Fisheries and Oceans Science | Pêches et Océans Sciences

C S A S | S C É S

Canadian Stock Assessment Secretariat | Secrétariat canadien pour l'évaluation des stocks

Research Document 2000/159 | Document de recherche 2000/159

Not to be cited without permission of the authors [1]

Ne pas citer sans autorisation des auteurs [1]

A Preliminary Review of A New Model Based On Test Fishing Data Analysis to Measure Abundance Of Returning Chum Stocks To The Fraser River

W.J. Gazey[1] and R.V. Palermo[2]

[1]W.J. Gazey Research
1214 Camas Court
Victoria, BC
V8X 4R1

[2]Fisheries and Oceans Canada
100 Annacis Parkway, Unit 3
Delta, B.C
V3M 6A2

ISSN 1480-4883
Ottawa, 2000

ABSTRACT

The test-fishery has operated at Albion on the Fraser River since 1978 to provide the means for an index of chum salmon abundance (escapement) within a season. Recent degradation of the accuracy and consistency of escapement estimates has seriously undermined the potential to evaluate clockwork management for the Fraser River chum salmon (CSAS Res. Doc. 99/169, Ryall et al.). To address this problem the cumulative catch-per-unit-effort (CPUE) was calculated to account for saturation, depletion in the second set and interpolation for missing sampling days. In addition, the test-fishery data were cast into a Bayesian framework that incorporated preseason knowledge of run size and migration timing, within season information on migration timing and a predictive regression to calibrate run size to the historical record. Based on a retrospective analysis of 1979-1998 data, the Bayesian procedure was judged to be superior to the classical test fisheries approach of using a simple predictive regression of cumulative CPUE on run size. However, the predictive ability of the either model was seriously compromised by the reliability of escapement enumeration (end of season minimal residual standard deviation of prediction was 256 thousand fish).

RÉSUMÉ

La pêche expérimentale effectuée dans le fleuve Fraser à Albion depuis 1978 sert à établir un indice d'abondance du saumon kéta (échappée) dans la saison. Le fléchissement récent de la précision et de la cohérence des estimations de l'échappée a fortement compromis la possibilité d'évaluer la gestion précise du saumon kéta du Fraser (document SCÉS 99/169, Ryall *et al.*). Pour résoudre ce problème, on a calculé les captures par unité d'effort (CPUE) cumulatives de sorte à tenir compte de l'excès de données, de la carence de la deuxième série de données et de l'interpolation pour les jours d'échantillonnage manquants. On a aussi fait une analyse bayesienne des données de pêche expérimentale incluant les renseignements obtenus avant la saison sur la taille des remontes et le moment de la migration, les renseignements sur le moment de la migration dans la saison et une régression prévisionnelle pour étalonner la taille des remontes par rapport aux données historiques. Selon une analyse rétrospective des données pour la période 1979-1998, on a déterminé que l'analyse bayesienne donnait de meilleurs résultats que l'approche classique des pêches expérimentales, qui fait appel à une simple régression prévisionnelle des CPUE cumulatives d'après la taille des remontes. Le niveau de fiabilité des dénombrements d'échappées a toutefois sérieusement compromis la capacité de prévision des deux modèles (écart-type résiduel minimal à la fin de la saison).

1.0 INTRODUCTION

Beginning in 1987, a stepped harvest rate management strategy (clockwork strategy) has been applied to the in-river commercial harvest of Fraser River chum salmon. A directed harvest is not allowed until predictive analysis of the cumulative test fishing catches at Albion indicate that the estimated abundance of chum salmon exceeds the gross escapement goal as determined from stock recruitment analysis.

The Fraser River chum salmon run was divided into early and late timed components (Oct. 1 – to Nov. 12 and Nov. 13 – Dec. 20 respectively) based on historical catches in a test fishery that used to run at Cottonwood (Palmer 1972). A linear model based on regression analysis using estimated escapement (dependent variable) and the test fishery's cumulative catch per unit effort (thousand fathom minutes) (independent variable) for weekly time blocks was developed to assess the estimated returns to each timing component with regard to satisfying minimum gross escapement.

Table 1. Clockwork harvest rate schedule for the Fraser River.

RUN TIMING	TERMINAL RUN SIZE	MANAGEMENT	POTENTIAL CATCH
Early Oct. 1 - Nov. 12	445,000 - 550,000	Minimum Gross Escapement Goal* = 410,000; one opening not to exceed 10% harvest rate	35,000 – 55,000
	>550,000	Set harvest rate at 15%	82,500+
Late Nov. 13 - Dec. 20	370,000 - 500,000	Minimum Escapement Goal* = 350,000; one opening not to exceed 10% harvest rate	35,000 – 50,000
	>500,000	Set harvest rate at 15%	75,000

* (Gross Escapement Goal = Net Escapement + Test Catch + Native Harvest)

Chum salmon gross escapement to the Fraser River has increased since 1987, meeting the rebuilding goal of the Clockwork strategy. In recent years however, the predictive linear model has been unable to estimate the run size within bounds needed to make fishery management decisions based on the Clockwork strategy. There is a growing discrepancy between the run size predicted from the Albion test fishery catches and the escapement estimated from spawning ground counts. The decision to allow a fishery depends on the run size prediction from the Albion test fishery and therefore, the potential to miss fishing opportunities because of inaccuracies in the prediction have become quite probable.

In addition to the problems with the differences between the overall escapement estimate and the Albion run size prediction, the assumptions concerning the two timing components have further complicated management of the in-river chum fishery. Catches in the Albion test fishery have not demonstrated the same bimodal distribution pattern observed in the historical catches in the Cottonwood test fishery (Palmer 1972). The clockwork model is based on two distinct timing runs and the gross escapement goal must be reached for each component before fishing can occur. The run size "count" is reset on Nov. 13 and must then reach the late run minimum escapement goal before fishing can take place even if the early component has vastly exceeded its goal. In recent years there has been a precipitous drop off in chum catches in the test fishery around 12 November (M. Sullivan pers. com. and Fig.4). With the restrictions imposed on the fishery due to coho and steelhead conservation concerns (essentially no fishing in September-October), this has meant that the clockwork criteria for a fishery cannot be met.

CSAS Res. Doc. 99/169 (Ryall *et. al*.) reviewed the status of clockwork chum stocks, including the Fraser River, and the Clockwork Management Strategy. DFO's Pacific Scientific Advice Review Committee (PSARC) recommended a review of the terminal in-season test fishery model with an objective of finding a more suitable protocol if the current model is deemed a failure. It was noted that recent degradation of the accuracy and consistency of escapement estimates is seriously undermining the potential to evaluate Clockwork management. The lack of reliable escapement data is also hindering development of new assessment methodology.

The purpose of this paper is to cast the inseason estimation problem into a Bayesian calculus such that all available information is used, a posterior probability distribution is available for evaluation of risk and estimates early in the season are consistent with prior expectations. Bayesian techniques were first applied to the inseason estimation of Bristol Bay, Alaska, sockeye salmon abundance by Fried and Hilborn (1988); however, their procedure is restricted to updating a weighted composite estimate of run size from several independent forecasting methods and does not make forecasts of timing. Similarly and more recently, Pella et al. (1998) used Bayesian techniques to update a preseason forecast of chinook salmon abundance index in Southeast Alaska with early catch rates. Conceptually, our approach is more comparable to the time density model to estimate run size and entry timing described by Springborn et al. (1998).

In the text that follows we describe the data collected by the Albion chum test fishery and subsequent calculation of the cumulative catch-per-unit-effort (CPUE) statistic, which is used as an unbiased index of abundance. Next, the time density model for fish passage at Albion is described and parameters of interest and available data are defined. Finally, the predictive posterior model is presented and all the ensuing distributional components are specified.

2.0 METHODS

2.1 Data Sources

The test fishing data was summarized by Farwell (1985) for the years 1963 through 1984 and put into electronic format. Data from 1984 through the present has not been published, however, the Fraser River test fishing data has been accumulated in spreadsheets and provided to us by Fisheries Management staff (M. Sullivan, Fraser River Management Biologist). Fraser River chum salmon escapement data was accumulated from numerous sources (Farwell 1993, 1995, 1995, Farwell *et. al.* 1992, and Ryall *et. al.* 1999, and Thomas 1996).

2.2 Test Fishing Data Collection

Originally, two test sites were fished in the lower Fraser River, one off the north shore of Tilbury Island and the other near Albion, BC (Figures 1 and 2). The Tilbury Island site, locally known as the "Cottonwood drift", was selected as the best location to assess chum abundance after an exploratory program in 1961 and 1962 (Palmer 1972). This site was fished until 1987 and had been fished annually since 1963. In 1978, other test fishing sites were investigated because of concerns that changes in river morphology and interference from shipping traffic were affecting the catches at Cottonwood. As a result of this investigation, the Albion location was selected and has been fished annually since 1979.

