Fishery Manuscript No. 14-08

Forecasting Annual Run Size of Chinook Salmon to the Taku River of Alaska and Canada

by

David R. Bernard

and

Edgar L. Jones III

December 2014

Alaska Department of Fish and Game

Divisions of Sport Fish and Commercial Fisheries

Symbols and Abbreviations

The following symbols and abbreviations, and others approved for the Système International d'Unités (SI), are used without definition in the following reports by the Divisions of Sport Fish and of Commercial Fisheries: Fishery Manuscripts, Fishery Data Series Reports, Fishery Management Reports, and Special Publications. All others, including deviations from definitions listed below, are noted in the text at first mention, as well as in the titles or footnotes of tables, and in figure or figure captions.

Weights and measures (metric)

centimeter	cm
deciliter	dL
gram	g
hectare	ha
kilogram	kg
kilometer	km
liter	L
meter	m
milliliter	mL
millimeter	mm

Weights and measures (English)

cubic feet per second	ft^3/s
foot	ft
gallon	gal
inch	in
mile	mi
nautical mile	nmi
ounce	oz
pound	lb
quart	qt
yard	yd

Time and temperature

day	d
degrees Celsius	°C
degrees Fahrenheit	°F
degrees kelvin	K
hour	h
minute	min
second	s

Physics and chemistry

all atomic symbols	
alternating current	AC
ampere	A
calorie	cal
direct current	DC
hertz	Hz
horsepower	hp
hydrogen ion activity (negative log of)	pH
parts per million	ppm
parts per thousand	ppt, ‰
volts	V
watts	W

General

Alaska Administrative Code	AAC
all commonly accepted abbreviations	e.g., Mr., Mrs., AM, PM, etc.
all commonly accepted professional titles	e.g., Dr., Ph.D., R.N., etc.
at	@
compass directions:	
east	E
north	N
south	S
west	W
copyright	©
corporate suffixes:	
Company	Co.
Corporation	Corp.
Incorporated	Inc.
Limited	Ltd.
District of Columbia	D.C.
et alii (and others)	et al.
et cetera (and so forth)	etc.
exempli gratia (for example)	e.g.
Federal Information Code	FIC
id est (that is)	i.e.
latitude or longitude	lat or long
monetary symbols (U.S.)	$, ¢
months (tables and figures): first three letters	Jan,...,Dec
registered trademark	®
trademark	™
United States (adjective)	U.S.
United States of America (noun)	USA
U.S.C.	United States Code
U.S. state	use two-letter abbreviations (e.g., AK, WA)

Mathematics, statistics

all standard mathematical signs, symbols and abbreviations	
alternate hypothesis	H_A
base of natural logarithm	*e*
catch per unit effort	CPUE
coefficient of variation	CV
common test statistics	(F, t, χ^2, etc.)
confidence interval	CI
correlation coefficient (multiple)	R
correlation coefficient (simple)	r
covariance	cov
degree (angular)	°
degrees of freedom	df
expected value	*E*
greater than	>
greater than or equal to	≥
harvest per unit effort	HPUE
less than	<
less than or equal to	≤
logarithm (natural)	ln
logarithm (base 10)	log
logarithm (specify base)	log_2, etc.
minute (angular)	'
not significant	NS
null hypothesis	H_O
percent	%
probability	P
probability of a type I error (rejection of the null hypothesis when true)	α
probability of a type II error (acceptance of the null hypothesis when false)	β
second (angular)	"
standard deviation	SD
standard error	SE
variance	
population	Var
sample	var

FISHERY MANUSCRIPT NO. 14–08

FORECASTING ANNUAL RUN SIZE OF CHINOOK SALMON TO THE TAKU RIVER OF ALASKA AND CANADA

by

David R. Bernard
D. R. Bernard Consulting LLC, Ankeny, Iowa

and

Edgar L. Jones III
Alaska Department of Fish and Game, Division of Sport Fish, Juneau, Alaska

Alaska Department of Fish and Game
Division of Sport Fish, Research and Technical Services
333 Raspberry Road, Anchorage, Alaska, 99518-1599

December 2014

Development and publication of this manuscript were financed by NOAA Grant No. NA10NMF4380279 appropriated by the U.S. Congress for implementation of the U.S. Chinook Letter of Agreement.

The Fishery Manuscript Series was established in 1987 by the Division of Sport Fish for the publication of technically oriented results of several years' work undertaken on a project to address common objectives, provide an overview of work undertaken through multiple projects to address specific research or management goal(s), or new and/or highly technical methods, and became a joint divisional series in 2004 with the Division of Commercial Fisheries. Fishery Manuscripts are intended for fishery and other technical professionals. Fishery Manuscripts are available through the Alaska State Library and on the Internet: http://www.adfg.alaska.gov/sf/publications/. This publication has undergone editorial and peer review.

David R. Bernard
D. R. Bernard Consulting, LLC
2481NW 87th Ave., Ankeny, Iowa 50023-8829, USA

and

Edgar L. Jones III [a]
ed.jones@alaska.gov
Alaska Department of Fish and Game, Division of Sport Fish,
802 3rd Street, Douglas, AK 99824, Juneau, AK 99811-0020, USA

[a] *Author to whom all correspondence should be addressed.*

This document should be cited as:
Bernard, D. R., and E. L. Jones III. 2014. Forecasting annual run size of Chinook salmon to the Taku River of Alaska and Canada. Alaska Department of Fish and Game, Fishery Manuscript No. 14-08, Anchorage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

An alternative to the current method of forecasting size of the annual terminal run of large (age 1.3−1.5, approximately >659 mm from mid eye to tail fork) Chinook salmon *Oncorhynchus tshawytscha* to the Taku River is demonstrated. Both the current method and the alternative are based on sibling relationships within brood years; however, the alternative is based on median forecasts from power functions with lognormal error. Bayesian analysis using the program WinBUGS was used to avoid bias in forecasted run size from measurement error, and to quantify conditional uncertainty in forecasts. Because of greater imprecision in estimates of annual run size in earlier years, only data from 1995 to 2013 were used to exercise both current and alternative methods. From jackknifed hindcasting the mean percent error for the current method was +13% and +7% for the alternative; the former had two negative predictions. A Bayesian forecast with the alternative method produced the posterior probability distribution for the estimated terminal run size in 2014 with median 24,440 and mean 25,980 large salmon. Management targets under the Pacific Salmon Treaty for Taku fisheries were described, and the forecast posterior probability distribution was used to calculate probabilities of meeting those targets in 2014. Relevance of forecasting with a truncated data series of recent years was discussed. Instructions were given on how to expand the WinBUGS code to produce a forecast for 2015 and for years beyond.

