured recruitment vector \mathbf{R}_t that is added to the population at time t . Survivorship is represented by a survivorship matrix \mathbf{S}_t , whose diagonal elements give the probability that an individual of length i will survive to the following time step. Finally, growth, or a change from one length class to another, is represented by a Markov chain based on a lower-diagonal transition matrix \mathbf{P} with nonzero elements $P_{l,l'}$, $1 \leq l \leq l' \leq n$, where n is the total number of length classes. Thus growth is viewed as the likelihood that an individual in length class l will grow into length class l' by the end of the time step. Based on these assumptions, a vector equation for the change in the abundance of the structured population over time is written

$$\mathbf{N}_t = \mathbf{P}\mathbf{S}_{t-1}\mathbf{N}_{t-1} + \mathbf{R}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{\Sigma}^{\mathbf{w}_t}), \quad (1)$$

which, when written in expanded notation, is as follows:

$$\begin{bmatrix} N_{1,t} \\ N_{2,t} \\ \vdots \\ N_{n,t} \end{bmatrix} =$$

$$\begin{bmatrix} P_{1,1} & 0 & & & 0 \\ P_{1,2} & P_{2,2} & & & \\ & & * & & \\ & & & P_{l,l'} & \\ & & & & * & 0 \\ P_{1,n} & P_{2,n} & & & P_{n-1,n} & P_{n,n} \end{bmatrix} \begin{bmatrix} S_{1,t-1} & 0 & & & 0 \\ 0 & S_{2,t-1} & & & \\ & & * & & \\ & & & S_{l,t-1} & \\ & 0 & & & * & 0 \\ & & & & 0 & S_{n,t-1} \end{bmatrix} \begin{bmatrix} N_{1,t-1} \\ N_{2,t-1} \\ \vdots \\ N_{n,t-1} \end{bmatrix} \\ + \begin{bmatrix} R_{1,t} \\ R_{2,t} \\ \vdots \\ R_{n,t} \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ \vdots \\ w_{n,t} \end{bmatrix},$$

where $R_{l,t} \geq 0$, $0 \leq P_{l,l'} \leq 1$, and $0 \leq S_{l,t} \leq 1$ for all lengths l and l' and times t , and where w_t is the multivariate Gaussian error associated with the system. The growth matrix \mathbf{P} is shown here as a lower-triangular matrix representing growth as an increasing process. Note that the columns of \mathbf{P} sum to 1, which represents the assumption that an individual starting off in one length class must either remain at its present length or grow into a larger length class, if it survives. The element $P_{n,n}$ representing the culmination of all individuals of length class n or larger is therefore identically equal to 1 if there is no migration out of the population. In the case of an age-structured population, with the time step equal to the age class length, the subdiagonal elements $P_{l,l+1}$ are all set equal to 1, with $P_{n,n}$ equal to 0 or 1 depending on whether or not there is an accumulation of older individuals in the last age category over time.

A second equation is developed in parallel to the first describing the set of observations made on the system over time. Let \mathbf{C}_t be the vector representation of the length (alternatively, weight or age) frequency distribution of individuals occurring in the catch. The catch by length is assumed proportional to the number of individuals by length in the population. Represent this constant of proportionality by the matrix μ_t of diagonal elements $\mu_{l,t}$. The relationship between catch and numbers is then concisely stated as follows:

$$\mathbf{C}_t = \mu_t \mathbf{N}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \Sigma^v), \quad (2)$$

or again, in the expanded matrix notation

$$\begin{bmatrix} C_{1,t} \\ C_{2,t} \\ \vdots \\ C_{n,t} \end{bmatrix} = \begin{bmatrix} \mu_{1,t} & 0 & & & 0 \\ 0 & & & & \\ & \mu_{2,t} & * & & \\ & & & \ddots & \\ & & & & \mu_{l,t} & * & 0 \\ 0 & & & & 0 & \mu_{n,t} \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ \vdots \\ N_{n,t} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ \vdots \\ v_{n,t} \end{bmatrix}$$

This population growth model is incorporated into a state-space statistical procedure known as the Kalman filter to obtain a maximum likelihood approach to estimating the population parameters associated with the transition matrices $\phi_t = \mathbf{P}\mathbf{S}_t$ and μ_t , to obtain conditional estimates of population abundance over time given the transition matrices ϕ_t and μ_t . Such estimation and prediction procedures are often crucial to the management of fish and wildlife populations, where observations are made in terms of harvest, or catch, at length over time. Although the argument developed here assumes the observations are given in terms of catch, observations can be made by other means such as by hydroacoustic

or trawl surveys, in which case the structure of the observation equation would change accordingly. For example, in the case where in addition to catch observations for each length class, one has other (possibly independent) observations on the number of individuals in each length class, these additional observations are incorporated directly into the formulation above as follows:

$$\begin{bmatrix} C_{1,t} \\ \vdots \\ C_{n,t} \\ H_{1,t} \\ \vdots \\ H_{n,t} \end{bmatrix} = \begin{bmatrix} \mu_{1,t} & & \\ & * & \\ & & \mu_{n,t} \\ \rho_{1,t} & & \\ & * & \\ & & \rho_{n,t} \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ \vdots \\ N_{n,t} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ \vdots \\ v_{n,t} \\ u_{1,t} \\ \vdots \\ u_{n,t} \end{bmatrix},$$

where \mathbf{H}_t is the vector of additional observations taken on the population, ρ_t is the proportionality matrix relating \mathbf{N}_t to \mathbf{H}_t , and \mathbf{u}_t is the error vector on those auxiliary observations. Without loss of generality we may subsume these values into the vectors \mathbf{C}_t , and the matrices μ_t and Σ^v , and proceed as before.

Note that in this application \mathbf{w}_t and \mathbf{v}_t are assumed to be uncorrelated. Correlation between the population and catch equations can be accounted for by specifying the joint distribution of \mathbf{w}_t and \mathbf{v}_t with the appropriate correlation factors. The conditional distribution of the one vector variable given the other is then what is used in the computations discussed below.

The population dynamics model shown in equations (1) and (2) is incorporated into a state-space Kalman filter framework (Kalman, 1960; Kalman and Bucy, 1961) to obtain a likelihood function for the state and observation variables \mathbf{N}_t and \mathbf{C}_t in terms of the underlying parameters, \mathbf{P} , \mathbf{S} , \mathbf{R} , and μ , where \mathbf{S} , \mathbf{R} , and μ represent the sets of \mathbf{S}_t , \mathbf{R}_t , and μ_t over the time periods considered in the estimation.

This likelihood function is maximized over the parameter space for a given set of observations to obtain the maximum likelihood estimates of the underlying parameters. Once estimates of these parameters are obtained, equations (1) and (2) are used to predict future values of the state and observation variables. These predicted values are later updated with the addition of new observations taken in time. Below we develop the theory leading to the likelihood function by first examining the theoretical basis for the conditional estimate of the state vector variable \mathbf{N}_t given the catch observations \mathbf{C}_τ for time periods $\tau = 1, \dots, t$ (abbreviated as \mathbf{C}^t). Following this is a discussion of the theory behind the conditional estimate of the observation vector variable \mathbf{C}_t at time t leading up to a formulation of the likelihood in terms of the product of the conditional observation probabilities.

3. The Kalman Filter Likelihood

The likelihood is now developed using conditional probability theory. Let the joint probability of the variables $\{C_0, C_1, \dots, C_T\}$ be represented as the product of a set of conditional probabilities:

$$p(C_0, C_1, \dots, C_T) = p(C_T | C_0, C_1, \dots, C_{T-1})p(C_{T-1} | C_0, C_1, \dots, C_{T-2}) \cdots p(C_0).$$

With this formulation, a likelihood function can be constructed that incorporates error structures more general than those commonly used in fisheries research. Such a formulation avoids the assumption of independence that is made when the joint probability is written as $p(C_0, C_1, \dots, C_T) = p(C_0)p(C_1) \cdots p(C_T)$ in order to represent the log-likelihood as a simple sum of squares. The conditional estimation maintains the description of the log-

likelihood as a sum, and retains the recursive nature of the likelihood formulation that is useful in its application.

The conditional representation is given first followed by the formation of the likelihood. A more specific parameterization, applicable to fisheries research, then follows.