At each test-fishing site a standard gill net fishes during a standard tide stage. Two test sets are conducted on a daily basis throughout the duration of the chum migrating period (October - December).

The gill nets are of 6 3/4" (171 mm) nylon mesh (measured between opposite knots of a stretched individual mesh) with dimensions of 150 fathoms (274 m) long and 60 meshes deep. The net was hung in a three to one ratio, where each fathom or meter of corkline supported three fathoms or meters of net. Size code f33 nylon twine is used and was coloured as closely as possible to colour code NG 12 (light green) of the Nichimo Company Limited (Farwell 1985). The fishing procedure is described in detail in Farwell (1985) and is briefly outlined below.

“The fishing procedures were standardized as much as possible. The nets were set approximately perpendicular to the shore of the river and allowed to drift with the prevailing currents. At each site, the first set accused while the tide was ebbing and was timed to permit the set to be completed before the slack period of the lowest low water tide of the day. The second set began immediately after the tide change. Each set consisted of three time periods: setting time, drift time and picking time. Setting time duration is variable but usually around five minutes and began when the first section of the net was placed in the water and ended when the last section of the net was in the water. Drift time duration was standardized at 30 minutes beginning as soon as the net was fully in the water. Picking time duration was dependent on the number of fish caught, began immediately following the expiration of the 30 minute drift and ended when the net was fully removed from the water.”

The exact set position and the distance covered in each of the sets was dependent on river discharge and the height of the two tides on either side of the low water slack.

Information recorded for each set included the time of set start, time when entire net was in the water, time when net picking began, and time when the entire net was retrieved. Total catch by species was also recorded.

2.3 Catch per Unit Effort Calculation

To determine a relative daily index of abundance, the catch in each set was related to the effort in that set. A calculated catch per unit of effort (CPUE), was calculated as follows (from Farwell, 1985): The catch (*C*) in each set was the total number of each species removed from the net. The amount of effort (*E*) for each set was calculated from the length of time the net was in the water (T_f), and the length of the net (*L*) and was recorded in thousand fathom minute (TFM) units:

$$(1) \qquad E = T_f \frac{L}{1000}$$

where the fishing time is T_f in minutes and was obtained by summing of the setting (T_s), drift (T_d) and picking times (T_p) as:

$$(2) \qquad T_f = \frac{T_s}{2} + T_d + \frac{T_p}{2}$$

The determination of CPUE for each species was then:

$$(3) \qquad CPUE = \frac{C}{E}$$

and was performed for each set separately.

The computation of cumulative daily CPUE had to depart from the usual methodologies (i.e., Farwell 1985) in order to account for the effects of saturation, depletion of catch in the second set and interpolation for missing sampling days. We first examined the plot of catch for all species versus CPUE for saturation effects, which causes a curvilinear relationship. We subsequently assumed the following saturation power function:

$$(4) \qquad C \propto (CPUE)^v$$

The saturation exponent, $v = 1.16$ ($R^2 = 0.98$) was then estimated using functional regression on the log-transformed values (Ricker 1973). The chum salmon CPUE for all sets were then adjusted by this exponent. Figure 3 illustrates the raw and adjusted relationships.

The adjusted CPUE for two sets taken on the same day were then combined by the average weighted by fishing time, i.e.,

$$(5) \qquad A = \frac{(CPUE_1') \cdot E_1 + (CPUE_2') \cdot E_2}{E_1 + E_2}$$

where *A* is the adjusted mean CPUE for a sampling day and $CPUE_d'$ is the adjusted CPUE for set $d = 1$ or 2.

On some days, only one set was taken. However, there turned out to be approximately a 9% reduction in CPUE in the second set on days that two sets were taken, presumably because of depletion. We therefore regressed the adjusted CPUE for the first set on the adjusted daily mean (A of equation 5) to obtain the predictive relationship ($R^2 = 0.85$):

(6) $$A = 0.4982 + 0.9267 \cdot CPUE_1'$$

which was used to estimate the CPUE on days with only a single set.

Finally, the daily adjusted CPUE were summed over the season to obtain the cumulative CPUE for each sampled day. We assumed that the start and the end of the run were sampled and any missing days in the cumulative CPUE were estimated by linear interpolation.

3.0 TIME DENSITY MODEL AND NOTATION

The migration of fish past some fixed geographical reference point has long been characterised in salmon management as a time density probability function (Mundy 1979). Our timing reference point is the test fishing site at Albion on the Fraser River and the mean time density function should be a normal distribution according to the central limit theorem. The characterisation of migration timing of salmon as a normal distribution is very orthodox (e.g., Cave and Gazey 1994). A comparable approach is used by Springborn et al. (1998) with an inverted exponential logistical model, which has properties similar to the cumulative normal distribution. In practice, there is a great deal of yearly variation in the run distribution as illustrated by plots of CPUE for 1979 to 1999 (Fig. 4). While a more complex model could be used to account for skewness (e.g., Schnute and Sibert 1983), our experience is that the shape of the distribution is exceedingly difficult to determine within season. In other words, the symmetrical normal distribution seems to be a good approximation in the absence of *a priori* information.

Accordingly, the time density function was assumed normal

(7) $$g(t|T,S) = \frac{1}{S\sqrt{2\Pi}} e^{-\frac{1}{2}\left(\frac{t-T}{S}\right)^2}$$

where T is the mean and S is the standard deviation. The probability that a fish has passed the reference point by time x is then the cumulative distribution:

(8) $$G(x|T,S) = \int_{-\infty}^{x} g(t|T,S)dt$$

These time density functions and parameters are illustrated in Fig. 5. The number of fish that have passed the reference point is then $N \cdot G(t|S,T)$, where N is the run size or escapement at Albion. The historical time density information can be reduced to the mean timing and variance because these statistics are sufficient for the normal distribution assumed for migration timing. Therefore,

(9) $$T_y = E[t] = \sum_t tp_{ty} \text{ and}$$

(10) $$S_y^2 = E[t^2] - (E[t])^2 = \sum_t t^2 p_{ty} - T_y^2$$

where

$$p_{ty} = \frac{B_{t+1,y} - B_{ty}}{B_{n_y y}}$$

and B_{ty} is the adjusted cumulative CPU sampled on day t in year y and n_y is the number of sampling days in year y. Thus, the data from previous years (see Table 2) consists of three observed vectors of the parameters to be predicted:

$\mathbf{N_p} = \{N_1, N_2, \ldots, N_n\}$ - escapement
$\mathbf{T_p} = \{T_1, T_2, \ldots T_n\}$ - mean timing
$\mathbf{S_p} = \{S_1, S_2, \ldots S_n\}$ - standard deviation or spread;

and a vector of cumulative CPU at time t:

$\mathbf{B_{pt}} = \{B_{t1}, B_{t2}, \ldots B_{tn}\}$ - cumulative CPU at time $t = 1, 2, 3, \ldots, n_y$

The data for the year to be predicted (current year) consists solely of the cumulative CPUE at Albion up to time t, i.e., $\mathbf{B}_t = \{B_1, B_2, \ldots, B_t\}$.

4.0 BAYES PREDICTIVE MODEL

4.1 Framework

The objective is to predict for the current year run size (N), timing mean (T) and the standard deviation or spread (S) based on observations (data) made in the current year and pervious years. The desired distribution is called the posterior predictive distribution, posterior because it is conditional on the observed data and predictive because it is a prediction for an observable N, T or S (Gelman et al. 1995):

(11) $$\begin{aligned} f(N|\mathbf{B_t}, \mathbf{X_{pt}}) &= \iint f(N,T,S|\mathbf{B_t}, \mathbf{X_{pt}})\,dTdS \\ f(T|\mathbf{B_t}, \mathbf{X_{pt}}) &= \iint f(N,T,S|\mathbf{B_t}, \mathbf{X_{pt}})\,dNdS \\ f(S|\mathbf{B_t}, \mathbf{X_{pt}}) &= \iint f(N,T,S|\mathbf{B_t}, \mathbf{X_{pt}})\,dNdT \end{aligned}$$

where $\mathbf{X_{pt}} = \{\mathbf{N_p}, \mathbf{T_p}, \mathbf{S_p}, \mathbf{B_{pt}}\}$ is an array defining all data collected in previous years. Note that the prediction for the parameter of interest is obtained by integrating out the other two parameters and that the equations relate to a specific time t during the season.