Key words: Chinook salmon, Taku River, forecast, sibling regressions, Bayesian regression, medians, MPE, MAPE, credibility intervals.

INTRODUCTION

Management of terminal fisheries and near-terminal fisheries on Chinook salmon (*Oncorhynchus tshawytscha*) migrating to the Taku River of Alaska and Canada (Figure 1) depends in part on a preseason forecast of annual terminal run size[1]. In 2011 and before, the preseason forecast depended on sibling relationships. The presumption was that numbers of older Chinook salmon are linearly related to numbers of younger Chinook salmon in the same brood year. The current model used is to regress estimated terminal run of age-1.3[2] Chinook salmon in a calendar year (*cy)* against the estimated terminal run of age-1.2 salmon in year *cy* – 1. A second linear regression involves age-1.4 and age-1.3 Chinook salmon from years *cy* and *cy* – 1, respectively. These regressions are updated annually, then used to predict the number of age-1.3 and age-1.4 Chinook salmon in the terminal run for the upcoming year (*cy* + 1). Residual error in these regressions is presumed additive. The sum of the two predictions is considered the preseason forecast for large Chinook salmon (>659 mm from mid eye to tail fork [METF]), the group relevant for fisheries management[3].

The shortcomings of the current method for forecasting are as follows:

- not all brood year information is used in the forecast of age-1.4 fish;
- effect of measurement error in independent variables on both the 1.3 vs. 1.2 and 1.4 vs. 1.3 regressions is ignored;
- negative predictions for an age group are possible;
- residual error follows a normal distribution and is additive, but it should follow a log-normal distribution and be multiplicative; and
- uncertainty in forecasts is incompletely expressed.

---

1 As mandated in ANNEX IV, Chapter 1, Section (3)(b)(3)(viii) of the Pacific Salmon Treaty, December 8, 2008.

2 European notation is used to describe age of salmon, such that an age *x.y* salmon spent *x + 1* winters in fresh water incubating and rearing, and *y* years at sea maturing, with an overall age of *x + y + 1* years.

3 A third regression of age-1.5 and age-1.4 in years *cy* and *cy* – 1 is not included because age-1.5 salmon make up a few hundred fish per year in the terminal run.

A point-by-point description of these shortcomings in the current method follows:

The run of age-1.4 fish is in part a function of the run of age-1.2 fish two years earlier; an age-1.4 vs. age-1.2 or -1.3) regression would potentially capture that added information.

Ignoring measurement error in independent variables can bias parameter estimates and hence bias forecasts from sibling regressions.

Given a steep enough slope in a regression, a negative intercept is possible, which can lead to a nonsensical prediction (i.e., a negative number of fish for a particular age class).

Both process error and measurement error in dependent variables are more likely log-normally distributed (i.e., estimates of run size cannot be negative); assuming additive normal error in sibling regressions tends to over-forecast small runs and under-forecast large ones.

Finally, a single-number forecast from the current method does not have an associated estimate of uncertainty.

The work described in this report was undertaken to improve the current preseason forecast. The resulting alternative method still depends upon sibling relationships, but measurement error in statistics and uncertainty in results is incorporated. This report begins with a description of available data and then continues with an exercise of the current method, the development of the alternative method, and an exercise of the alternative method relative to management of Taku fisheries.

FORECASTING METHODS

THE DATA

The 35-year time series of basic demographic estimates of annual run size of the Taku stock is displayed in Figure 2 (for a description of estimation methods and for a source of statistics, see Jones et al. 2010; McPherson et al. 1996–1999, 2000, 2010; Pahlke and Bernard 1996). Sampling to estimate age and sex composition of spawning abundance and harvest extends back to calendar year 1977. Spawning abundance was indexed annually through aerial surveys from 1977 to the present (2013 at the time of this writing). Mark–recapture studies were used to estimate spawning abundance for 1989, 1990, 1995–1997, and 1999 to the present; these estimates also allowed estimation of an expansion factor used to convert aerial indices from other years to spawning abundances. A result of this expansion is that precision of statistics for years prior to 1995 was generally considerably less (average CV for estimated run size is 31%) than for years after (average CV for estimated run size is 14%) (Figure 2). Because of this early imprecision, only the more precise data collected from 1995 forward have been used in forecasting. Also, because the abundance of age-1.5 salmon is almost exclusively <5% of large salmon, this age group was ignored in forecasts.

Terminal run size consists of spawning abundance, inriver harvest, commercial harvest in District 111, and harvest in the Juneau area recreational fishery (Figure 1). Harvests in ocean fisheries are estimated through an ongoing tagging program based on coded wire tags (CWTs) and on fishery sampling to recover CWTs (McPherson et al. 2000). The estimated fraction of the total harvest of Chinook salmon composed of members of the Taku stock is multiplied by the total harvest to estimate the harvest specific to the stock. Harvest of all stocks in commercial fisheries is a tally of landings, whereas harvest of all stocks in the recreational fishery is estimated through an annual postal survey of fishing households (see Jennings et al. 2010 for a description of the survey). Because there is approximately a 1-year delay in getting estimates from this postal survey, the latest year with direct estimates of total sport harvest is 2013. Annual spawning escapement is estimated by subtracting estimated inriver harvest in Canada from the size of the inriver run at Canyon Island in Alaska just downstream of the border (Figure 1) as estimated from a series of mark–recapture studies. Annual catches in a small personal use fishery (a hundred or so fish) downstream of Canyon Island are tallied from required permits, and because these catches occur below Canyon Island and the area germane to the inriver run estimate, they are tallied with other marine harvests.

CURRENT METHOD

The current method is used to forecast terminal run size of mature, large (age-1.3–1.4) Chinook salmon. The two regressions based on estimated run size by age over calendar years 1995–2013 and brood years 1990–2007 are graphed in Figure 3 and are:

$$\hat{R}_{1.3,cy} = 10{,}173 + 1.788\hat{R}_{1.2,cy-1} + \varepsilon_{cy} \tag{1a}$$

$$\hat{R}_{1.4,cy} = -9{,}336 + 0.800\hat{R}_{1.3,cy-1} + \varepsilon_{cy} \tag{1b}$$

where $\hat{R}$ represents estimated run size by age and year, not the forecasted run, and $\varepsilon_{cy} \sim norm(0,\sigma_\varepsilon^2)$ represents additive error. Regressions are based on n consecutive years of data and $n - 1$ data pairs. Here $n = 19$. When cy represents the fishing season that just ended, both (1a) and (1b) are then used to predict the abundance of age-1.3 and age-1.4 salmon in the next fishing season (year $n + 1$ in the series). The sum of these two predictions is the forecast of terminal run size, here

$$\widetilde{R}_{n+1} = \widetilde{R}_{1.4,n+1} + \widetilde{R}_{1.3,n+1} \tag{1c}$$

with $\widetilde{R}_{n+1}$ being the forecasted number of large (age-1.3−1.4) Chinook salmon in the upcoming run. The forecast for 2014 with the current method was 24,040.