3.1 Conditional Estimation

In state-space terminology we have two equations, a state equation (1) describing the state of the population at time t , and an observation equation (2) describing the observations made on the system over time. Assume that the observations are taken over equally spaced time increments so that time is marked off in unit increments (e.g., $t - 1, t, t + 1, \dots$). (This assumption may be relaxed to cover the case where time increments are not equally spaced when differences in the length of the time interval can be accounted for elsewhere. One way this may be accomplished is through respecification of the projection matrices used in the formulation of the state and observation equations. For example, in the application to follow, this may be accomplished by allowing the scale of the mortality and growth rates to vary in correspondence with variations in the time increment, effectively making all the time increments unit time steps.) Assume that the errors in the state and observation equations are multivariate Gaussian. Then the conditional distribution of abundance N_{t-1} at time $t - 1$ given the matrix of catch vectors $C^{t-1} = \{C_\tau: \tau \leq t - 1\}$ for times up to and including time $t - 1$ is given by

$$(N_{t-1} | C^{t-1}) \sim \text{MVN}(N_{t-1}^*, \Sigma^{N_{t-1} | C^{t-1}}), \quad (3)$$

where $(X|Y) \sim f$ will be used to denote the statement that the conditional distribution of X given Y is f , where $N_t^* = E(N_t | C^t)$ —that is, the expected value of N_t given all the observations up to and including those taken at time t .

Equation (3) results from the following argument. If the distribution of N_t is multivariate Gaussian, linear transformations of N_t are also multivariate Gaussian [for example, the distribution of the linearly transformed variable given by equation (1) is multivariate Gaussian]. Furthermore, the conditional distribution of C_t given N_t (i.e., $f_{C_t | N_t}$) is also multivariate Gaussian [for example, the conditional distribution given by (2)]. As a consequence, the conditional distribution of N_t given the observation C_t , $f_{N_t | C_t}$, which is proportional to $f_{C_t | N_t} f_{N_t}$ by Bayes' rule, is also multivariate Gaussian (Lipster and Shirayev, 1978). The derivation of equation (3) and the subsequent equations is recursive in the sense that if we are given a conditional random vector variable of the form (3), then in the course of evaluating the random variable at subsequent steps, we will eventually return to a variable of the form given in (3) with time t incremented by 1. Assume for the moment that the variable $(N_{t-1} | C^{t-1})$ is given and that the parameters in the projection matrices in equation (1) are known. We can calculate the expected value of N_t conditioned on the observed catch up to and including time $t - 1$ as follows:

$$(N_t | C^{t-1}) \sim \text{MVN}(\mathbf{P}\mathbf{S}_{t-1}\mathbf{N}_{t-1}^* + \mathbf{R}_t, \Sigma^{N_t | C^{t-1}}), \quad (4)$$

where

$$\Sigma^{N_t | C^{t-1}} = \mathbf{P}\mathbf{S}_{t-1}\Sigma^{N_{t-1} | C^{t-1}}(\mathbf{P}\mathbf{S}_{t-1})^T + \Sigma^w_t.$$

Note that the index of time for \mathbf{N} has changed in the latest variance-covariance matrix.

Once the catch has been observed for the next time period t , the conditional distribution of abundance depicted in (3) is updated with the latest observation taken at time t . Thus $E(N_t | C^t)$ incorporates the latest information that is present in observation C_t to get

$$\mathbf{N}_t^* = (\mathbf{P}\mathbf{S}_{t-1}\mathbf{N}_{t-1}^* + \mathbf{R}_t) + \Sigma^{N_t | C^{t-1}} \mu_t^T (\mu_t \Sigma^{N_t | C^{t-1}} \mu_t^T + \Sigma^v_t)^{-1} \{C_t - \mu_t(\mathbf{P}\mathbf{S}_{t-1}\mathbf{N}_{t-1}^* + \mathbf{R}_t)\}, \quad (5)$$

with the conditional variance–covariance matrix $\text{var}(\mathbf{N}_t | \mathbf{C}')$ expressed as

$$\Sigma^{\mathbf{N}_t | \mathbf{C}'} = \Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}} - \Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}} \boldsymbol{\mu}_t^T (\boldsymbol{\mu}_t \Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}} \boldsymbol{\mu}_t^T + \Sigma^{\mathbf{v}_t})^{-1} \boldsymbol{\mu}_t \Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}}. \quad (6)$$

Consequently, given the initial conditions $\mathbf{N}_0^* = E(\mathbf{N}_0 | \mathbf{C}_0)$ and $\Sigma^{\mathbf{N}_0} = \text{var}(\mathbf{N}_0 | \mathbf{C}_0)$, we can derive equations (3)–(6) for any time t conditioned on the observations available up through that time [Lipster and Shirayev (1978); see also the derivation by Meinhold and Singpurwalla (1983)]. Note that equation (5) is just the sum of the predicted state, given the previous set of observations, with a weighted correction for the current observation.

Equation (5) is interpreted as an estimate of where the system should be, based on \mathbf{P} , \mathbf{S}_{t-1} , and \mathbf{R}_t , given its previous state \mathbf{N}_{t-1}^* , and updated by the most recent observation \mathbf{C}_t . The influence of the observation depends on the difference between what was observed and what was expected, weighted by the Kalman gain, which is roughly the ratio of the prediction error to the total error. Thus if the variability in the observation error \mathbf{v}_t is high, through $\Sigma^{\mathbf{v}_t}$, the weight of the contribution of the correction term is low. Conversely, if the variability in the state error \mathbf{w}_t is high (through $\Sigma^{\mathbf{w}_t}$), leading to high variability in the conditional estimate of \mathbf{N}_t^* (through $\Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}}$), then contribution to the estimate of the state variables, which is based on the most recent observation vector \mathbf{C}_t , is accordingly given more weight. In an analogous way, equation (6) represents an update—that is, a reduction—in the estimate of the variance of the state given the most current observation. Again, if the state error is high relative to the observation error then the second term on the right of the equality will be large, contributing greatly to reducing the variance with each observation. Conversely, if the observation error is large, then $(\boldsymbol{\mu}_t \Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}} \boldsymbol{\mu}_t^T + \Sigma^{\mathbf{v}_t})^{-1}$ will be small and the initial variance will be reduced only slightly.

In biology, and in particular in fisheries biology, variation in the estimates of population abundance can result from variation in the state process as well as variation in the observation process. State variability, for example, may result from environmentally induced factors operating on survivorship, growth, and recruitment, whereas observation error may occur as a result of underreported catch, unreported discard, or sampling error.

3.2 Likelihood Formation

To determine the maximum likelihood estimates of the matrices \mathbf{P} , \mathbf{S} , \mathbf{R} , and $\boldsymbol{\mu}$, we first examine the conditional distribution of \mathbf{C}_t given \mathbf{C}^{t-1} . By an argument similar to that presented above, we see that the conditional distribution of the observation vector \mathbf{C}_t given \mathbf{C}^{t-1} is multivariate Gaussian with mean

$$E(\mathbf{C}_t | \mathbf{C}^{t-1}) = \boldsymbol{\mu}_t (\mathbf{P} \mathbf{S}_{t-1} \mathbf{N}_{t-1}^* + \mathbf{R}_t) \quad (7)$$

and variance

$$\text{var}(\mathbf{C}_t | \mathbf{C}^{t-1}) = \boldsymbol{\mu}_t \Sigma^{\mathbf{N}_t | \mathbf{C}^{t-1}} \boldsymbol{\mu}_t^T + \Sigma^{\mathbf{v}_t}. \quad (8)$$

Under the assumption of a multivariate Gaussian distribution, the conditional probability density of the observation vector variable \mathbf{C}_t can be written as

$$p(\mathbf{C}_t | \mathbf{C}^{t-1}) = (2\pi)^{-n/2} |\Sigma^{\mathbf{C}_t}|^{-1/2} \exp\{-\frac{1}{2}(\mathbf{C}_t - \mathbf{C}_t^*)^T |\Sigma^{\mathbf{C}_t}|^{-1}(\mathbf{C}_t - \mathbf{C}_t^*)\}, \quad (9)$$

where \mathbf{C}_t^* and $\Sigma^{\mathbf{C}_t}$ are the vector mean and the variance–covariance matrix given by equations (7) and (8), respectively. Given \mathbf{N}_0 , and treating it as a variable with known distribution by assuming a Bayesian prior, or treating it as a parameter vector to be estimated by assuming $\Sigma^{\mathbf{N}_0}$ identically equals 0, the initial catch distribution becomes

$$(\mathbf{C}_0 | \mathbf{N}_0) \sim \text{MVN}(\boldsymbol{\mu}_0 \mathbf{N}_0^*, \Sigma),$$

where $\Sigma = \boldsymbol{\mu}_0 \Sigma^{\mathbf{N}_0} \boldsymbol{\mu}_0^T + \Sigma^{\mathbf{v}_0}$.