The joint posterior distribution is partitioned into the sampling and prior distributions using Bayes rule to obtain:

(12) $$f(N,T,S|\mathbf{B}_t,\mathbf{X}_{\mathbf{pt}}) \propto g(\mathbf{B}_{\mathbf{t}}|N,T,S)h(N,T,S|\mathbf{X}_{\mathbf{pt}})$$

The notation follows the convention that the posterior distribution and all direct derivations be denoted by $f(\cdot|\cdot)$, the sampling distribution by $g(\cdot|\cdot)$ and the prior distribution by $h(\cdot|\cdot)$. Note that the likelihood of observing the data in the current year depends on the parameters and not on data from previous years.

The prior distribution can be partitioned as follows:

(13) $$\begin{aligned} h(N,T,S|\mathbf{X}_{\mathbf{pt}}) &= h(N|T,S,\mathbf{X}_{\mathbf{pt}})h(T,S|\mathbf{X}_{\mathbf{pt}}) \\ &\propto h_0(N)h_1(N|T,S,\mathbf{X}_{\mathbf{pt}})h(T,S|\mathbf{X}_{\mathbf{pt}}) \end{aligned}$$

The first line in (13) factors out run size from the time density parameters. The second line specifies an initial prior of run size based on independent preseason information other than that described here (h_0) and a prior of run size calibrated to the time density parameters, the cumulative CPUE and previous years of information (h_1).

Substitution of equation (13) into (12) yields the joint posterior partitioned into four distributions:

(14) $$f(N,T,S|\mathbf{B}_{\mathbf{t}},\mathbf{X}_{\mathbf{pt}}) \propto g(\mathbf{B}_{\mathbf{t}}|N,T,S)h_0(N)h_1(N|T,S,\mathbf{X}_{\mathbf{pt}})h(T,S|\mathbf{X}_{\mathbf{pt}})$$

Note that result (14) only depends on the weak assumption of an independent preseason estimate. The utility of the predictive model depends on the assumptions and procedures specifying these four distributions, which we next explore

4.2 Initial Prior for Run Size

A pre-season estimate of run size (N_0) with an associated coefficient of variation (cv) is usually available. We assume a lognormal distribution and that the cv provides an approximation of the standard deviation of the log-transformed variable. Thus, ignoring constant terms the prior distribution can be evaluated as follows:

(15) $$h_0(N) \propto \frac{1}{N}\exp\left\{-\frac{1}{2}\left(\frac{\ln(N)-\ln(N_0)}{cv}\right)^2\right\}$$

Setting the $cv > 0.7$, approximately, will cause the prior to have very little impact on the posterior (i.e., uninformative).

4.3 Run Size Calibrated to the Historical Record

The prior of run size calibrated to the time density parameters and previous information (h_1) was accomplished by Bayesian regression relating cumulative CPUE and timing information to total run size. We examined the impact of cumulative CPUE, timing mean and spread with and without log-transformations through ordinary linear predictive regression. The regression model with the best fit to the data (in terms of adjusted R^2) was

$$N_y = a_0 + a_1 T_y + a_3 B_{yt} \tag{16}$$

for any day. At the end of the season the regression appeared to have good predictive ability (adjusted $R^2 = 0.86$); however, one year (1998) exerted large leverage (see Fig 6). Table 3 provides the parameter estimates and analysis of variance table. It is important to note the residual variance is large and that the ultimate predictive capacity (assuming the regression model were known precisely) is ±508 thousand chum at a confidence level of 95%.

Four new parameters have been introduced by the regression, i.e., a_0, a_1, a_2 and σ^2 the residual variance. Fortunately, if we assume that the joint prior for these parameters is a noninformative uniform distribution, namely

$$h(a_0, a_1, a_2, \sigma^2) \propto \frac{1}{\sigma^2} \tag{17}$$

then the Bayesian regression predictive posterior distribution is a univariate Student-t distribution with n-3 degrees of freedom with the parameter estimates obtained from ordinary least-squares (Gelman et al. 1995, p236-239). Therefore, the desired distributional kernel is

$$h_1(N \mid S, T, \mathbf{X}_{\mathbf{pt}}) = h_1(N \mid T, B_t, \mathbf{N}_{\mathbf{p}}, \mathbf{T}_{\mathbf{p}}) \propto \left[1 + \frac{(N - \hat{N})^2}{(n-3)\cdot \hat{V}}\right]^{-(n-2)/2} \tag{18}$$

where n is the number of years of observations, $\hat{N}$ is the prediction from the regression and $\hat{V}$ is the predictive variance for a given B_t and T (see any statistical text for the computational formulas, e.g. Draper and Smith 1981).

4.4 Prior for Temporal Run Distribution

The distribution of historical timing and spread (Table 2) were examined with normal probability plots (see Fig. 7). The linear relationships indicate marginal normal distributions; however, timing and spread were significantly correlated ($\rho = 0.79$, $P < 0.001$). Therefore, we assumed a bivariate normal distribution for the joint prior, i.e.,

(19)

$$h(T, S \mid \mathbf{X}_{\mathbf{pt}}) = h(T, S \mid \mathbf{T}_{\mathbf{p}}, \mathbf{S}_{\mathbf{p}}) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{T-\mu_T}{\sigma_T}\right)^2 - 2\rho\left(\frac{T-\mu_T}{\sigma_T}\right)\left(\frac{S-\mu_S}{\sigma_S}\right) + \left(\frac{S-\mu_S}{\sigma_S}\right)^2\right]\right\}$$

where ρ is the correlation, μ_T the timing mean, σ_T the timing standard deviation, μ_S the spread mean and σ_S the spread standard deviation.

4.5 Likelihood of Observations

The likelihood of observing the data can only be calculated conditional on a model. Our model is extremely simple in that we assume that the adjusted cumulative CPUE is proportional to the run that has passed the Albion test-fishing site. Mathematically,

$$B_i \propto G(i \mid S, T) \cdot N \cdot e^{\varepsilon} \propto G(i \mid S, T) \cdot e^{\varepsilon} \qquad \text{for } i = 1, 2, \ldots, t \tag{20}$$

where ε is a normally distributed error term with mean 0. Note that the total run size N is effectively constant within a year and need not be considered in the computation of the likelihood. We also assume a log-normal multiplicative error structure since the variance must increase with the cumulative CPUE. A slight rearrangement of (20) yields:

$$(21) \qquad \ln\left\{\frac{B_i}{G(i \mid S,T)}\right\} - q = \varepsilon \qquad \text{for } i = 1, 2, \ldots, t$$

where q is a proportionality constant.

Two new parameters have been introduced, i.e., q and its variance $\sigma_q^{\;2}$. If we assume that the joint prior for these parameters is a noninformative uniform distribution, namely

$$(22) \qquad h(q, \sigma_q^{\;2}) \propto \frac{1}{\sigma_q^{\;2}}$$

and that the ε's are independent and normally distributed then Walters and Ludwig (1994) showed that the likelihood kernel for the observations is as follows:

$$(23) \qquad g(\mathbf{B_t} \mid N, S, T) = g(\mathbf{B_t} \mid S, T) \propto (s^2)^{-(n-1)/2}$$

where

$$(24) \qquad s^2 = \frac{\sum (Z_i - \hat{q})^2}{n-1}$$

$$(25) \qquad \hat{q} = \frac{\sum Z_i}{n}$$

$$(26) \qquad Z_i = \ln\left\{\frac{B_i}{G(i \mid S,T)}\right\}$$

4.6 Computation

The calculation of the marginal posterior can require massive computational effort particularly if there are many parameters (more than three, say). Complex problems can be dealt by approaches such as Markov chain simulation (Gelman et al. 1995) or the sampling-importance-resampling algorithm (Smith and Gelfand 1992). While nine parameters have been identified in the above problem, six of the parameters were integrated (removed) from the posterior through least-squares estimation leaving only the three time density parameters *(N, S* and *T*). Therefore, a straightforward grid system can be used to calculate the posterior using the following nested algorithm:

1. Estimate regression parameters and associated error terms for historical data (equation 16).
2. Calculate the priors at all grid points using equations (15) and (19).

3. Loop over all time evaluation points:
4. Estimate run size and variance from regression parameters (step 1) using B_t and T as inputs and calculate the kernel for all evaluation points for N using equation (18)
5. Loop over all spread evaluation points:
6. Calculate the likelihood of the observations for every S and T using equations (23)-(26)
7. Loop over all run size evaluation points:
8. Update the joint posterior by the product of steps (2), (4) and (6) for every T, S and N.
9. Sum the total posterior and sum into each of the three marginal posteriors.

When these steps are complete, the marginal distributions need to be normalised (scaled such that their sum is 1). Finally, any statistics of interest (e.g., mean, median, mode, highest probability regions) may be calculated directly from the marginal distributions.