The accuracy and precision of forecasts from the current method were judged through a jackknife hindcast of data on runs from 1995 through 2013 (see Efron and Tibshirani 1993 for methods based on jackknifing). Statistics were "excluded" from the data 1 calendar year at a time from the series 1996–2013; sibling regressions as described above were recalculated from statistics for the "remaining" years, then these new sibling regressions were used with statistics from the year immediately preceding the "excluded" year to "forecast" the annual run size for the "excluded" year. Table 2 presents a simple example to demonstrate how the jackknifing and subsequent forecasting were organized. An individual percent error (IPE) was calculated for each jackknifed year as [IPE = (Forecast−Estimate) / Estimate × 100]. The mean of the IPEs, the mean percent error (MPE) for the entire $n - 1$ (18 years), is considered a measure of accuracy with an anticipated value of 0. The mean absolute IPEs (MAPE) is a measure of both precision and accuracy.

Hindcasted forecasts using the current method were on average +13% (= MPE) higher than their comparable estimates of run size with an expected MAPE of 35% (Table 3). Forecasts by age were positive for all "excluded" years except for the predicted runs of age-1.4 salmon in 1999 and in 2008 (shaded cells in Table 3). The empirical distribution of IPEs from the current method was broad and highly skewed with its mode less than zero and its mean (13%) and median (9%) above zero (Figure 4).

ALTERNATIVE METHOD

At its core the alternative method is still based on sibling relationships, but a set of 2 power functions are used to express those relationships:

$$\hat{R}_{1.3,cy} = a_1\left(\ln \hat{R}_{1.2,cy-1}\right)^{b_1} \exp(\varepsilon_{cy}) \tag{2a}$$

$$\hat{R}_{1.4,cy} = a_2\left(\hat{R}_{1.3,cy-1} + \hat{R}_{1.2,cy-2}\right)^{b_2} \exp(\varepsilon_{cy}) \tag{2b}$$

where a and b are parameters and ε_{cy} a variate following a normal probability distribution with mean 0 and variance σ_{ε}^2. The use of power functions avoids the possibility that a forecast for either age group would be negative; this change also allows for lognormal error. Summing estimated run sizes for 4- and 5-year-old salmon within the same brood year to produce an independent variable for forecasting the number of 6-year-olds in the same brood year takes advantage of all available information for that brood year. As a further refinement, the alternative method is based on predicting medians, not means. Medians as a measure of central tendency are more relevant than means for predicting a lognormal variate such as annual run by age. Predictions are still produced through addition as in Equation (1c) above.

These changes alone represent an improvement over the current method. Hindcasted predictions using the alternative method as so far described were on average 7% (= MPE) higher than their comparable estimates of run size with an expected MAPE of 33% (Table 4). Table 5 presents a simple example to demonstrate how the jackknifing and subsequent hindcasting were organized. The empirical distribution of IPEs was relatively flat (Figure 4, lower panel), partly the result of the IPEs having a trend in time (Figure 5) with predictions tending to be less than estimates early in the series and tending to be more later.

An additional difference in the alternative method involves the modeling of error. The linear form of a power function with generic notation is

$$\ln(y_{cy}) = \ln a + b\ln(x_{cy}) + \lambda_{cy} \tag{3}$$

where $\lambda_{cy} \sim norm(0,\sigma_{\lambda}^2)$ and represents what is commonly called "process error." The dependent variable y in Equation (3) is assumed to be known *without* error, which is not so when forecasting run size to the Taku River. Run size is estimated with measurement error such that the estimate $Y_{cy} = y_{cy} \exp(u_{cy})$ where $u_{cy} \sim norm(0,\sigma_u^2)$. Substituting $Y \rightarrow y$ in Equation (3) produces

$$\ln(Y_{cy}) = \ln a + b\ln(x_{cy}) + \lambda_{cy} + u_{cy}\,. \tag{4}$$

Measurement error u_{cy} is added to the right-hand side of the equation (3) to balance its incorporation into the left hand side resulting from the substitution. The terms λ_{cy} and u_{cy} collectively become the residuals in a fit of Equation (4) and represent both process and measurement error combined in the dependent variable such that $\varepsilon_{cy} \sim norm(0,\sigma_{\lambda}^2 + \sigma_u^2)$ in

Equation (2). However, the independent variable x is also an estimate and is known with error, such that estimate $X_{cy} = x_{cy}\ \exp(v_{cy})$ where $v_{cy} \sim\ norm(0, \sigma_v^2)$. Substitution of $X \rightarrow x$ into Equation (4) presents an error-in-variables problem that, if ignored, as was the case in Equations (1) and (2), could result in strongly biased estimates of b in regression (Fuller 1987) and subsequently biased predictions.

Bayesian analysis based on simulation was used to account for both measurement error in estimates and to express uncertainty in forecasts. In a Bayesian analysis, estimates of run size are treated as being known with certainty, whereas values of parameters and actual run sizes are considered to follow probability distributions (see Gelman et al. 1995, Carlin and Louis 2000 for explanations of Bayesian methods). Bayesian analysis begins with each parameter (i.e., the ln a, b, σ_ε^2, σ_v^2) and actual run size (i. e., the $R_{1.2}$ and $R_{1.3}$) each expressed as a "prior" probability distribution that represents current uncertainty in their values (Table 6). Simplistically expressed, the likelihoods of observing the estimates of run size are then calculated and multiplied by the prior distributions to get "posterior" probability distributions on parameters and functions of parameters, such as forecasts. Put another way, prior knowledge about a parameter is updated with data to reduce uncertainty.