The initial conditional catch density is

$$p(\mathbf{C}_0 | \mathbf{N}_0) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{C}_0 - \boldsymbol{\mu}_0 \mathbf{N}_0^*)^T |\boldsymbol{\Sigma}|^{-1} (\mathbf{C}_0 - \boldsymbol{\mu}_0 \mathbf{N}_0^*)\right\}. \quad (10)$$

Consequently, the likelihood of the vector of parameters $\boldsymbol{\Theta} = (\mathbf{P}, \mathbf{S}, \mathbf{R}, \boldsymbol{\mu}, \mathbf{N}_0^*)$, given the observed catch up to time t , may be expressed as follows:

$$L(\boldsymbol{\Theta}) = \prod_{\tau=1}^t p(\mathbf{C}_\tau | \mathbf{C}^{\tau-1}) p(\mathbf{C}_0 | \mathbf{N}_0) \quad (11)$$

and the log-likelihood becomes

$$\begin{aligned} \ln L(\boldsymbol{\Theta}) &= \sum_{\tau=1}^t \ln p(\mathbf{C}_\tau | \mathbf{C}^{\tau-1}) + \ln p(\mathbf{C}_0 | \mathbf{N}_0) \\ &= \text{Constant} - \frac{1}{2} \sum_{\tau=0}^t \ln |\boldsymbol{\Sigma}^{\mathbf{C}_\tau}| - \frac{1}{2} \sum_{\tau=0}^t (\mathbf{C}_\tau - \mathbf{C}_\tau^*)^T |\boldsymbol{\Sigma}^{\mathbf{C}_\tau}|^{-1} (\mathbf{C}_\tau - \mathbf{C}_\tau^*), \end{aligned} \quad (12)$$

which is maximized to obtain the maximum likelihood estimate of the parameter vector $\boldsymbol{\Theta}$. A simple algorithm for computing the components of the likelihood is provided in the Appendix.

3.3 Parameter Specification Based on Biological Relationships

The parameters $\boldsymbol{\Theta}$ are now transformed into a parameter set that represents standard fisheries population dynamics and growth equations. The objective is primarily to constrain the parameters to an estimable parameter space and secondarily to relate the parameters used in the model to those conventionally used in the fisheries literature. The corresponding fisheries equations are now briefly summarized. A detailed description of these equations and their derivation is given elsewhere (Sullivan et al., 1989; Sullivan, unpublished Ph.D. dissertation, University of Washington, Seattle, 1988). The Baranov catch equation applied to length classes may be used to define the elements of the exploitation matrix $\boldsymbol{\mu}_t$. The length- and time-specific element $\mu_{l,t}$ represents the proportion of those fish that die as a direct result of fishing:

$$\mu_{l,t} = \frac{F_{l,t}}{Z_{l,t}} (1 - e^{-Z_{l,t}}), \quad (13)$$

where $F_{l,t}$ is the length- and time-specific instantaneous fishing mortality rate and $Z_{l,t}$ is the length- and time-specific total instantaneous mortality rate. The total instantaneous mortality rate $Z_{l,t}$ is the sum of a known natural mortality rate $M_{l,t}$ and the fishing mortality rate $F_{l,t}$, and it is used in defining the elements of the $l \times l$ diagonal survivorship matrix \mathbf{S}_t as

$$S_{l,t} = e^{-Z_{l,t}}. \quad (14)$$

Furthermore, the instantaneous fishing mortality rate is assumed to be separable into a product of a length-specific selectivity coefficient s_l and a full-recruitment fishing mortality rate f_t at time t following the formulation of Doubleday (1976), modified to apply to length:

$$F_{l,t} = s_l f_t.$$

The selectivity term s_l is represented as a two-parameter gamma function:

$$s_l = \frac{l^{\alpha_s} e^{-\beta_s l}}{\max_k (k^{\alpha_s} e^{-\beta_s k})},$$

where $k \in \{1, \dots, n\}$, based on an age-structured selectivity model suggested by Deriso, Quinn, and Neal (1985).

Each column of the growth transition matrix \mathbf{P} is constrained to follow an integrated version of a two-parameter gamma distribution for growth:

$$f(x | \alpha_l, \beta) = \frac{1}{\beta^{\alpha_l} \Gamma(\alpha_l)} x^{\alpha_l-1} e^{-x/\beta}, \quad (15)$$

where the length-specific parameters α_l are defined in terms of the von Bertalanffy (1934) growth parameters L_∞ and k by defining the mean change in length to be

$$\bar{\Delta}_l = (L_\infty - l)(1 - e^{-k}).$$

Thus, the probability of an individual's growing from length class l to length class l' is determined by integrating over the range (l'_1, l'_2) of the receiving length class l' at the beginning of the next time period:

$$P_{l,l'} = \int_{l'_1}^{l'_2} f(x | \alpha_l, \beta) dx, \quad (16)$$

and individuals entering a length class with midlength greater than L_∞ are assumed to remain in that length class subject to survival.

The distribution of new recruits into the population and the initial population distribution can also be constrained to follow a parametric distribution such as the one shown in equation (15) or a nonparametric distribution depending on the level of information available. In the example to follow, recruitment \mathbf{R}_t is separated into a time-dependent variable R_t and a length-dependent variable ρ_l , representing the proportion of recruits going into each length class. The proportion going into each class is approximated by a gamma distribution as in equation (15), but specified using two recruitment distribution parameters, α_r and β_r . Recruitment may be summarized then as the integral over the range (l_1, l_2) of the length class l times the expected total recruitment for that time period:

$$R_{l,t} = R_t \int_{l_1}^{l_2} f(x | \alpha_r, \beta_r) dx. \quad (17)$$

To summarize, equations (13), (14), (16), and (17) specify the transformation of the state-space parameters μ_t , \mathbf{P} , \mathbf{S}_t , and \mathbf{R}_t to the fisheries parameters $M_{l,t}$, f_t , α_s , β_s , L_∞ , k , β , R_t , α_r , β_r . Except where it becomes necessary to differentiate between the two, both the original and the transformed parameter sets will be represented simply by the vector Θ .

3.4 Example Error Structure

The multivariate error structure of this system of equations, as characterized by the variance-covariance matrices associated with the state and observation equation error variables, is general enough to encompass a wide range of correlation formulations. Each formulation leads to different constructions of the likelihood and consequently can result in different estimates of the parameter vector Θ .

As an example, an approximation to the likelihood equation can be developed when the error associated with the state equation arises from the sum of a series of multinomial random variables. Such an error structure follows naturally when growth is viewed as a stochastic process and the population is viewed in terms of individuals each having an independent probability of growing into a given length category in the following time step. In this case the surviving individual grows out of class l and into class l' with probability $P_{l,l'}$. It follows that, given N_l surviving individuals in length class l , we would expect to find

$N_l P_{l,l'}$ of these individuals in length class l' at the next time step. That is, the expected number of surviving individuals going from length class l to length class l' is

$$E(N_{l,l'}) = N_l P_{l,l'}$$

with

$$\text{var}(N_{l,l'}) = N_l P_{l,l'}(1 - P_{l,l'}) \quad \text{and} \quad \text{cov}(N_{l,l'}, N_{l,l''}) = -N_l P_{l,l'} P_{l,l''},$$

for $l' \neq l''$.

By a central limit theorem (see, for example, Rao, 1973), a multinomial random vector converges to a multivariate Gaussian random vector with the expectation given above, and the variance–covariance matrix specified in terms of the multinomial variance and covariance parameters as follows:

$$\begin{bmatrix} N_{l,1} \\ N_{l,2} \\ \vdots \\ N_{l,n} \end{bmatrix} \sim \text{MVN} \left(\begin{bmatrix} N_l P_{l,1} \\ N_l P_{l,2} \\ \vdots \\ N_l P_{l,n} \end{bmatrix}, \begin{bmatrix} N_l P_{l,1}(1 - P_{l,1}) & & -N_l P_{l,1} P_{l,n} \\ & N_l P_{l,2}(1 - P_{l,2}) & \\ -N_l P_{l,1} P_{l,n} & & N_l P_{l,n}(1 - P_{l,n}) \end{bmatrix} \right).$$

To determine the expected total number contributed to a given length class l' by all length classes ($l \in \{1, \dots, l'\}$), we sum over the expected number of individuals growing into length class l' , for all initial length classes l :

$$N_{l'} = \sum_{l=1}^{l'} N_{l,l'},$$

or, in matrix notation, for all length classes $l' \in \{1, \dots, n\}$:

$$\begin{bmatrix} N_1 \\ N_2 \\ \vdots \\ N_{l'} \\ \vdots \\ N_n \end{bmatrix} = \mathbf{I}^T \begin{bmatrix} N_{1,1} \\ \vdots \\ N_{1,l'} \\ \vdots \\ N_{1,n} \\ \vdots \\ N_{2,1} \\ \vdots \\ N_{2,l'} \\ \vdots \\ N_{2,n} \\ \vdots \\ N_{n,n} \end{bmatrix},$$

where the vector on the far right is a combination of the N_l vectors discussed in the previous equation stacked on top of one another, all $N_{l,k}$ equal 0 for $k > l'$ following the growth assumption made earlier, and the matrix \mathbf{I} is a concatenation of identity matrices I each of order n ,

$$\mathbf{I} = \begin{bmatrix} I \\ I \\ I \\ \vdots \\ I \end{bmatrix}.$$

If Ψ_l is the variance-covariance matrix for N_l , then the variance-covariance matrix for the linear combination of the N_l vectors is

$$\Psi = \mathbf{I}^T \text{diag}(\Psi_l) \mathbf{I} = \sum_{l=1}^n \Psi_l,$$

since the vectors N_l are assumed to be mutually independent.