5.0 RETROSPECTIVE ANALYSIS

Historical information can be used to evaluate the performance of a predictive model. In our case, there were 20 years of data (1979-1998). In a retrospective analysis, each year is chosen for prediction with all other years assumed known. For example, if 1985 was chosen then the 1979-1984 and 1986-1998 data represent the "historical record" and the 1985 cumulative CPUE up to time t represent the "current year" data that are used for prediction. The estimated escapement for all years and days within a year can then be compared to the observed escapement. In order to standardise yearly timing variation and to reduce the computation burden we chose to evaluate all sampling occurrences 7 days before and 30 days after the yearly timing mean (peak). This 38-day duration should encompass the period when most management decisions are necessary. The latest run occurred in 1985 with an evaluation period of 30 October to 6 December while the earliest run occurred in 1998 with an evaluation period of 13 October to 19 November.

The classical test-fishing model of simple linear regression using escapement (run size) and cumulative CPUE, i.e.,

$$N_y = b_0 + b_1 B_{yt} \tag{27}$$

where the b's are the regression parameters, was first evaluated as a standard for comparison with the Bayes model. Figure 8 plots the percent residual (relative) error from the actual (observed) escapement as a function of days from the peak (timing mean) for this regression model. A line connecting points signifies a common year. Similarly, Fig. 9 plots the residual error of the posterior escapement mean from the Bayes model. The parameters for the prior distributions were identical in all years and were obtained from the historical data displayed in Table 2. In order to facilitate the interpretation of Figs. 8 and 9, Fig. 10 plots the mean, absolute mean and standard deviation of the residual errors by day from the peak. For both the simple regression and Bayes models, the accuracy does not improve despite the accumulation of more information with the progression of the season. In addition, both models exhibit a tendency to overestimate with respect to percentages (each year has equal weight despite large differences in run size) with the

Bayes model somewhat worse than the regression model. However, the Bayes model is more accurate (closer to the actual escapement) as illustrated by the smaller absolute residual error and standard deviation (see Fig. 10).

As pointed out previously, the residual variance of the CPUE calibrated to historical record (equation 16) is large and overwhelms all other components. Perhaps the uncertainty is generated by the upriver escapement estimates and the end of season cumulative CPUE from the test-fishery provides a more precise index of the run size at Albion. Under this hypothesis, the inseason goal should be the prediction of year-end CPUE. (The final CPUE need only be scaled to the mean run size, e.g., equation 16 or 27; moreover, prediction of the mean run becomes more precise as more years are added.) The adjustment to the Bayes model in order to predict final CPUE is uncomplicated, i.e., simply replace the escapement vector by the final CPUE vector (see Table 2). The subsequent percent residual (relative) errors from the predicted and actual (observed) final cumulative CPUE are plotted in Fig.11. Note that the residual errors are considerably smaller than the escapement estimates and that the CPUE predictions converge towards the true value as the season progresses. The daily mean profile of the 90% highest probability density (HPD) regions of the posterior distribution for the prediction of escapement and final CPUE are contrasted in Fig. 12.

6.0 DISCUSSION AND RECOMMENDATIONS

The Bayesian procedure presented in this paper is superior to the classical test fisheries approach of using a simple predictive regression of cumulative CPUE on run size. Not only does our procedure use a predictive regression to calibrate the run size but also incorporates preseason knowledge and within season information on run timing. In addition, the posterior distribution provides the necessary precursor for risk assessment and policy evaluation.

The normal time-density assumption for migration is the weakest aspect of the Bayes model as presently constructed. Distributions that are more flexible to account for skewness and kurtosis might provide some improvement in the predictions. Alternatively, partitioning the composite run into timing components with stock composition information may yield distributions that are more normal-like. However, regardless of the time-density presumption, the predictive ability of the model will be overwhelmed by the reliability of escapement counts (at least ±508 thousand at 95% confidence).

The clockwork management regime for the Fraser River chum salmon (Table 1) can not be supported by the current information on run timing and abundance. First, the timing at Albion has not been markedly bimodal particularly during the 1990's (Fig. 4) despite the presence of stocks with different timing characteristics (Palmer 1972). The current practice of estimating the composite run and then assuming a fixed proportion to partition the run into early and late components only serves to obscure the problem. The management plan needs to be rationalised for a single composite or separate inseason estimation procedures and data collection (e.g., stock composition information) must be developed for the separate timing components.

Second, with regards to abundance estimation, the uncertainty problem might be addressed with one or more of the following strategies:

1) Use a coarser grain trigger in the clockwork plan to account for uncertainty (e.g., raise the trigger to 1 million) or incorporate uncertainty directly into the plan (e.g., 90% chance that the run is 500 thousand or larger). The downside of such a plan is that fishing opportunities will be reduced.

2) Reduce uncertainty in the escapement data through increased sampling effort for escapement enumeration. However, this would most certainly entail increased costs for sampling.

3) Manage the fishery under the hypothesis that the test-fishery provides a precise index of run size at Albion.

Managers will have to weigh the costs, benefits and risks of alternative strategies. Our analysis to date confirms the need to examine the collection and reliability of the escapement estimates for Fraser River chum salmon.

7.0 ACKNOWLEDGMENTS

We would like to thank Melanie Sullivan for the difficult task of getting all the data from the Fraser River test fisheries together into a form suitable for analysis.

8.0 FIGURES

Figure 1. Site Map of the Cottonwood test fishing area.
Figure 2. Site Map of the Albion test fishing area.
Figure 3. Plots of raw CPUE and CPUE adjusted for saturation versus catch by set.
Figure 4. Historical run distribution of chum salmon from the Albion test fishery (day 1 = 1 September).
Figure 5. Time density at the Albion test fishery site.
Figure 6. Plot of adjusted CPUE and timing at Albion versus escapement
Figure 7. Normal probability plots for timing and spread (straight line indicates data distributed normally).
Figure 8. Percent residual (relative) error from the actual (observed) escapement as a function of days from the peak (timing mean) for simple linear regression model (each line represents a common year).
Figure 9. Percent residual (relative) error from the actual (observed) escapement as a function of days from the peak (timing mean) for the Bayes model (each line represent a common year).
Figure 10. Mean daily profile of the percent error, absolute value of percent error and standard deviation derived from the regression and Bayes models.
Figure 11. Percent residual (relative) error from the actual (observed) final CPUE as a function of days from the peak (timing mean) for the Bayes model (each line represent a common year).
Figure 12. Mean daily profile of the 90% highest probability density (HPD) regions of the posterior distribution for prediction of escapement and final CPUE.

Figure 1. Site Map of the Cottonwood test fishing area.