With one exception, all prior distributions in our analysis (Table 6) are non-informative (often called flat priors). For those variables with flat priors, posterior distributions will reflect only the effect of likelihoods based on data. The one exception concerns the informative prior on measurement error of x—the $\sigma_v^2 \sim$ 1/gamma(c, d) where c and d are shape parameters. From Evans et al. (1993), $c = (\overline{\sigma}_v^2)^2 / \mathrm{V}(\sigma_v^2)$ and $d = \overline{\sigma}_v^2 / V(\sigma_v^2)$. Sampling variances $v(\hat{R}_{1.2,cy})$ and $v(\hat{R}_{1.3,cy})$ were used to approximate these shape parameters. Again from Evans et al. (1993), $\hat{\sigma}_{v,cy}^2 = \ln[cv^2(\hat{R}_{cy}) + 1]$ to transform estimated sampling variances to the log scale. The average and variance of the $\hat{\sigma}_{v,cy}^2$ from 1995 through 2013 were used as surrogates for $\overline{\sigma}_v^2$ and $V(\sigma_v^2)$ to calculate the shape parameters for this informative prior.

Uncertainty in forecasts was modeled in the alternative method with consideration of past IPEs. Equation (1c) provides a prediction of terminal run size but not the best forecast. As described above, the IPE is:

$$\mathrm{IPE} = \frac{\mathrm{Prediction} - \mathrm{Actual}}{\mathrm{Actual}} \times 100$$

Note that in "fitting" a time series of n years under the alternative method there will be $n - 2$ data triplets and $n - 2$ residuals with which to transform into $n - 2$ IPEs. Dropping percent in favor of a simple fraction, the equation above can be reformulated as:

$$(\mathrm{IPE}_{cy} + 1) = \frac{\mathrm{Prediction}_{cy}}{\mathrm{Actual}_{cy}} \tag{5}$$

with Actual$_{cy}$ being the known estimate of terminal run size (from field sampling), and with $\infty > (\mathrm{IPE}_{cy} + 1) \geq 0$. Note that Prediction$_{cy}$ is the sum of two lognormal variates $\widetilde{R}_{1.3,cy}$ and $\widetilde{R}_{1.4,cy}$, and is therefore assumed to be a lognormal variate itself, but with unknown mean and variance (Beaulieu and Xie 2004). As a result the (IPE$_{cy}$ + 1) is also a lognormal variate with mean $\mu_{(\mathrm{IPE}+1)}$

and variance $\sigma^2_{(IPE+1)}$. During each iteration (sample, update) in the Bayesian simulation of the alternative method, (IPE_{cy} + 1) is calculated for all *cy* except for 1995 and 1996, thereby providing $n - 2$ individual values with which to calculate a mean and a variance of the (IPE + 1). And that mean and variance were used to calculate values for $\mu_{(IPE+1)}$ and $\sigma^2_{(IPE+1)}$ for each iteration in the simulation with the methods from Evans et al. (1993) noted above. This allows the generation of a value ~ lognormal($\mu_{(IPE+1)}$, $\sigma^2_{(IPE+1)}$) in each iteration for the year to be forecasted, the calendar year one beyond the end of the data series (here that year is 2014). Rearranging Equation (5) to be used in the forecast for the year beyond the data:

$$\text{Actual}_{(2014)} = \frac{\text{Prediction}_{(2014)}}{(\text{IPE}_{(2014)} + 1) - 1} \tag{6}$$

The subtraction of 1 from the denominator in Equation (6) is to rescale the forecasted IPE such that $\infty > \text{IPE} \geq -1$ (or $\infty > \text{IPE} \geq -100\%$). The posterior distribution on the variable $\text{Actual}_{(2014)}$ is the probability distribution for the estimated terminal run size for the Taku stock of large Chinook salmon in 2014, and it is the forecast under the alternative method.

The computer program WinBUGS[4] version 1.4.2 based on Markov Chain Monte Carlo (MCMC) algorithm was used to simulate posterior distributions that are part of the alternative method as described above (see Appendix A1 for a listing of the code). Each simulation contained 3 chains, each with different sets of starting values for priors. Each simulation was iterated (updated) 480,000 times after a burn-in of 20,000 updates, providing 1,440,000 samples from the joint probability distribution for parameters (referred to as nodes in the WinBUGS language). Posterior distributions in longer simulations (up to 3 million samples) produced the same means in posterior distributions for forecasts as did the shorter simulations with 1.44 million samples.

Using the data from 1995 through 2013, the estimated terminal run size in 2014 ($\text{Actual}_{(2014)}$) has a 90% chance of being between 13,830 and 43,330 large salmon (Table 7 contains this credibility interval (CI) along with statistics for a subset of parameters in the Bayesian analysis). The mean of the posterior distribution of $\text{Actual}_{(2014)}$ is 25,980 and its median a bit lower at 24,440. As indicated by these results, the posterior is slightly skewed with a longer tail corresponding to higher numbers (Figure 6).

Because IPEs for the later years in the time series differed than those for earlier years (Figure 5), the Bayesian analysis was also run on more recent years. Using the data from 2003 through 2013, the estimated terminal run size in 2014 ($\text{Actual}_{(2014)}$) has a 90% chance of being between 13,560 and 33,390 large salmon (Table 8 contains this credibility interval along with statistics for a subset of parameters in the Bayesian analysis). The mean of the posterior distribution of $\text{Actual}_{(2014)}$ is 22,090 and its median a bit lower at 21,270. The lower bound of the 90% CI was about the same as with the full data set (13,560 vs. 13,830), but the upper bound was almost 10,000 fish lower (33,390 vs. 43,330). Subsequently, the mean of the posterior distribution was about 4,000 fish lower for the shorter series than the full series (22,090 vs. 25,980). The posterior distribution for the forecast from the shorter series is also skewed, but less so than for the full series (Figure 6).

---

[4] © Medical Research Council, Imperial College, London, U. K. 2007. Product names used in this report are included for scientific completeness but do not constitute a product endorsement.

DISCUSSION

To recap, the stated shortcomings of the current forecast method include the following:

- negative predictions for an age group are possible;
- residual error follows a normal distribution and is additive, but should follow a lognormal distribution and be multiplicative;
- not all brood year information is used in the forecast of age-1.4 fish;
- effect of measurement error in independent variables on both the 1.3 vs 1.2 and 1.4 vs 1.3 regressions is ignored; and
- uncertainty in forecasts is incompletely expressed.

No negative predictions for an age group are possible with the alternative method, and by design, residual error is modeled with lognormal distributions. The alternative method does include more information from each brood year than the current method, but there are other ways to handle additional information than as was done here. Obviously, separating abundance estimates for age-1.2 and age-1.3 salmon into two independent variables instead of one is another way. Preliminary work along those lines (not reported here) showed little benefit to keeping estimates separate instead of summing them. If in the future a better expression of sibling relationships is found, the alternative method can be easily modified.