Under the present parameterization, the $P_{l,t}$ are constant through time, but since N_l may change from time step to time step, so too will Ψ change from time step to time step. If we denote the Ψ for each time step by Ψ_t , then this corresponds to Σ^{w_t} used in equation (1). A similar, though simpler, argument can be made for Σ^{v_t} by assuming independence from w_t and estimating the variance matrix as a Gaussian approximation to a binomial process with diagonal elements equal to $N_{l,t} \mu_{l,t} (1 - \mu_{l,t})$ and off-diagonal elements equal to 0.

This argument can be taken one step further in order to sustain the Gaussian relationships used in the formation of the likelihood discussed in Section 3. Assume that the variances of both the observation and state error variables v_t and w_t are functions of the expected population level at the previous time step N_{t-1}^* , so that

$$v_t \sim \text{MVN}(\mathbf{0}, \mu_t \mathbf{P} \mathbf{S}_{t-1} \mathbf{N}_{t-1}^* (\mathbf{I} - \mu_t)^T)$$

and

$$w_t \sim \text{MVN}(\mathbf{0}, \mathbf{P} \mathbf{S}_{t-1} \mathbf{N}_{t-1}^* (\mathbf{I} - \mathbf{P})^T).$$

By assuming that the variances depend on an expected value, rather than on a random variable, the Bayesian argument using the prior and posterior distributions remains valid as do the conditioned and unconditioned Gaussian distributions that result. Alternatively, one might wish to have the variance, of the observation error for example, be a function of the current level of the state, in which case a generalized Kalman filter approach (Zehnwirth, 1988) might provide better theoretical motivation for the results presented here. Assuming that the variances depend on the previous expected population size, however, is not a severe restriction, especially for larger population sizes, and so this is the approach that will be taken in the application to follow.

4. Application

The performance of the estimation procedure is now examined under contrasting assumptions about the state and observation error structure, and with and without auxiliary information. The estimation procedure is applied to data collected from length frequency observations associated with total catch for western Gulf of Alaska walleye pollock (*Theragra chalcogramma*). The data are used here to exemplify the developed technique. The present example does not make use of the age information available for this species other than for comparing population parameter estimates.

Monthly length frequency data on pollock were collected by observers from the U.S. Foreign Fisheries Observer Program on board foreign trawlers and joint venture fishing vessels fishing in the Western Gulf of Alaska, during the years 1983, 1984, and 1985 (Nelson et al., 1984; Berger, Nelson, and Wall, 1985; Berger et al., 1986). The total foreign and joint venture fishery catch in biomass and number are available for each year with catch biomass also available on a monthly basis. To estimate the total number of individuals caught per month, the total annual catch number was distributed by month according to the percent biomass observed per month. The observer data, given in numbers of fish observed per one-centimeter length category per month, were converted into percent catch and multiplied by the total number of fish caught per month by the foreign and joint venture fishery. The monthly data sets were then combined into quarterly data sets (four

per year) with the observations classified into 22 two-centimeter wide length categories, with midlengths ranging from 19.5 cm to 61.5 cm. Independent estimates of stock abundance were obtained using hydroacoustic data collected during the first quarter of each year considered (Nunnallee and Williamson, 1989). Natural mortality was assumed to be constant for all lengths and set equal to .1 per quarter (Megrey, 1989).

Table 1 shows the parameter estimates obtained from six applications of the procedure. Standard error estimates and selected pairwise correlations (i.e., $|\rho| > .4$) for the first application, the full model with auxiliary information, are shown in Table 2. The estimates shown in Table 2 were obtained from the inverse Hessian that results from the optimization routine used in numerically maximizing the likelihood.

The first set of applications is based on a Kalman likelihood for the full model, which assumes a multivariate normal approximation of the multinomial stochasticity reflected in growth, with no error in the recruitment or survival processes, and with an independent multivariate Gaussian approximation of the multinomial stochasticity reflected in the catch sampling, where the diagonal elements of the observation covariance matrix are set equal to $\mu_{l,t}(1 - \mu_{l,t})P_lS_{l,t-1}N_{l,t-1}^*$ and the off-diagonal elements are set equal to 0. The initial covariance matrix Σ^{N_0} is set identically equal to 0 and the initial population vector N_0^* is estimated by the algorithm. These estimates are compared with a second set of estimates

Table 1
Parameter estimates under three Kalman likelihoods with and without auxiliary information

| | Full model | | Observation error only | | State error only | |
|------------|--------------------------|-----------------------------|--------------------------|-----------------------------|--------------------------|-----------------------------|
| | With aux. information | Without aux. information | With aux. information | Without aux. information | With aux. information | Without aux. information |
| N_0 | 4.40×10^9 | 1.08×10^{10} | 4.60×10^9 | 6.01×10^9 | 4.66×10^9 | 4.58×10^9 |
| R_1 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_2 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_3 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_4 | 1.94×10^8 | 1.00×10^9 | 2.56×10^8 | 0 | 2.99×10^8 | 2.81×10^8 |
| R_5 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_6 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_7 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_8 | 6.90×10^7 | 7.17×10^8 | 0 | 8.80×10^8 | 6.82×10^7 | 9.15×10^7 |
| R_9 | 0 | 0 | 0 | 0 | 0 | 0 |
| R_{10} | 0 | 0 | 0 | 0 | 0 | 0 |
| R_{11} | 0 | 0 | 0 | 0 | 0 | 0 |
| α_s | 1.24 | .29 | .00 | 2.03 | 1.08 | 1.51 |
| β_s | .42 | .34 | .00 | .00 | .00 | .00 |
| f_0 | .057 | .023 | .053 | .052 | .009 | .007 |
| f_1 | .005 | .002 | .002 | .003 | .010 | .007 |
| f_2 | .018 | .006 | .012 | .012 | .019 | .020 |
| f_3 | .038 | .013 | .031 | .026 | .025 | .022 |
| f_4 | .132 | .037 | .138 | .079 | .105 | .088 |
| f_5 | .006 | .002 | .056 | .003 | .002 | .002 |
| f_6 | .060 | .015 | .125 | .034 | .028 | .027 |
| f_7 | .083 | .020 | .177 | .042 | .031 | .034 |
| f_8 | .221 | .055 | .293 | .091 | .426 | .406 |
| f_9 | .018 | .004 | .027 | .007 | .023 | .028 |
| f_{10} | .012 | .003 | .014 | .004 | .011 | .012 |
| f_{11} | .071 | .016 | .105 | .026 | .010 | .074 |
| L_∞ | 58.59 | 57.45 | 51.11 | 64.09 | 63.71 | 61.76 |
| k | .102 | .121 | .174 | .096 | .104 | .125 |
| β | 1.57 | 1.45 | 7.26 | .38 | 1.06 | 1.14 |
| α_r | 25.8 | 16.7 | .0 | 197.8 | 29.8 | 25.0 |
| β_r | 1.06 | 1.93 | .08 | .22 | 1.17 | 1.34 |
| M | .1 ^a | .1 ^a | .1 ^a | .1 ^a | .1 ^a | .1 ^a |
| SSE | 3.14×10^8 | 1.70×10^8 | 8.84×10^{15} | 6.84×10^{14} | 2.71×10^{16} | 3.60×10^{16} |

^a Parameters fixed during the estimation procedure.