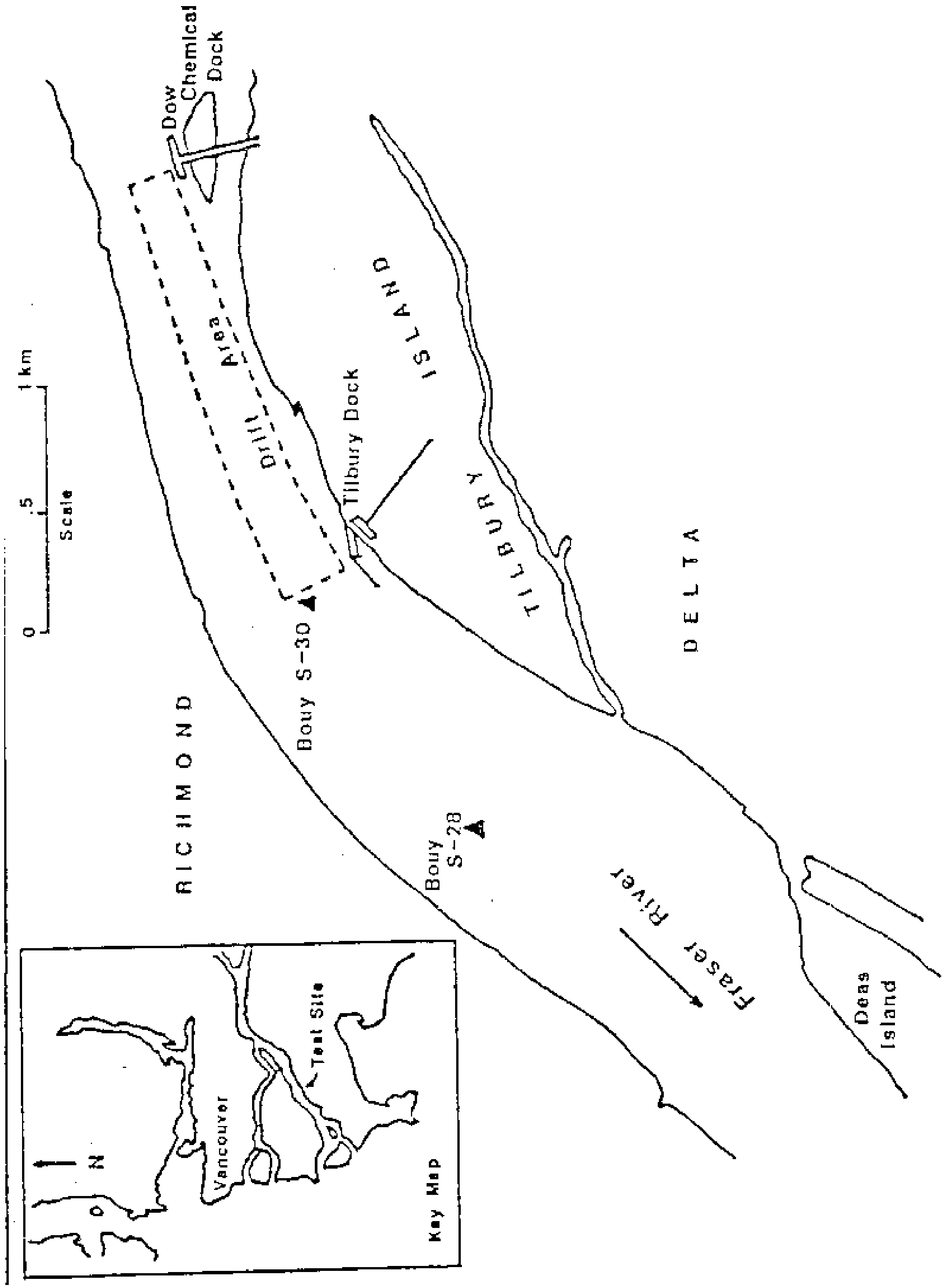


Figure 2. Site Map of the Albion test fishing area.

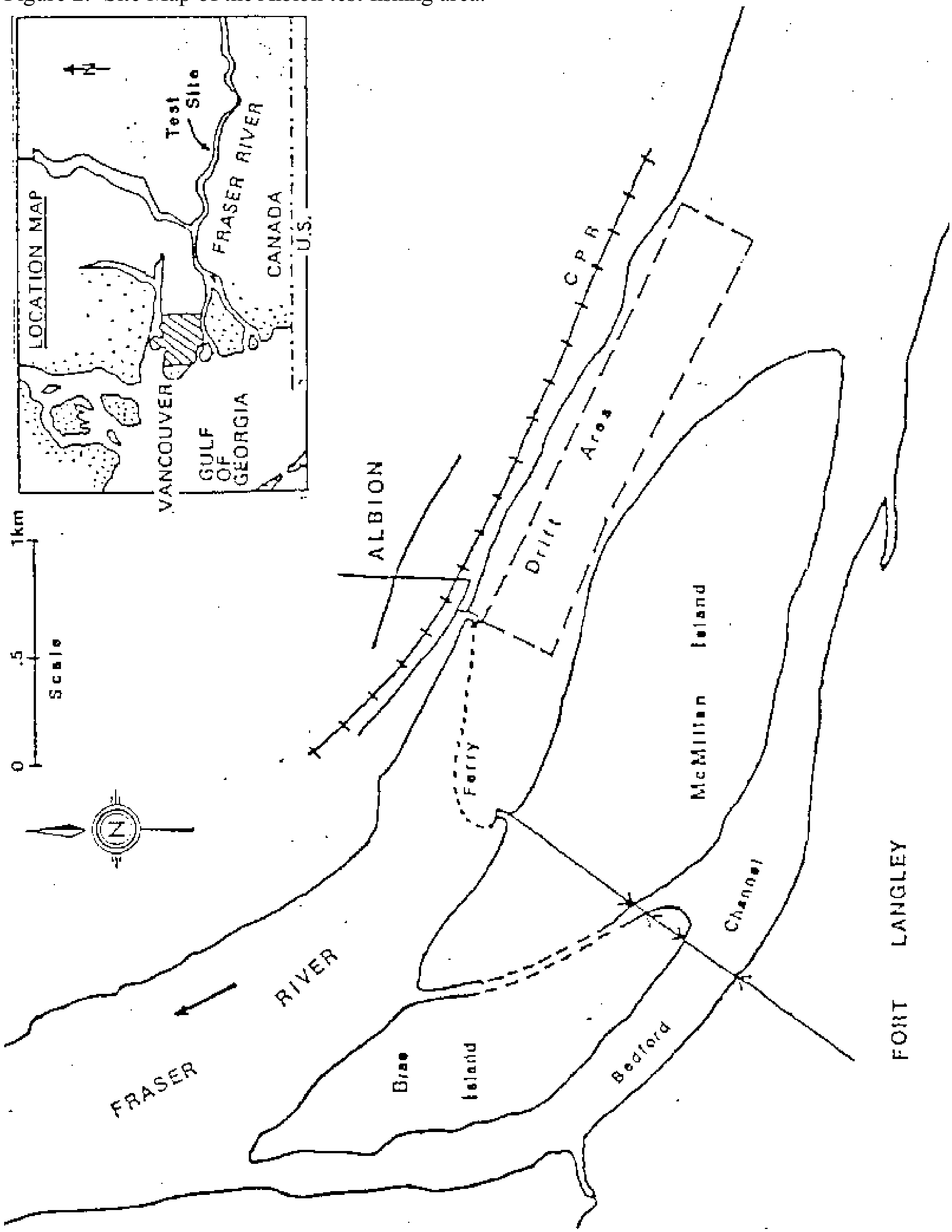


Figure 3. Plots of raw CPUE and CPUE adjusted for saturation versus catch by set.

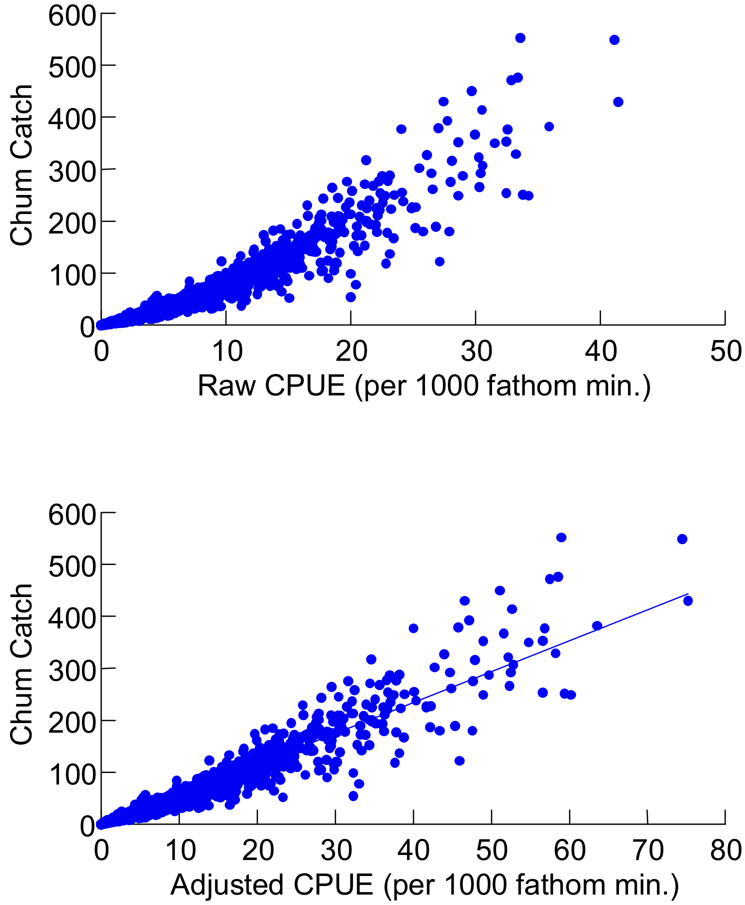


Figure 4. Historical run distribution of chum salmon from the Albion test fishery (day 1 = 1 September).