The major improvements in the alternative method involve the Bayesian approach used in forecasting. Measurement error in data is modeled, and results have readily understandable expressions of uncertainty. For instance, the accepted escapement goal for the Taku River is the range from a lower goal (LG) of 19,000 to a higher goal (HG) of 36,000 large salmon a year (CTC 2014). By negotiation, the base level catch (BLC) is 6,400 fish per year shared between Canadian and U.S. fisheries. The midpoint of the escapement goal [(LG + HG)/2 = 27,500] plus the BLC constitutes the base terminal run (BTR), which is 33,900 fish. If the preseason forecast is greater than 33,900 there will be additional, directed fishing. The mean of the node **Prob.GT.BTR** expressed as a percent in Table 7 is the probability that in 2014 the estimated terminal run size will be greater than the BTR—put another way, the chance of having directed fishing in 2014. That chance is about 17%. Another relevant question is whether the run will be so low as to cut into the base level catch. The mean of the node **Prob.GT.LGBLC** from Table 7 indicates that there is a 46% chance of the estimated run in 2014 being above the LG by at least the BLC. And finally, the mean of the node **Prob.GT.LGBLC** from Table 7 indicates that there is a 23% chance that the terminal run will be less than the LG. These probability nodes were assigned either a 1 or zero for each update in the simulation whether or not the forecast (node **Actual**) was greater than a specified management target. The probability is the mean over the number of updates. A summary of these points for 2014 follows:

- chance of having directed fishing is 17%;
- chance of meeting at least the LG and realizing at least the BLC is 46%; and
- chance of not meeting the LG even with prohibition of all fishing is 23%.

All of these chances are based on simulations involving the full time series of estimated terminal run sizes from 1995 through 2013.

As with all forecasting, some decisions must be made as to the data used. Here we ignored data collected from 1977 through 1994 because they were too "noisy" to provide useful information. These early data were also, by definition, from a time that was perhaps not as relevant as more recent information. The unusual persistence of over-forecasts in later years demonstrated in Figure 5 is consistent with a shift in ocean survival that bears on our ability to forecast. The Bayesian forecast on the truncated time series bears this out (Figure 8). The posterior distribution from the truncated series is much more narrow than the distribution from the entire series. The consequence is a better, but different, forecast. Probabilities relative to management targets calculated from 2003 to 2013 data differ substantially with those reported above in this section:

- chance of having directed fishing is 5%;
- chance of meeting at least the LG and realizing at least the BLC is 26%; and
- chance of not meeting the LG even with prohibition of all fishing is 34%:

A one-third chance of not meeting the LG even with no fishing is a sobering prospect, even with the implication there is a two-thirds chance of meeting the LG with some fishing in 2014.

Management protocols for fisheries on Taku River Chinook salmon mandate a single number as a pre-season forecast. There are two logical choices: the mean or the median of the posterior probability distribution for the node **Actual**. The mean has been the traditional choice for forecasting in general; however, the median is attractive here because there is a 50/50 chance of the estimated run size being above or below it. In the analysis using the truncated data series, there is a difference of 820 fish between median and mean (see Table 8). In the analysis using the data series 1995–2013, the difference is 1,540 (see Table 7).

The uncertainty expressed in this analysis, or any like analysis, is conditioned on the variation seen in the past being representative of variation to be experienced in the future. Our decisions to exclude some of the available data were in part to make forecasts reflect contemporary variation. However, the future may hold surprises with the introduction of "black swans," events that by their very nature are unpredictable (Taleb 2007). Very anomalous events have happened with salmon runs before (see Quinn et al. 2007) and will obviously happen again. These anomalous events are more likely to involve large runs, but it is impossible to know in advance.

The WinBUGS program listed in the Appendix was coded to be easily updated to produce forecasts for 2015 and for years beyond. The program has three sections—model, data, and initial conditions. A simple extension of the analysis to forecast an additional year into the future (say a forecast for 2015) requires changes in the data and initial values sections only. Instructions for such an extension are given as comments in code between data and initial values sections listed in the Appendix. In the data section, new estimates of run size by age for 2014 would be added to the end of vectors **R12**, **R13**, and **R14;** and the sample size **n** would be incremented by one year. Values of **c1**, **d1**, **c2**, and **d2** may be updated as well, but such updates are not necessary in all years. How to calculate **c1**, **d1**, **c2**, and **d2** is described above. In the initial values section, vectors of initial values **xR12** and **xR13** need to be extended one element each by repeating the previous element or by continuing the pattern by one element. Such extensions are necessary for both vectors in each of three chains (sets of initial values).

ACKNOWLEDGMENTS

The authors thank Paul Kissner, formerly of ADF&G, for implementing the stock assessment program for Chinook salmon on the Taku River in the early 1970s. Sandy Johnston and Pat Milligan of DFO and Scott McPherson, Andy McGregor, and Scott Kelley of ADF&G helped develop the Chinook salmon stock assessment program on the Taku River. We thank Phil Timpany of TRTFN for his 30 years of participation in Chinook salmon stock assessment on the Taku River in various capacities for ADF&G, TRTFN, and DFO. We thank Keith Pahlke for assistance in field sampling and project design in the 1990s, and completion of the 1989 and 1990 mark–recapture analyses. We thank Dan Reed for analysis assistance and reporting of the 1999–2002 and 2004–2007 mark–recapture studies. We thank the following project leaders for their efforts in stock assessment: Richard Yanusz and Dave Magnus (ADF&G); Richard Erhardt (TRTFN); contractor Brian Mercer (DFO); crew leaders Heather Stilwell, Jerry Owens, James Lobdell, Dale Brandenburger, Scott Duffy, Cheryl Dion, Mike Lafollette, David Dreyer, and Aaron Vochowski (ADF&G); Patrick Jackson, Rick Ferguson, Marty Strachan, Sean Stark, Scott Herron, Matt Waugh, Al Mcleod, Tyler Rudolph, and Lars Jessup (DFO); Mark Connor, Terry Jack, Mike Smarch, and Lino Battaja (TRTFN). We thank Al Demartini, Mark Olsen, Jarbo Crete, Shane Rear, Krista Kissner, Jaime Kissner, Christie Hendrich, Kercia Schroeder (Rear), Nicole Zeiser, Lee Close, Sidon Ranger, Wayne Lonn, and Roger Hayward (ADF&G), and Kirstie Falkevitch, Mark McFarland, and Christian and Mathieu Decharme (DFO) for smolt tagging since 1993. We thank Clyde Andrews and Norm Miller for expediting supplies to the crews. We thank Ward Air and Coastal Helicopters air charter and Taku River Dirtbags barge services for transporting supplies. Scott McPherson, Sue Millard, Mark Olsen, and Darlene Gillespie, along with and other personnel from the DFO Nanaimo Lab, performed most of the scale age analyses over the years. We are grateful to Stacey Poulson for helping with editorial and final preparation of this manuscript and to two anonymous reviewers for their insightful comments.