Table 2
*Estimated standard errors for Kalman likelihood
estimates shown in column 1 of Table 1 and
selected pairwise correlations*
Full model with auxiliary information

| Standard error | | Correlation | |
|----------------|--------------------|-----------------------|-------|
| N_0 | 6.34×10^7 | (N_0, R_4) | -.46 |
| R_4 | 4.53×10^7 | (β_s, f_0) | .97 |
| R_8 | 2.79×10^7 | (R_4, α_r) | -.52 |
| α_s | 6.70 | (R_4, β_r) | .49 |
| β_s | 3.15 | (α_r, β_r) | -1.00 |
| f_0 | .004 | (k, β) | .94 |
| f_1 | .001 | (L_∞, β) | -.96 |
| f_2 | .003 | (L_∞, k) | -.99 |
| f_3 | .004 | | |
| f_4 | .009 | | |
| f_5 | .002 | | |
| f_6 | .007 | | |
| f_7 | .009 | | |
| f_8 | .016 | | |
| f_9 | .005 | | |
| f_{10} | .004 | | |
| f_{11} | .011 | | |
| L_∞ | 5.47 | | |
| k | .031 | | |
| β | .57 | | |
| α_r | 21.74 | | |
| β_r | .92 | | |

Note: Fixed parameters and those with near-zero variances not included.

obtained by ordinary least squares, which represents a special case of the Kalman filter likelihood that assumes a deterministic process for the state, and independent and identically distributed observation error. A third set of estimates is provided that represents another special case of the Kalman likelihood that explicitly models state error only and assumes zero error for the observations and the multivariate normal approximation of the growth process as described above for the full model. These three sets of estimates are denoted as being from the full model, the model with observation error only, and the model with state error only. The performance of each approach is examined with and without the inclusion of auxiliary hydroacoustic information.

For the full model the auxiliary information, when used, takes the form of total abundance obtained during the first quarter of each year and is assumed to have independent error with variance proportional to the observed abundance. For the approach using observation error only (the least squares case), the auxiliary information is assumed to be independently and identically distributed. For the case that represents state error only, the auxiliary information is used in a way that follows the full-scale model where the fit is weighted by the estimated variance in abundance.

The differences in the error structure used in representing the auxiliary information reflect differences in assumptions made about the nature of the state and observation error. In the approaches that model state error, the observations are weighted; in the case where observation error only is modeled, they are not. Under the Kalman filter formulation, the nature of the system and observation error can be modeled explicitly. As a result, more appropriate weightings can be given the observations in determining parameter estimates.

Since the error is modeled explicitly and state-space information is passed from one time step to the next (through the updated variances) the stability problem commonly observed with weighted least squares estimates (where the weights are set proportional to the estimates) is not as severe. For this example error is being modeled as occurring only through growth; however, more sophisticated approaches that include survivorship and recruitment error may be attempted as well.

Notable in the comparison between estimates is the consistency with which recruitment is determined for the Kalman likelihoods that take state error into account. These values can be compared to estimated recruitment values of 3.2×10^8 and 9.0×10^7 three-year-old recruits obtained using catch-at-age analysis (Megrey, 1989) and estimates of 2.6×10^8 and 9.2×10^7 using hydroacoustic midwater trawl survey information (Nunnallee and Williamson, 1989). The consistency of these estimates is directly related to the more appropriate weighting provided by the conditionally updated variances. The recruitment distribution parameters are also indicative of better recruitment estimates with mean values (corresponding to a mean length of α_r, β_r) of 27.35 and 32.23 cm, respectively, for the full likelihood model in contrast to 0.0 and 43.52 cm for the least squares estimates, and 34.87 and 33.50 cm for the likelihood representing state error only. Midwater trawl survey information for these years indicates a range in mean length of three-year-olds of 32 to 35 cm.

In reference to the estimated growth parameters, we note that Hughes and Hirschhorn (1979), in examining growth relationships over the years 1973–1975 for western Gulf of Alaska pollock, found the arithmetic mean of estimates of k for ages 2–8 to be .33 and .30 per year for males and females, respectively (.08 per quarter for comparison), and the geometric mean of estimates of L_∞ to be 50.39 and 54.06 cm, respectively. The differences among the estimates shown in Table 1, and between these estimates and the independently estimated range of means just discussed, do not appear to be significant given the standard error estimates shown in Table 2 for the first application. The estimate of β in the application using observation error only with auxiliary information indicates a much wider dispersion of individuals during growth, spreading the modeled recruitment of time period 4 over a much broader range.

The fishing mortality estimates are relatively consistent across applications, over time, with noticeably higher mortality rates occurring at time periods 4 and 8, representing the first quarter of each year. The fishing mortalities are quite a bit lower in the initial year for the applications that model state error only. The selectivity parameters (α_s and β_s) appear to be quite sensitive to the inclusion of the auxiliary information and the error model used.

Table 2 indicates other aspects of model performance. High standard error estimates for the recruitment parameters α_r and β_r coupled with their high pairwise correlation ($\rho = -.995$) indicates that a better formulation of this component of the model might be found. Similarly, there is a high correlation among the growth parameters L_∞ , k , and β . This is a fairly common feature of the classic von Bertalanffy formulation and these results suggest that an alternative parameterization might be more appropriate (Ratkowsky, 1986).

For model comparison the residual (or error) sum of squares from each fit is provided (see SSE in Table 1). The residual sum of squares is proportional to the log of the likelihood shown in equation (9) and represents the sum of the squared differences between the observations and model estimates weighted by the process variation shown in (8). Care must be taken in interpreting model performance based solely on the residual sum of squares for each fit, however, since the Kalman formulation accounts for both process and observation error.

Figure 1 shows the observed versus estimated catches in numbers as determined by the Kalman likelihood estimation procedure that uses the full model and includes auxiliary information. Figure 2 shows the fit of the other five applications on the same scale as that

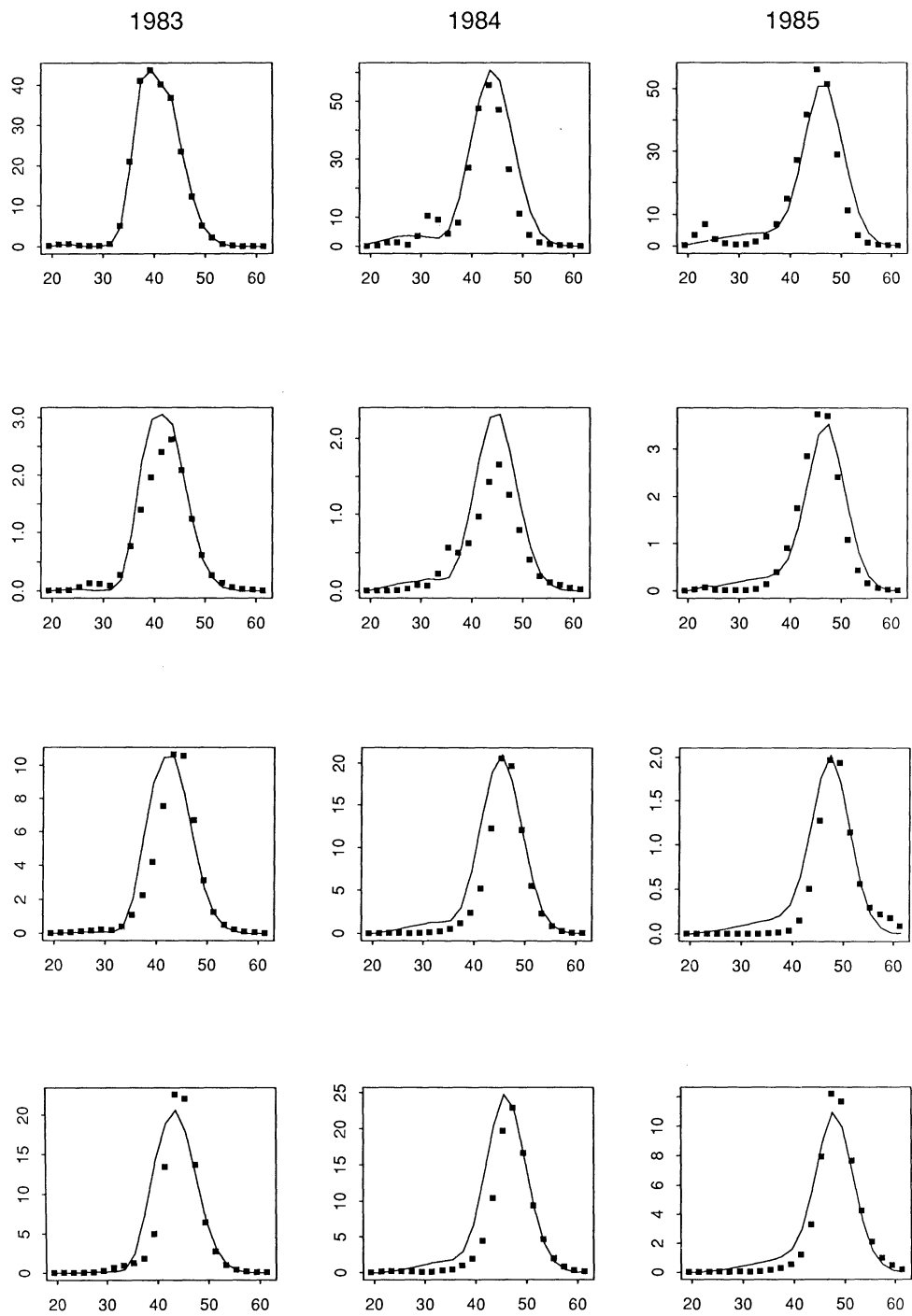


Figure 1. Walleye pollock catch in number (millions) at length (cm) contrasted with estimates from a Kalman filter likelihood (—) based on the full model representing both state and observation error and incorporating auxiliary information. Parameter estimates correspond to those given in column 1 of Table 1. Observations are presented quarterly (top to bottom) over a 3-year period starting with the first quarter of 1983 at the top left.