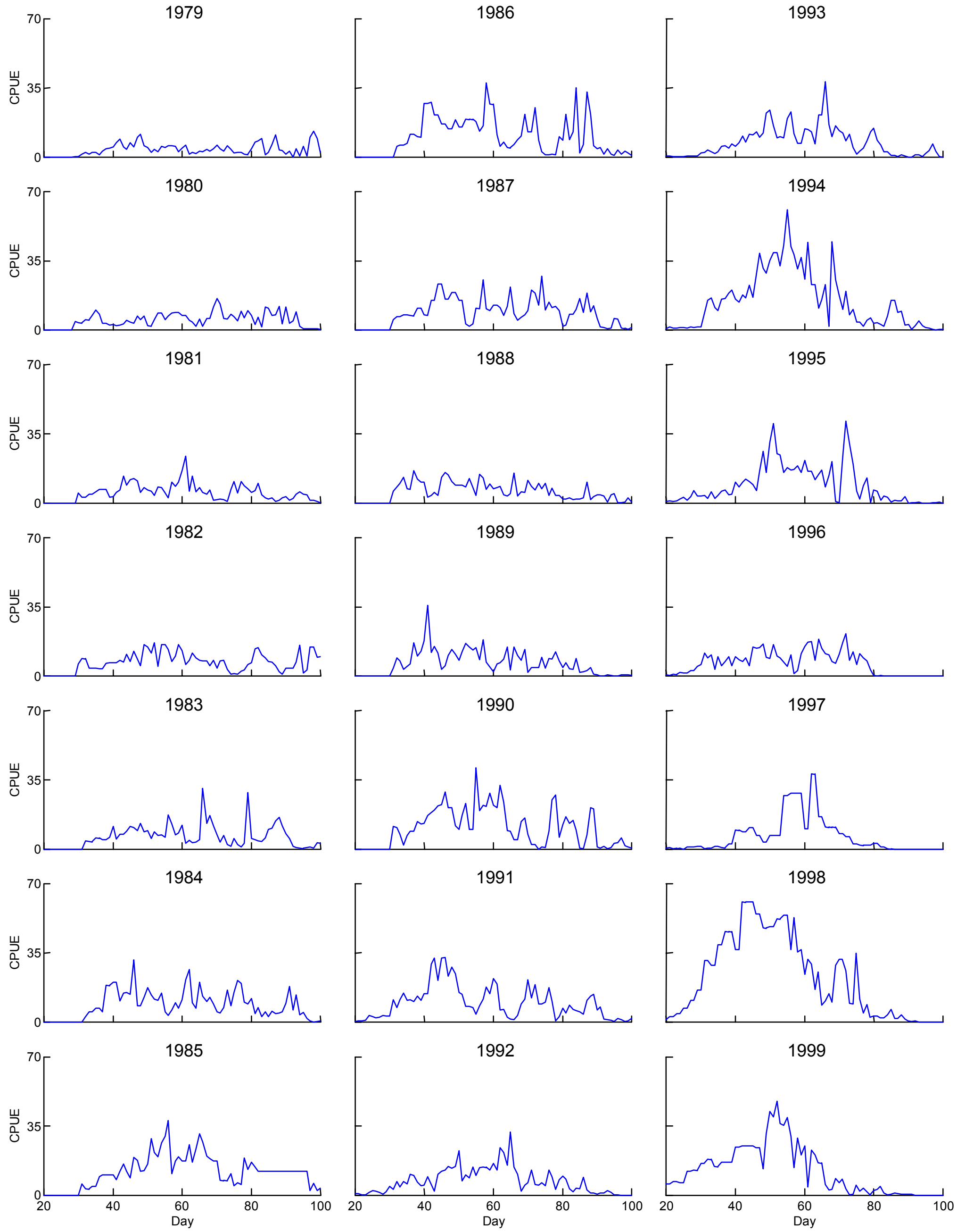


Figure 5. Time density at the Albion test fishery site.

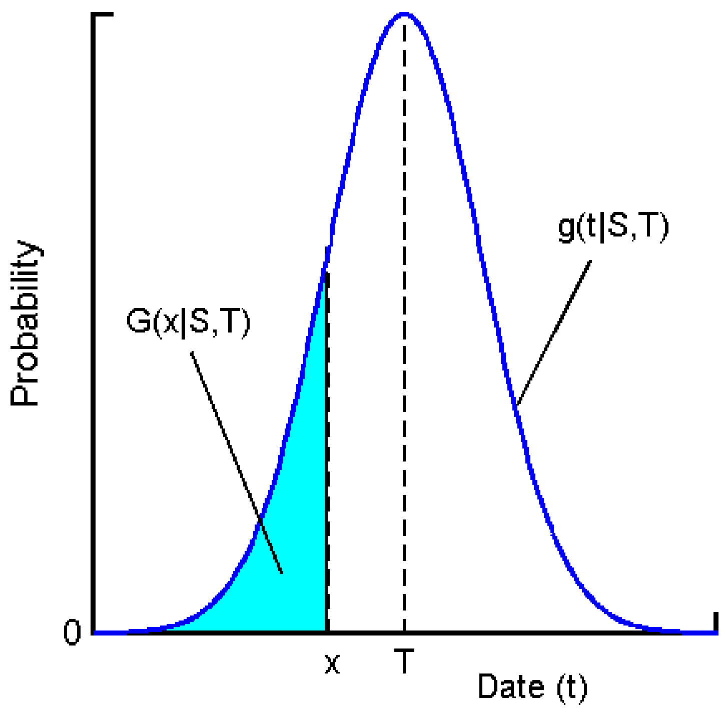


Figure 6. Plot of adjusted CPUE and timing at Albion versus escapement.

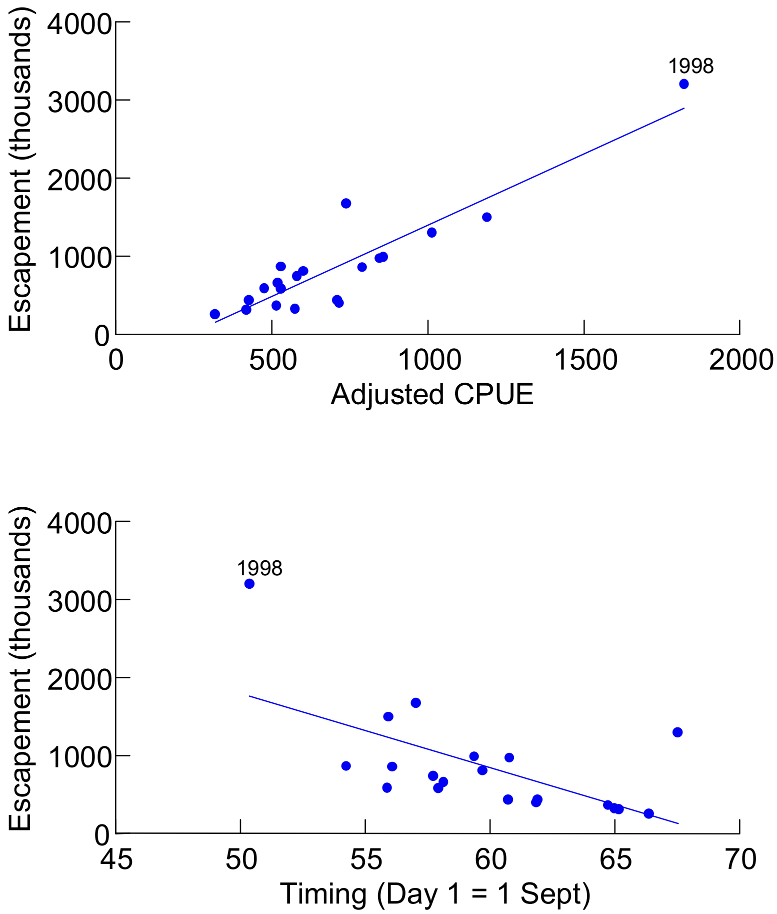


Figure 7. Normal probability plots for timing and spread (straight line indicates data distributed normally).

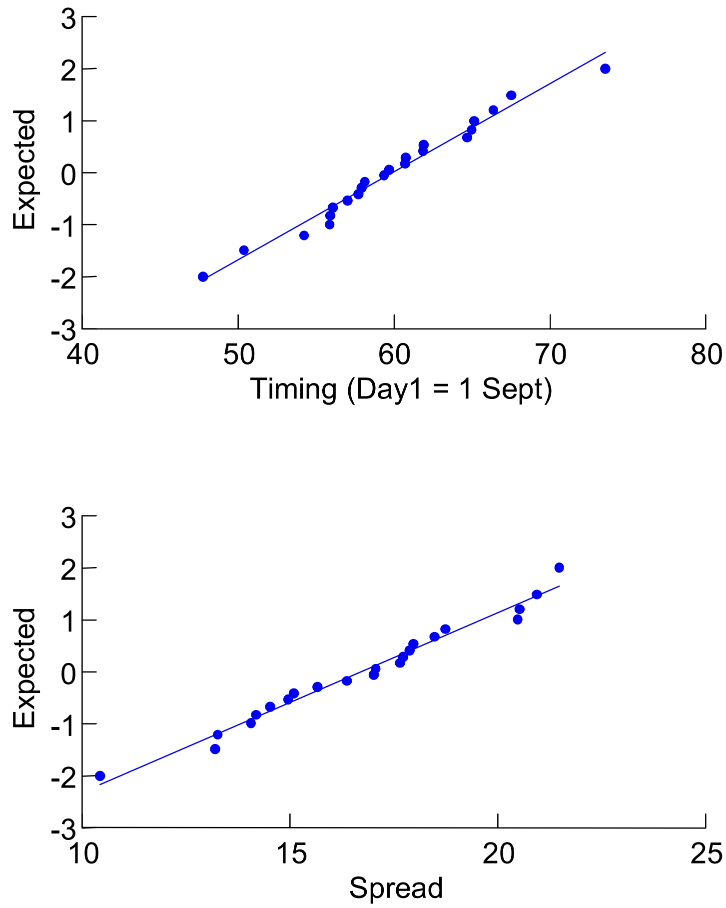


Figure 8. Percent residual (relative) error from the actual (observed) escapement as a function of days from the peak (timing mean) for simple linear regression model (each line represents a common year).