REFERENCES CITED

Beaulieu, N., and Q. Xie. 2004. An optimal lognormal approximation to lognormal sum distributions. Vehicular Technology, IEEE Transactions 53(2): 479-489.

Carlin, B. P., and T. A. Louis. 2000. Bayes and empirical Bayes methods for data analysis. Chapman & Hall, New York.

CTC (Chinook Technical Committee). 2014. Annual Report of Catch and Escapement for 2013. Pacific Salmon Commission, Report TCCHINOOK (14)–2. Vancouver, BC.

Efron, E., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, New York.

Evans, M., N. Hastings, and B. Peacock. 1995. Statistical distributions. 2nd ed. Wiley Interscience, New York.

Fuller, W. A. 1987. Measurement error models. John Wiley & Sons, New York.

Jennings, G. B., K. Sundet, and A. E. Bingham. 2010. Estimates of participation, catch, and harvest in Alaska sport fisheries during 2008. Alaska Department of Fish and Game, Fishery Data Series No. 10-22, Anchorage.

Jones III, E. L., S. A. McPherson, D. J. Reed, and I. M. Boyce. 2010. Spawning abundance of Chinook salmon in the Taku River from 1997 to 2007, Alaska Department of Fish and Game, Fishery Data Series No. 10-70, Anchorage.

Gelman, A., J. B. Carlin, H. S. Stern, and D. R. Rubin. 1995. Bayesian data analysis. Chapman & Hall, New York.

McPherson, S. A, E. L. Jones III, S. J. Fleischman, and I. M. Boyce. 2010. Optimal production of Chinook salmon from the Taku River through the 2001 year class. Alaska Department of Fish and Game, Fishery Manuscript No. 10-3, Anchorage.

McPherson, S. A., D. R. Bernard, and J. H. Clark. 2000. Optimal production of Chinook salmon from the Taku River. Alaska Department of Fish and Game, Fishery Manuscript No. 00-2, Anchorage.

McPherson, S. A., D. R. Bernard, R. J. Yanusz, P. A. Milligan, and P. Timpany. 1999. Spawning abundance of Chinook salmon in the Taku River in 1998. Alaska Department of Fish and Game, Fishery Data Series No. 99-26, Anchorage.

McPherson, S. A., D. R. Bernard, M. S. Kelley, P. A. Milligan, and P. Timpany. 1998. Spawning abundance of Chinook salmon in the Taku River in 1997. Alaska Department of Fish and Game, Fishery Data Series No. 98-41, Anchorage.

McPherson, S. A., D. R. Bernard, M. S. Kelley, P. A. Milligan, and P. Timpany. 1997. Spawning abundance of Chinook salmon in the Taku River in 1996. Alaska Department of Fish and Game, Fishery Data Series No. 97-14, Anchorage.

McPherson, S. A., D. R. Bernard, M. S. Kelley, P. A. Milligan, and P. Timpany. 1996. Spawning abundance of Chinook salmon in the Taku River in 1995. Alaska Department of Fish and Game, Fishery Data Series No. 96-36, Anchorage.

Pahlke, K. A., and D. R. Bernard. 1996. Abundance of the Chinook salmon escapement in the Taku River, 1989 to 1990. Alaska Fishery Research Bulletin 3(1):8-19, Juneau.

Quinn, T. P., D. M. Eggers, J. H. Clark, and H.B. Rich, Jr. 2007. Density, climate, and the processes of prespawning mortality and egg retention in Pacific salmon (*Oncorhynchus* spp.). Canadian Journal Of Fisheries and Aquatic Sciences 64:574-582.

Taleb, N. N. 2007. The black swan: the impact of the highly improbable. Penguin Books, New York.

TABLES AND FIGURES

Table 1.–Estimated run size by age for the Taku stock of Chinook salmon from calendar years 1995-2013.

	Run size by age				Run size ages	Standard errors by age				SE ages
CY	1.2	1.3	1.4	1.5	1.3-1.5	1.2	1.3	1.4	1.5	1.3-1.5
1995	34,471	16,559	22,105	753	39,417	3,870	2,001	2,610	180	3,294
1996	8,685	79,109	10,969	214	90,291	1,230	7,721	1,324	89	7,834
1997	2,880	47,381	78,242	0	125,623	641	6,605	11,128	0	12,941
1998	8,763	10,156	22,721	860	33,737	2,006	2,902	7,132	271	7,704
1999	12,009	14,139	4,538	253	18,930	1,476	2,212	711	105	2,326
2000	10,778	27,455	10,300	104	37,859	1,768	3,748	1,484	62	4,031
2001	5,918	40,142	10,729	80	50,952	908	5,192	1,443	30	5,389
2002	8,221	36,649	23,332	247	60,228	1,139	6,529	4,248	82	7,790
2003	18,842	25,388	15,565	131	41,085	1,994	3,984	2,371	57	4,637
2004	27,462	62,156	15,562	330	78,049	2,318	7,399	1,971	112	7,658
2005	10,123	48,274	18,654	194	67,122	841	3,368	1,312	66	3,615
2006	4,363	29,536	31,482	467	61,484	537	2,690	2,779	100	3,870
2007	8,348	10,842	7,434	283	18,559	1,481	1,784	1,335	119	2,231
2008	13,888	26,737	4,851	31	31,619	999	1,749	411	24	1,797
2009	12,627	24,175	9,115	114	33,404	1,240	1,761	650	61	1,878
2010	9,098	27,123	6,712	488	34,322	798	1,847	494	12	1,912
2011	15,216	27,028	5,700	143	32,872	1,680	2,824	541	18	2,875
2012	5,165	18,162	6,999	61	25,223	836	1,632	679	44	1,769
2013	5,646	16,379	6,060	88	22,526	835	1,625	677	45	1,761

Note: Age is expressed in European notation.