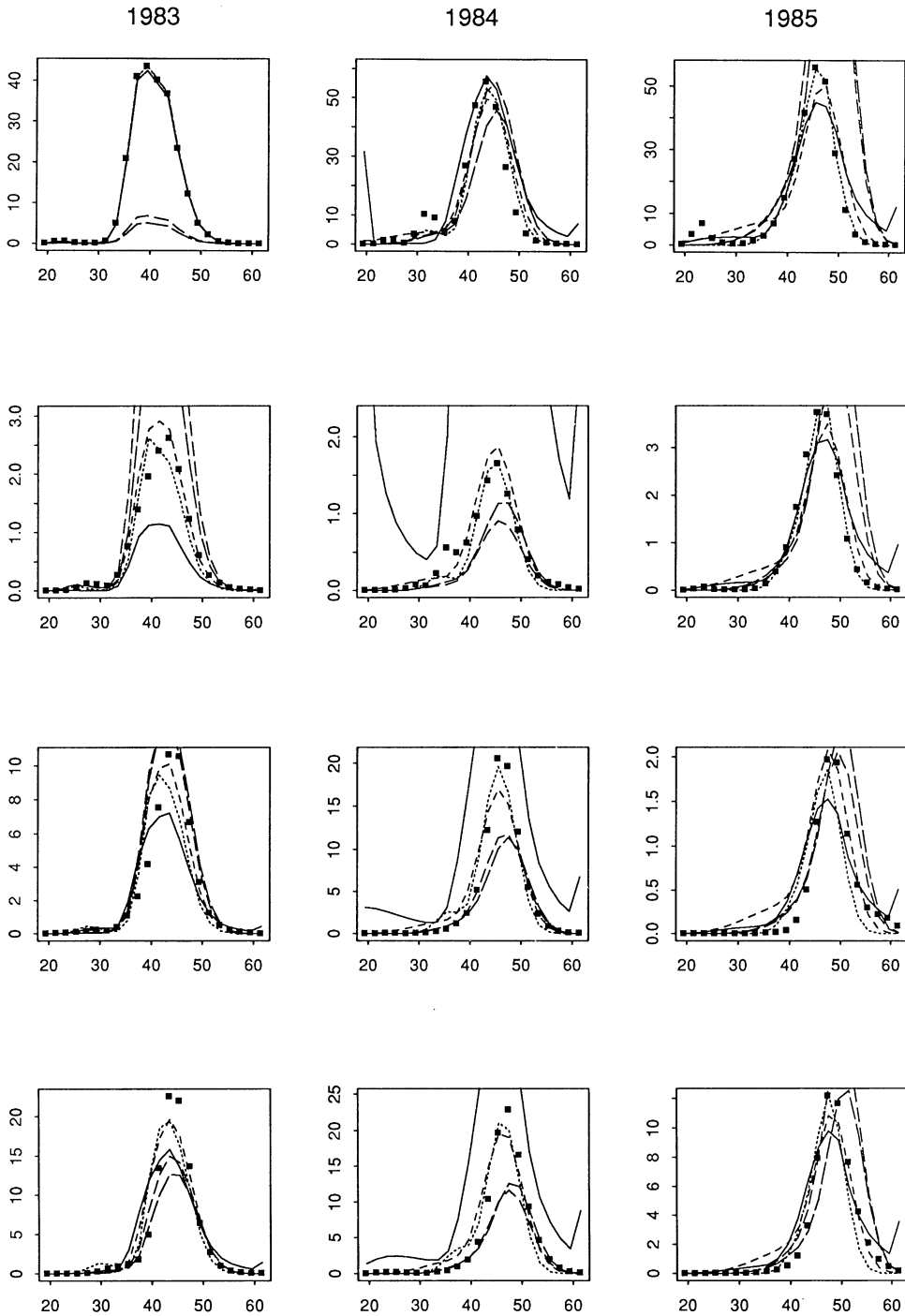


Figure 2. Walleye pollock catch in number (millions) at length (cm) contrasted with the other five estimates presented in Table 1, and shown on the same scale as that used in Figure 1. The estimates are based on the full model without auxiliary information (— · —), the model including observation error only with auxiliary information (— — —) and without (—), and the model including state error only with auxiliary information (— — —) and without (— · · —).

used in Figure 1. Differences between applications, shown as with and without auxiliary information, are due to the inclusion of the three survey observations. The residual sum of squares shown for the application using state error only is perhaps better interpreted as the difference between the estimated and predicted population abundance at length since the observations are assumed to be taken without error and are used in directly updating the abundance predictions. A dramatic reduction in the residual sums of squares is seen for the full model. This is due in part to the differential weighting caused by the process variation and due in part to a lack of fit in the other applications.

The observation and full error structure models without auxiliary information provide the best fit among the remaining models shown in Figure 2. The observation error model with auxiliary information gives a noticeably poorer fit, at the scale shown, in the second year in part because of changes in the magnitude of the catch. The difference of 3 to 4 million in numbers shown at time period 6 would probably go unnoticed in the graph for time period 5, emphasizing the importance of the relative weighting scheme used to modify the residuals for the fit. Note also the aberrant build-up of estimated catch in the largest length class due to the greater diffusion in growth (noted in the large value of the estimated β) attained under this model. The applications that use models representing state error only, though they may capture the recruitment process well, provide an erratic fit. This results from alternatively compensating for overestimating, then underestimating, the catch. The instability in the abundance estimate is propagated throughout the time series. Modifications to the existing models could be explored to improve the fits shown in Figure 2 (for example, by giving more, or less, weight to the auxiliary information or by providing a more sophisticated state process that uses information from the previous two or three time steps rather than just one), but the motivation here was to provide a contrast among several simple alternatives.

The influence of the auxiliary information on each set of estimates is apparent in the magnitude of the estimated initial abundances and recruitment values, and in the estimated quarterly instantaneous fishing mortality rates. This finding is consistent with the work of Deriso et al. (1985) for catch-at-age analysis.

5. Discussion

One advantage of the Kalman filter approach to modeling a dynamic age- or length-structured population is that process stochasticity can be dealt with separately from observation error in the formulation of the likelihood. This provides, for example, an explicit measure of the error in the process (population) estimates, which can be useful for assessing risk in management. A potential drawback is that the structure of the state and observation stochasticity must be specified, and this structure may be ill-defined, poorly understood, or in other ways difficult to model. In the application presented in this paper, for example, the state and observation stochasticity were modeled using multivariate normal approximations of multinomial processes. As a result, the variance in these processes at each time step is dependent on the level of the process at the previous time step. Since the variances are based on the prior means, the forecasts, using the model, are conditioned on a known value. This approach has been used extensively in modeling observation stochasticity, as has been discussed by West and Harrison (1989). In the formulation presented here, this method of modeling the variance is also used in the representation of state variation. The problem of an ill-defined system appears to be overcome by having auxiliary information available to stabilize the estimates of the trend in the series. Stochastic changes in the variance might also be incorporated (West and Harrison, 1989) as could correlation between the state and observation stochasticity. In the application discussed above, the

model was kept simple by letting the variance be conditioned on the prior mean and by assuming that the variances were uncorrelated between the state and the observation.

The specification of the initial variance–covariance matrix raises several interesting points regarding the modeling state stochasticity. When the *a priori* variance is known, as would be the case when variation in population size is independently estimated from a series of surveys, for example, then the estimated variance may be used for the initial variance in the Kalman formulation. When the variance is unknown but replicate realizations of the process are available, then the initial variance may be estimated as part of the procedure (Shumway, Olsen, and Levy, 1981). If the variance is unknown, and if it is not reasonable to obtain replicate realizations of the process, then a diffuse initial variance may better represent the initial information level. In that case, using one of the approaches discussed by Ansley and Kohn (1985) may be appropriate. For the pollock data discussed above, an application of the estimation procedure was made using a large positive initial variance. The result, as one might expect, was that the updated initial population estimate was based solely on the initial catch observation, leaving the initial population parameter estimates arbitrarily defined. Because interest focused on the initial population size estimate in the example discussed in the application, the initial variance was set equal to 0. This forced the initial estimate to be more consistent with the rest of the series.