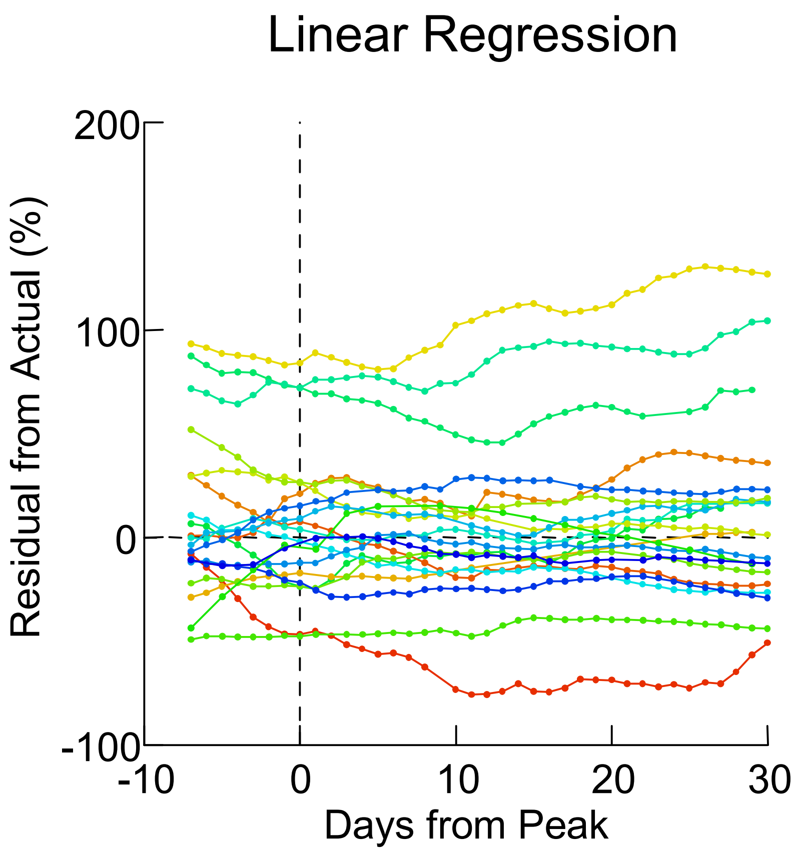


Figure 9. Percent residual (relative) error from the actual (observed) escapement as a function of days from the peak (timing mean) for the Bayes model (each line represent a common year).

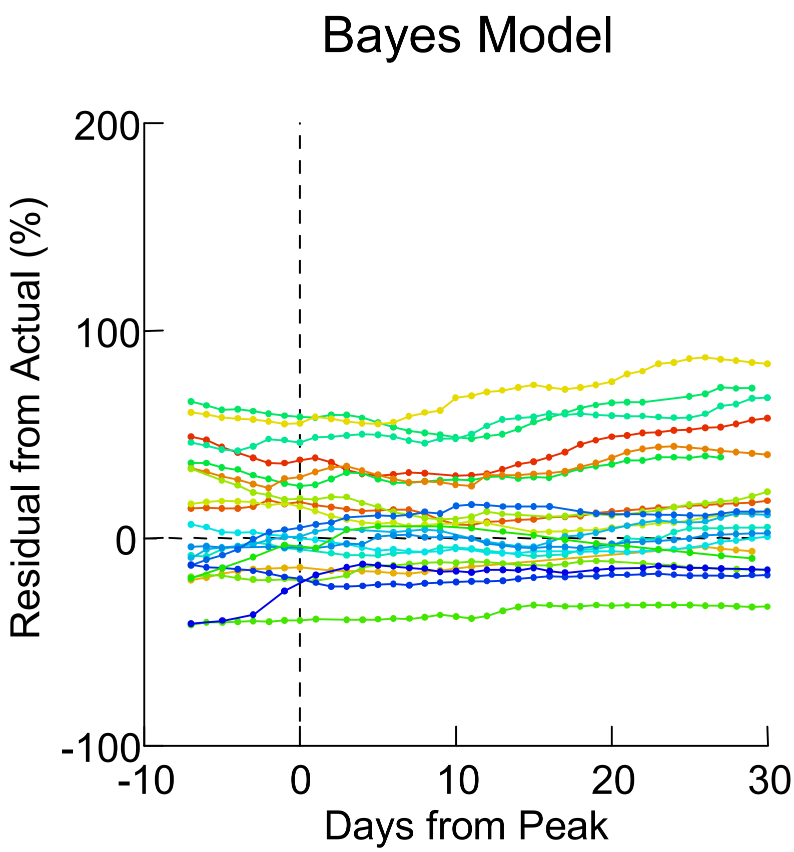


Figure 10. Mean daily profile of the percent error, absolute value of percent error and standard deviation derived from the regression and Bayes models.

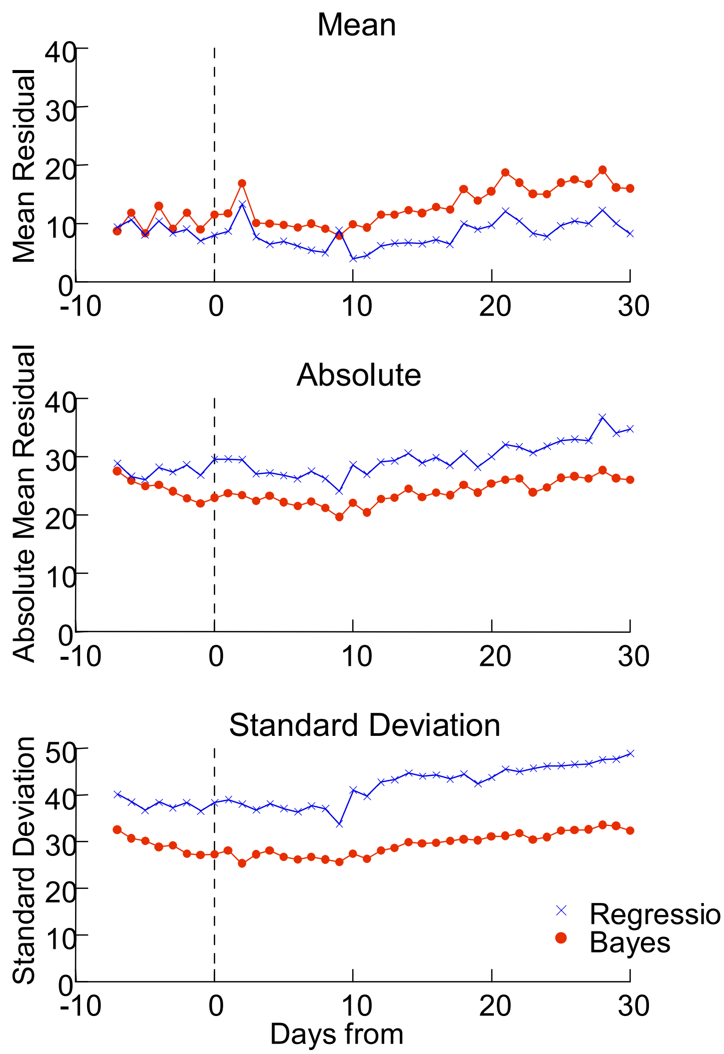


Figure 11. Percent residual (relative) error from the actual (observed) final CPUE as a function of days from the peak (timing mean) for the Bayes model (each line represent a common year).