Note: McPherson et al. (2010) provide a general description of how estimates were derived.

Note: Run size is estimated as the sum of estimated spawning abundance, Canadian in-river harvest, harvest in the U.S. terminal gillnet fishery (Alaska District 111), and estimated harvest of the Taku stock in the U.S. sport fishery (Juneau marine creel). Estimates in marine fisheries are based on catch sampling to recover coded wire tags.

Table 2.–Example of calendar years (CY) removed, run size forecasted, data pairs no longer complete and dropped, and complete pairs used to build jackknife regressions for hindcasting from a time series of 5 years of paired data with the current method.

Remove	Forecast	Dropped Pairs	Pairs Used in Jackknife Regression
CY1	No Forecast Possible		
CY2	CY2 from CY1	(CY2,CY1) and (CY3,CY2)	(CY4,CY3) and (CY5,CY4)
CY3	CY3 from CY2	(CY3,CY2) and (CY4,CY3)	(CY2,CY1) and (CY5,CY4)
CY4	CY4 from CY3	(CY4,CY3) and (CY5,CY4)	(CY2,CY1) and (CY3,CY2)
CY5	CY5 from CY4	(CY5,CY4)	(CY2,CY1), (CY3,CY2), and (CY4,CY3)

Note: With a time series of n years, there will be $n - 1$ data pairs; dropping a year would therefore affect 2 years, leaving $n - 3$ data pairs to jackknife in all but the last year ($n - 2$ pairs left).

Table 3.–Mean percent error (MPE) and mean absolute error (MAPE) from hindcasting with the current method of forecasting the run size of large (age-1.3–1.5) Chinook salmon to the Taku River.

CY	Regression 1 1.3 vs. 1.2		Regression 2 1.4 vs. 1.3		Predictions			Estimated	% Error	
excluded	Slope	Intercept	Slope	Intercept	Age 1.3	Age 1.4	Total	Total	IPE	Absolute
1996	1.562	10,810	0.439	-550	64,665	6,715	71,380	90,077	-21%	21%
1997	1.827	8,641	0.405	497	24,506	32,539	57,045	125,623	-55%	55%
1998	1.688	12,511	0.845	-10,746	17,373	29,268	46,641	32,877	42%	42%
1999	1.749	11,641	0.865	-12,417	26,968	-3,634	23,334	18,677	25%	25%
2000	1.801	9,611	0.828	-10,664	31,235	1,048	32,283	37,755	-14%	14%
2001	1.905	7,042	0.798	-9,180	27,578	12,732	40,310	50,872	-21%	21%
2002	1.894	7,804	0.804	-9,208	19,013	23,070	42,083	59,981	-30%	30%
2003	1.666	10,444	0.810	-9,638	24,139	20,032	44,171	40,953	8%	8%
2004	1.833	9,236	0.956	-12,977	43,773	11,304	55,077	77,719	-29%	29%
2005	2.002	8,379	0.961	-12,731	63,367	46,989	110,356	66,928	65%	65%
2006	1.733	11,230	0.792	-8,771	28,777	29,439	58,216	61,018	-5%	5%
2007	1.736	11,180	0.819	-9,886	18,753	14,319	33,072	18,276	81%	81%
2008	1.811	10,465	0.821	-10,180	25,584	-1,280	24,304	31,587	-23%	23%
2009	1.811	10,921	0.792	-8,683	36,078	12,500	48,578	33,290	46%	46%
2010	1.792	10,440	0.789	-8,340	33,065	10,732	43,797	33,834	29%	29%
2011	1.846	10,633	0.789	-8,224	27,432	13,178	40,610	32,729	24%	24%
2012	1.822	11,149	0.797	-8,947	38,867	12,593	51,461	25,161	105%	105%
2013	1.767	10,615	0.803	-9,461	19,740	5,117	24,857	22,438	11%	11%
									MPE = 13%	
									MAPE = 35%	

Note: Sibling regressions are based on abundance by age within the same brood year.

Note: The IPE = (Prediction – Estimated)/Estimated x 100.

Note: Age-1.5 Chinook salmon were not used in the calculations.

Note: CY = calendar year; grey shading indicates negative predictions.

Table 4.–Mean error (MPE) and mean absolute error (MAPE) from hindcasting with the alternative method (measurement error unaddressed) of forecasting the run size of large (age-1.3–1.5) Chinook salmon to the Taku River.

CY	Regression 1 ln(1.3) vs. ln(1.2)		Regression 2 ln(1.4) vs. ln(1.3+1.2)		Predictions			Estimated	% Error	
excluded	Slope	Intercept	Slope	Intercept	Age 1.3	Age 1.4	Total	Total	IPE	Absolute
1997	0.602	4.622	0.962	-0.903	23,869	29,733	53,602	125,623	-57%	57%
1998	0.545	5.231	1.375	-5.275	14,318	17,388	31,707	32,877	-4%	4%
1999	0.637	4.377	1.396	-5.478	25,739	2,324	28,063	18,677	50%	50%
2000	0.631	4.370	1.203	-3.413	29,555	5,795	35,350	37,755	-6%	6%
2001	0.697	3.715	1.134	-2.663	26,494	11,400	37,894	50,872	-26%	26%
2002	0.711	3.598	1.149	-2.862	17,633	14,692	32,325	59,981	-46%	46%
2003	0.564	4.961	1.237	-3.732	23,036	12,761	35,797	40,953	-13%	13%
2004	0.539	5.186	1.208	-3.423	36,133	9,552	45,685	77,719	-41%	41%
2005	0.652	4.189	1.213	-3.413	51,821	29,682	81,503	66,928	22%	22%
2006	0.564	5.019	1.180	-3.107	27,484	25,512	52,996	61,018	-13%	13%
2007	0.568	4.984	1.206	-3.363	17,033	12,156	29,189	18,276	60%	60%
2008	0.668	4.059	1.199	-3.286	24,149	3,867	28,016	31,587	-11%	11%
2009	0.679	3.980	1.153	-2.730	34,690	11,377	46,067	33,290	38%	38%
2010	0.647	4.241	1.149	-2.662	31,353	12,769	44,122	33,834	30%	30%
2011	0.712	3.682	1.133	-2.496	26,259	13,376	39,635	32,729	21%	21%
2012	0.702	3.781	1.134	-2.565	37,942	11,339	49,280	25,161	96%	96%
2013	0.628	4.421	1.144	-2.698	17,788	10,041	27,829	22,438	24%	24%
Reg	0.638	4.322	1.161	-2.912					MPE = 7%	
									MAPE = 33%	

Note: Sibling regressions are based on abundance by age within the same brood year.