Recruitment, the influx of new individuals into the population, was dealt with here as a series of points to be estimated. In many applications, recruitment can be viewed as a function of total population density (Beverton and Holt, 1957; Ricker, 1975) with the effect subject to time lags. The approach presented here is flexible enough to include a recruitment function if such a relationship is desired. For example, various linear approaches could be taken using autoregressive integrated moving average (ARIMA) models relating recruitment to previous recruitment levels or population sizes. Nonlinear relationships can also be incorporated by using an extended Kalman filter approach to approximate such processes [see Gelb et al. (1988) for a discussion of such approaches]. Research based on models of recruitment is of great interest to resource managers who often rely on such relationships for the assessment of risk under various harvest strategies.

In the application discussed above, estimates of initial population size, fishing mortality over time, selectivity by length, and von Bertalanffy growth were obtained. Given these estimates, one can obtain population abundance estimates for the rest of the time period. Abundance estimates are useful for monitoring population trends, examining environmental and density-dependent effects, and formulating management objectives for maintaining the viability of the stock. One way to obtain population abundance estimates is to use the filtered state variable estimates. The filtered estimates, as indicated by equations (5) and (6), are computed as intermediate variables by the routine and thus are readily available as output from the algorithm. However, these optimum linear filtered estimates use only the data up to and including the time of the estimate of interest. To obtain estimates at any time step that use all of the data, an optimal linear smoother may be more appropriate. The optimal linear smoothed estimates are the minimum variance unbiased estimates of the states as well as the maximum likelihood estimates (Fraser and Potter, 1969). The smoothed estimates may be arrived at by first obtaining the forward filtered estimates (the ones discussed above) and then applying a backward filter to obtain the result. A derivation of the optimal linear smoother from two optimal linear filters is given by Fraser and Potter (1969) and discussed generally by Gelb et al. (1988). Mendelssohn (1988) found that the smoothed and filtered estimates were often indistinguishable although there were some significant differences. If one is interested only in the general trend exhibited by the population, then the filtered estimates alone may be adequate. Also, if having a good estimate of the most recent year's population abundance is important, say for management

purposes, then again the filtered final estimate is adequate because it already is based on the entire data series. However, if it becomes important to obtain precise estimates of past abundances (for establishing a relationship between environmental conditions and stock size, for example) then the smoothed estimates are the ones to use. The information needed to run the backward filter is present in the intermediate variables used in formulating the forward filter and thus may be computed by extending the algorithm to cover the reverse process.

The likelihood function used here resulted in a fitting criterion based on the squared differences in the catches weighted by an estimate of their variance. Others [Doubleday (1976), for example] use a fitting criterion based on the logarithm of catch. Such transformations of the data can be useful in stabilizing the performance of fitting algorithms and may serve to change the characteristics of the error distribution, often making the distribution appear more Gaussian. The derivation presented here, however, presents the characteristics of the distribution as they result from a multinomial process, leading to a Gaussian approximation. The variance component is developed in the same manner, resulting in error terms that are proportional to the mean. Thus the multiplicative nature of the error is dealt with in a slightly different, but equally justifiable, manner. In the application of the algorithm discussed above, the parameters were constrained to be positive, but the population abundances and associated catches were numerically large, reducing the likelihood of negative estimates. This may not hold true for smaller populations and in that instance the Gaussian assumptions are less likely to hold and some other approach may have to be taken. A discussion of non-Gaussian filters has recently been given by Kitagawa (1987).

Respecifying the state and observation parameters into a set of conventional fisheries research parameters proved useful in the application discussed above. This respecification satisfied several objectives, as outlined in Section 3.3. The primary objective, that of constraining the parameters to an estimable parameter space, was nicely met. In the pollock example, the respecification enabled the 22×22 dimensioned matrix of length transition probabilities $P_{l,l'}$ to be represented simply by a three-parameter model for the entire time series. However, respecification of the parameters in this manner also brings with it some problems, as any examination of the fisheries research literature will indicate. For example, natural mortality was assumed to be known. Incorrect specification of this parameter can affect the estimates of fishing mortality and consequently affect the estimates of initial population size and recruitment. [Consider the work of Lapointe, Peterman, and MacCall (1989) for a recent discussion of this problem with respect to virtual population analysis.] Therefore in this, as in any, application one should endeavor to understand the consequences of using conventional approaches.

6. Conclusion

The use of the Kalman filter state-space formulation to describe a dynamic population and an associated set of observations (the catch or harvest) facilitates the incorporation of system stochasticity as well as observation error. Within this structure, parameter specification in terms of models commonly used in fisheries and population dynamics serves to logically constrain the parameter space as well as to relate parameter estimates to those traditionally obtained by other methods. Explicitly modeling state stochasticity in addition to the observation error allows the researcher greater flexibility in formulating likelihood equations, possibly avoiding less specific procedures that deal with variance trade-offs in the residuals. Modeling the variance as a conditionally updated structure through time is a method for dealing with variance that is somewhere between the traditional weighted least squares approach, where the weights are simultaneously estimated with the parameters,

and the ordinary least squares approach, which assumes that the variance is constant over the range of population values. Model structure and auxiliary information act to define the compromise between these two approaches.

ACKNOWLEDGEMENTS

I thank Dr Peter Guttorp and Dr Han-Lin Lai for many valuable discussions related to this work. I thank Dr Ana Parma, Dr Bernard Megrey, Dr Roy Mendelssohn, and Dr Jon Schnute for reviewing this manuscript and providing many helpful suggestions.

RÉSUMÉ

Une représentation spatiale d'une population structurée en classes selon des critères commerciaux est décrite et un filtre de Kalman est utilisé pour développer l'équation de vraisemblance conditionnelle nécessaire pour estimer les paramètres du système sous-jacent. L'état du système est caractérisé en utilisant la théorie conventionnelle des pêches, des collectes commerciales représentant des observations faites sur la population. La structure de vraisemblance conditionnelle incorporée dans le filtre de Kalman facilite à la fois l'introduction de l'aléatoire du système aussi bien que d'un terme représentant l'erreur d'observation, dans l'équation de vraisemblance totale. Dans ce cadre, une approche du maximum de vraisemblance est utilisée pour estimer les paramètres de la population en tenant compte des deux sources d'erreur.

REFERENCES

- Ansley, C. F. and Kohn, R. (1985). Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions. *Annals of Statistics* **13**, 1286–1316.
- Berger, J., Murai, S., Nelson, R., and Wall, J. (1986). Summaries of provisional foreign and joint-venture groundfish catches (metric tons) in the Northeast Pacific Ocean and Bering Sea, 1985. Seattle, Washington: National Marine Fisheries Service, Northwest and Alaska Fisheries Center.
- Berger, J., Nelson, R., and Wall, J. (1985). Summaries of provisional foreign and joint-venture groundfish catches (metric tons) in the Northeast Pacific Ocean and Bering Sea, 1984. Seattle, Washington: National Marine Fisheries Service, Northwest and Alaska Fisheries Center.
- von Bertalanffy, L. (1934). Untersuchungen über die Gesetzmäßigkeit des Wachstums. *I. Roux' Archiv* **131**, 613.
- Beverton, R. J. H. and Holt, S. J. (1957). *On the Dynamics of Exploited Fish Populations*. Fishery Investigations Series II, Volume XIX. London: Her Majesty's Stationery Office.
- Collie, J. S. and Sissenwine, M. P. (1983). Estimating population size from relative abundance data measured with error. *Canadian Journal of Fisheries and Aquatic Sciences* **40**, 1871–1879.
- DeLury, D. B. (1947). On the estimation of biological populations. *Biometrics* **3**, 145–167.
- Deriso, R. B., Quinn, T. J. II, and Neal, P. R. (1985). Catch-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 815–824.
- Doubleday, W. G. (1976). A least squares approach to analyzing catch at age data. *Research Bulletin of the International Commission for the Northwest Atlantic Fisheries* **12**, 69–81.
- Fraser, D. C. and Potter, J. E. (1969). The optimum linear smoother as a combination of two optimum linear filters. *IEEE Transactions on Automatic Control* **7**, 387–390.
- Gelb, A., Kasper, J. F., Jr., Nash, R. A., Jr., Price, C. F., and Sutherland, A. A., Jr. (1988). *Applied Optimal Estimation*. Cambridge, Massachusetts: M.I.T. Press.
- Getz, W. M. and Haight, R. G. (1989). *Population Harvesting: Demographic Models of Fish, Forest, and Animal Resources*. Princeton, New Jersey: Princeton University Press.
- Gutreuter, S. J. and Anderson, R. O. (1985). Importance of body size to the recruitment process in largemouth bass populations. *Transactions of the American Fisheries Society* **114**, 317–327.
- Hughes, S. E. and Hirschhorn, G. (1979). Biology of walleye pollock, *Theragra chalcogramma*, in the western Gulf of Alaska, 1973–75. *Fishery Bulletin* **77**, 263–274.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the American Society of Mechanical Engineers, Series D* **82**, 35–45.
- Kalman, R. E. and Bucy, R. (1961). New results in linear filtering and prediction. *Transactions of the American Society of Mechanical Engineers, Series D* **83**, 95–108.
- Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series. *Journal of the American Statistical Association* **82**, 1032–1041.