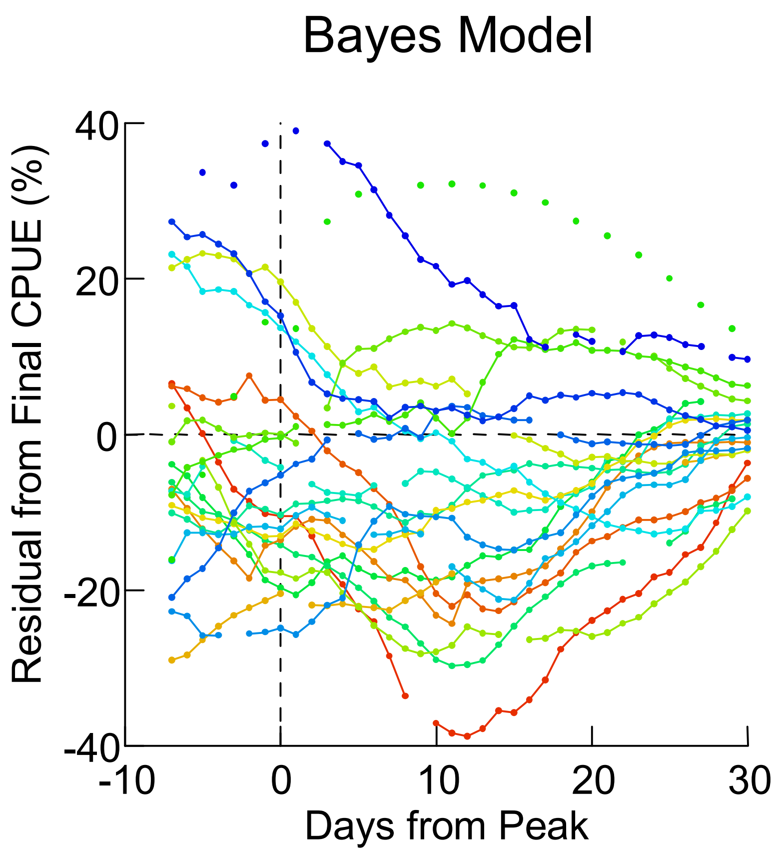
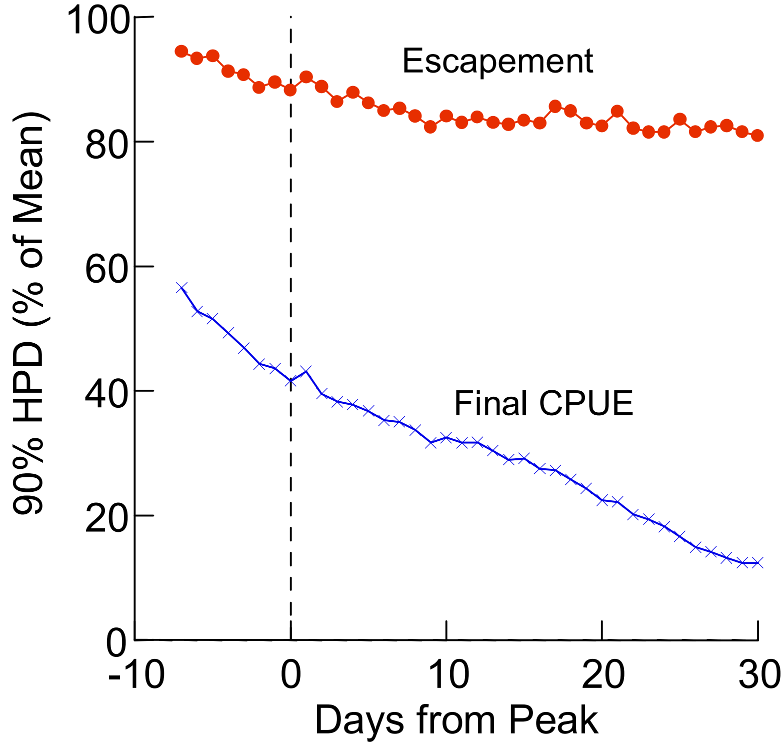


Figure 12. Mean daily profile of the 90% highest probability density (HPD) regions of the posterior distribution for prediction of escapement and final CPUE.



9.0 TABLES

Table 2. Elementary timing and run statistics.

Year	Escapement (thousands)	Cumulative CPUE	Timing Mean[1]	Timing Spread
1978		523.7	73.9	20.2
1979	255.6	318.8	67.1	20.9
1980	312.1	418.8	65.2	18.7
1981	435.3	426.9	60.7	17.6
1982	320.3	573.3	65.0	20.4
1983	365.0	514.5	64.9	18.3
1984	433.3	710.4	62.0	17.9
1985	1295.3	1012.4	68.1	20.6
1986	972.9	844.7	60.8	17.0
1987	398.3	715.2	61.9	17.0
1988	585.3	474.9	58.1	17.8
1989	585.6	528.7	56.0	15.6
1990	988.9	855.9	59.3	16.4
1991	859.3	789.1	56.1	17.7
1992	741.6	579.2	57.7	14.9
1993	809.6	601.6	59.7	14.5
1994	1498.4	1188.6	55.9	14.1
1995	1673.1	738.6	57.0	14.2
1996	867.6	529.9	54.2	15.1
1997	656.9	519.4	58.1	10.4
1998	3200.0	1820.4	50.4	13.3
1999		1014.0	47.8	13.2
Mean	862.7	713.6	60.0	16.6
Std. Dev	678.6	330.1	5.9	2.8
CV	0.79	0.46	0.10	0.17

[1] Julian date starting 1 September.

Table 3. Regression results for equation (16) at the end of season (units in thousand of chum).

Parameter	Estimate	t-value	Probability
a_0	1952.55	1.98	0.063
a_1	-37.1	-2.45	0.025
a_2	1.595	7.99	<0.001

Source	df	Mean-Square	F-ratio	Probability
Regression	2	3.81E+06	5.73	<0.001
Residual	17	6.64E+04		

10.0 LITERATURE CITED

Cave, J.D., and W.J. Gazey. 1994. A preseason simulation model for fisheries on Fraser River sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci., 51: 1535-1549.

Draper, N.R., and W. Smith. 1981. Applied Regression Analysis (Second Ed.). John Wiley and Sons. 709p.

Farwell, M.K. 1985. The Chum Salmon Test Fishery in the Fraser River: Catch and Effort Summary, 1963 and 1965 to 1984. Canadian Data Report on Fisheries and Aquatic Sciences, No. 529.

Farwell, M.K., M. Joyce, A. Charlie. 1992. Enumeration of the 1991-1992 Harrison river System Chum Salmon Escapement. Chehalis Band Report.

Farwell, M.K., 1993. Enumeration of the 1992-1993 Harrison river System Chum Salmon Escapement. Chehalis Band Report.

Farwell, M.K., 1994. Enumeration of the 1993-1994 Harrison river System Chum Salmon Escapement. Chehalis Band Report.

Farwell, M.K., 1995. Enumeration of the 1994-1995 Harrison river System Chum Salmon Escapement. Chehalis Band Report

Fried, S.M. and R. Hilborn. 1988. Inseason forecasting of Bristol Bay, Alaska, sockeye salmon (*Oncorhynchus nerka*) abundance using Bayesian probability theory. Can. J. Fish. Aquat. Sci., 45: 850-855.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 1995. Bayesian Data Analysis. Chapman and Hall, London. 526p.

Munday, P.R. 1979. Computation of migratory timing illustrated by application to the management of commercial salmon fisheries. Doctoral dissertation. University of Washington, Seattle.

Palmer, R.N. 1972. Fraser River Chum Salmon. Canadian Department of the Environment, Fisheries Service, Tech. Report 1971-2, 284p.

Pella, J.J., M.M. Masuda and D.G. Chen. 1998. Forecast methods for inseason management of the Southeast Alaskan chinook salmon troll fishery. *In*: F. Funk, T.J. Quinn II, J Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.I. Zhang (eds.), Fishery stock assessment models. University of Alaska Sea Grant, AK-98-01: 287-314.

Ricker, W.E. 1973. Linear regressions in fisheries research. J. Fish. Res. Board Can., 30: 409-434.

Ryall, P.J. and A. Cass. 1995. An evaluation of run size estimation techniques used for the Clockwork chum stocks. PSARC Working Paper S95-08.

Ryall, P.J.,C. Murray, R.V. Palermo, D. Baily and D. Chin. 1999. Status of Clockwork Chum Salmon Stock and Review of the Clockwork Management Strategy. Canadian Stock Assessment Secretariat Research Document 99/169.

Schnute, J., and J. Sibert. 1983. The salmon terminal fishery: a practical comprehensive timing model. Can. J. Fish. Aquat. Sci., 40: 835-853.

Springborn, R.B., N.D. Lampsakis and V.F. Gallucci. 1998. A time density model to estimate run size and entry timing in a salmon fishery. N. Amer. J. of Fish. Mgmt., 18: 391-405.

Smith, A.F.M. and A.E. Gelfand. 1992. Bayesian statistics without the tears: a sampling-resampling perspective. The American Statistician, 46: 84-88.

Thomas, J.O. 1996. Enumeration of the 1995-1996 Harrison river System Chum Salmon Escapement. Chehalis Band Report

Walters, C., and D. Ludwig. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci., 51: 713-722.