- Lapointe, M. F., Peterman, R. M., and MacCall, A. D. (1989). Trends in fishing mortality rate along with errors in natural mortality rate can cause spurious time trends in fish stock abundances estimated by virtual population analysis (VPA). *Canadian Journal of Fisheries and Aquatic Sciences* **46**, 2129–2139.
- Lipster, R. S. and Shirayayev, A. N. (1978). *Statistics of Random Processes II*. New York: Springer-Verlag.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* **11**, 431–441.
- Megrey, B. A. (1989). Gulf of Alaska walleye pollock: Population assessment and status of the resource as estimated in 1988. In *Condition of Groundfish Resources of the Gulf of Alaska in 1988*, T. K. Wilderbuer (ed.), 1–54. NOAA Technical Report F/NWC-165. Washington, D.C.: U.S. Department of Commerce.
- Meinhold, R. J. and Singpurwalla, N. D. (1983). Understanding the Kalman filter. *The American Statistician* **37**, 123–127.
- Mendelsohn, R. (1988). Some problems in estimating population sizes from catch-at-age data. *Fishery Bulletin* **86**, 617–630.
- Nelson, R., Wall, J., Berger, J., and Gibbs, B. (1984). Summaries of provisional foreign and joint-venture groundfish catches (metric tons) in the Northeast Pacific Ocean and Bering Sea, 1983. Seattle, Washington: National Marine Fisheries Service, Northwest and Alaska Fisheries Center.
- Nunnallee, E. P. and Williamson, N. J. (1989). Results of acoustic/midwater trawl surveys of walleye pollock in Shelikof Strait, Alaska, in 1988. In *Condition of Groundfish Resources of the Gulf of Alaska in 1988*, T. K. Wilderbuer (ed.), 225–242. NOAA Technical Report F/NWC-165. Washington, D.C.: United States Department of Commerce.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edition. New York: Wiley.
- Ratkowsky, D. A. (1986). Statistical properties of alternative parameterizations of the von Bertalanffy growth curve. *Canadian Journal of Fisheries and Aquatic Sciences* **43**, 742–747.
- Ricker, W. E. (1975). *Computations and Interpretation of Biological Statistics of Fish Populations*. Bulletin 191. Ottawa, Canada: Department of the Environment, Fisheries and Marine Service.
- Sainsbury, K. R. (1982). Population dynamics and fishery management of the paua, *Haliotis iris* II. Dynamics and management as examined using a size class population model. *New Zealand Journal of Marine and Freshwater Research* **16**, 163–173.
- Schnute, J. (1987). A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences* **44**, 924–940.
- Schnute, J. T., Richards, L. J., and Cass, A. J. (1989a). Fish growth: Investigations based on a size-structured model. *Canadian Journal of Fisheries and Aquatic Sciences* **46**, 730–742.
- Schnute, J. T., Richards, L. J., and Cass, A. J. (1989b). Fish survival and recruitment: Investigations based on a size-structured model. *Canadian Journal of Fisheries and Aquatic Sciences* **46**, 743–769.
- Shumway, R. H., Olsen, D. E., and Levy, L. J. (1981). Estimation and tests of hypotheses for the initial mean and covariance in the Kalman filter model. *Communications in Statistics—Theory and Methods* **10**, 1625–1641.
- Sullivan, P. J., Lai, H. L., and Gallucci, V. F. (1989). A catch-at-length analysis that incorporates a stochastic model of growth. *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 184–198.
- West, M. and Harrison, J. (1989). *Bayesian Forecasting and Dynamic Models*. New York: Springer-Verlag.
- Zehnwirth, B. (1988). A generalization of the Kalman filter for models with state-dependent observation variance. *Journal of the American Statistical Association* **83**, 164–167.

Received November 1989; revised July 1990 and January 1991; accepted April 1991.

APPENDIX

A recursive algorithm for computing the components of the likelihood function, equation (12), is provided here.

Given the observations \mathbf{C}_τ and the state-space parameters μ_τ , \mathbf{P} , \mathbf{S}_τ , and \mathbf{R}_τ , where $\tau \in \{0, \dots, t\}$, and given the parameters \mathbf{N}_0 and $\Sigma^{\mathbf{N}_0}$, the current state of the system, starting with $\tau = 1$, is

$$(\mathbf{N}_{\tau-1} \mid \mathbf{C}^{\tau-1}), \quad (\text{A.1a})$$

$$\Sigma^{\mathbf{N}_{\tau-1} \mid \mathbf{C}^{\tau-1}}, \quad (\text{A.1b})$$

with the variance–covariance matrices either given or computed as

$$\Sigma^w_\tau = \mathbf{P}\mathbf{S}_{\tau-1}(\mathbf{N}_{\tau-1} | \mathbf{C}^{\tau-1})(\mathbf{I} - \mathbf{P})^T, \quad (\text{A.2a})$$

$$\Sigma^v_\tau = \mu_\tau \mathbf{P}\mathbf{S}_{\tau-1}(\mathbf{N}_{\tau-1} | \mathbf{C}^{\tau-1})(\mathbf{I} - \mu_\tau)^T, \quad (\text{A.2b})$$

with \mathbf{I} defined as the identity matrix. The one-step-ahead predictions are then

$$(\mathbf{N}_\tau | \mathbf{C}^{\tau-1}) = \mathbf{P}\mathbf{S}_{\tau-1}(\mathbf{N}_{\tau-1} | \mathbf{C}^{\tau-1}) + \mathbf{R}_\tau, \quad (\text{A.3a})$$

$$\Sigma^{\mathbf{N}_\tau | \mathbf{C}^{\tau-1}} = \mathbf{P}\mathbf{S}_{\tau-1} \Sigma^{\mathbf{N}_{\tau-1} | \mathbf{C}^{\tau-1}} (\mathbf{P}\mathbf{S}_{\tau-1})^T + \Sigma^w_\tau, \quad (\text{A.3b})$$

with the corresponding one-step-ahead observation predictions

$$\mathbf{C}^*_\tau = \mu_\tau (\mathbf{N}_\tau | \mathbf{C}^{\tau-1}), \quad (\text{A.4a})$$

$$\Sigma^{\mathbf{C}_\tau} = \mu_\tau \Sigma^{\mathbf{N}_\tau | \mathbf{C}^{\tau-1}} \mu_\tau^T + \Sigma^v_\tau. \quad (\text{A.4b})$$

The updated estimates for the next iteration (i.e., $\tau = 2, 3, 4, \dots, t$) are then computed as

$$(\mathbf{N}_\tau | \mathbf{C}^\tau) = (\mathbf{N}_\tau | \mathbf{C}^{\tau-1}) + \Sigma^{\mathbf{N}_\tau | \mathbf{C}^{\tau-1}} \mu_\tau^T [\Sigma^{\mathbf{C}_\tau}]^{-1} (\mathbf{C}_\tau - \mathbf{C}^*_\tau), \quad (\text{A.5a})$$

$$\Sigma^{\mathbf{N}_\tau | \mathbf{C}^\tau} = \Sigma^{\mathbf{N}_\tau | \mathbf{C}^{\tau-1}} - \Sigma^{\mathbf{N}_\tau | \mathbf{C}^{\tau-1}} \mu_\tau^T [\Sigma^{\mathbf{C}_\tau}]^{-1} \mu_\tau \Sigma^{\mathbf{N}_\tau | \mathbf{C}^{\tau-1}}. \quad (\text{A.5b})$$

The process is repeated until all of the necessary components of the likelihood are computed—that is, $\{\mathbf{C}^*_\tau, \Sigma^{\mathbf{C}_\tau}; \tau \in \{0, \dots, t\}\}$.

Note that the components of the weighted residual $(\mathbf{C}_\tau - \mathbf{C}^*_\tau) [\Sigma^{\mathbf{C}_\tau}]^{-1/2}$, namely, $(\mathbf{C}_\tau - \mathbf{C}^*_\tau)$ and $[\Sigma^{\mathbf{C}_\tau}]^{-1/2}$, can be more useful and easier to manipulate than the form of the components given in equation (A.4). They can be more useful since the array of weighted residuals can be used in nonlinear maximization algorithms such as those based on the technique of Marquardt (1963). They can be easier to manipulate since variance–covariance matrices in general can be represented as the product of a complementary set of upper- and lower-triangular matrices that are more efficiently stored during computation and are easier to use in nested computational algorithms.