Note: The IPE = (Prediction – Estimated)/Estimated x 100.

Note: Age-1.5 Chinook salmon were not used in the calculations.

Note: CY = calendar year.

Table 5.–Example of years (CY) removed, run size forecasted, data triplets no longer complete and dropped, and complete triplets used to build hindcasting jackknife regressions for a time series of 7 years of data.

Remove	Forecast	Dropped Triplets	Triplets in Jackknife Regression
CY1	No Forecast Possible		
CY2	No Forecast Possible		
CY3	CY3 from (CY2,CY1)	(CY3,CY2,CY1),(CY4,CY3,CY2), (CY5,CY4,CY3)	(CY7,CY6,CY5) and (CY6,CY5,CY4)
CY4	CY4 from (CY3,CY2)	(CY4,CY3,CY2),(CY5,CY4,CY3), (CY6,CY5,CY4)	(CY3,CY2,CY1) and (CY7,CY6,CY4)
CY5	CY5 from (CY4,CY3)	(CY5,CY4,CY3),(CY6,CY5,CY4), (CY7,CY6,CY5)	(CY4,CY3,CY2) and (CY3,CY2,CY1)
CY6	CY6 from (CY5,CY4)	(CY6,CY5,CY4) and (CY7,CY6,CY5)	(CY5,CY4,CY3),(CY4,CY3,CY2), (CY3,CY2,CY1)
CY7	CY7 from (CY6,CY5)	(CY7,CY6,CY5)	(CY6,CY5,CY4),(CY5,CY4,CY3), (CY4,CY3,CY2),(CY3,Y2,Y1)

Note: With a time series of *n* years there will be *n* – 2 data triplets under the alternative method; dropping a year would therefore affect 3 years, leaving *n* – 5 data triplets to jackknife in all but the last (*n* – 4 triplets left) and penultimate years (*n* – 3 triplets left).

Table 6.–Prior probability distributions, likelihoods, and data inputs used in the Bayesian analysis involving forecasts under the alternative method.

Prior probability distributions:	Comments:
lna1,lna2 ~ uniform(-10,10)	Flat priors for intercept in each transformed sibling regression
b1,b2 ~ uniform(0,2)	Flat priors for slope in each transformed sibling regression
tau.Y1,tau.Y2 ~ gamma(0.001,0.001)	Flat priors representing precision due to combined process and measurement error in the dependent estimates for each sibling regression. Note that precision is defined as $1/\sigma^2$.
tau.R12,tau.R13 ~ gamma(c,d)	Informative priors representing precision due to measurement error in the independent estimates for each sibling regression. Note that precision is defined as $1/\sigma^2$.
xR12[cy], xR13[cy] ~ uniform(0,20)	Flat priors for the logs of the actual values of terminal run size by age. Note the actual values are considered as variables here in this analysis.
Likelihoods:	
R12[cy] ~ lognormal(xR12[cy],tau.R12)	Estimates from sampling programs of terminal run size by age (the Rs here) are considered to be known with certainty, whereas the means of the lognormal distributions (xR or mu) are variables.
R13[cy] ~ lognormal(xR13[cy],tau.R13)	
R13[cy] ~ lognormal(mu1[cy],tau.Y1)	The estimate R13[cy] is involved with two likelihoods—one to express measurement error and the other to express uncertainty in regression parameters ln a and b.
R14[cy] ~ lognormal(mu2[cy],tau.Y2)	

Data inputs:	
n	Number of years in the time series.
R12[cy]	Estimated number of Chinook salmon of age 1.2 in the terminal run in year cy.
R13[cy]	Estimated number of Chinook salmon of age 1.3 in the terminal run in year cy.
R14[cy]	Estimated number of Chinook salmon of age 1.4 in the terminal run in year cy.
c1,d1	Shape parameters to define the informative prior for measurement error in R12.
c2,d2	Shape parameters to define the informative prior for measurement error in R13.

Note: Notation is defined in the text or is similar to identifiers in WinBUGS code (see Appendix).

Table 7.–Statistics on posterior distributions for a subset of variables (nodes) from a simulation run with WinBUGS on years 1995–2013 with a forecast for estimated terminal run size in 2014 of large Chinook salmon (age-1.3 and age-1.4) to the Taku River (Actual).

Node (variable)	Mean	SD	Median	Lower CI 90%	Higher CI 90%	MC Error [a]
Actual	25,980	9,341	24,440	13,830	43,330	39.6
b1	0.639	0.190	0.638	0.323	0.948	0.00299
b2	1.160	0.219	1.163	0.793	1.514	0.00374
lna1	4.313	1.739	4.322	1.482	7.207	0.02743
lna2	-2.896	2.316	-2.936	-6.654	0.982	0.03957
MPE	0.066	0.085	0.062	-0.066	0.211	0.00019
Prob.GT.BTR × 100	17.2%	–	–	–	–	0.00118
Prob.GT.LGBLC × 100	45.6%	–	–	–	–	0.00170
Prob.LT.LG × 100	23.3%	–	–	–	–	0.00133

[a] MC Error is the standard error of the mean of the posterior distribution for each node (variable) and should be < 1% of the standard deviation for that distribution in a successful simulation.

Table 8.–Statistics on posterior distributions for a subset of variables (nodes) from a simulation run with WinBUGS on years 2003–2013 with a forecast for estimated terminal run size in 2014 of large Chinook salmon (age-1.3 and age-1.4) to the Taku River (Actual).

Node (variable)	Mean	SD	Median	Lower CI 90%	Higher CI 90%	MC Error[a]
Actual	22,090	6,207	21,270	13,560	33,390	37.75
b1	0.602	0.200	0.597	0.283	0.932	0.00320
b2	1.059	0.281	1.066	0.579	1.517	0.00501
lna1	4.505	1.850	4.552	1.449	7.458	0.02957
lna2	-2.126	2.979	-2.199	-6.980	2.959	0.05303
MPE	0.031	0.091	0.026	-0.106	0.184	0.00016
Prob.GT.BTR × 100	4.5%	–	–	–	–	0.00070
Prob.GT.LGBLC × 100	66.3%	–	–	–	–	0.00221
Prob.LT.LG × 100	33.7%	–	–	–	–	0.00221

[a] MC Error is the standard error of the mean of the posterior distribution for each node (variable) and should be < 1% of the standard deviation for that distribution in a successful simulation.

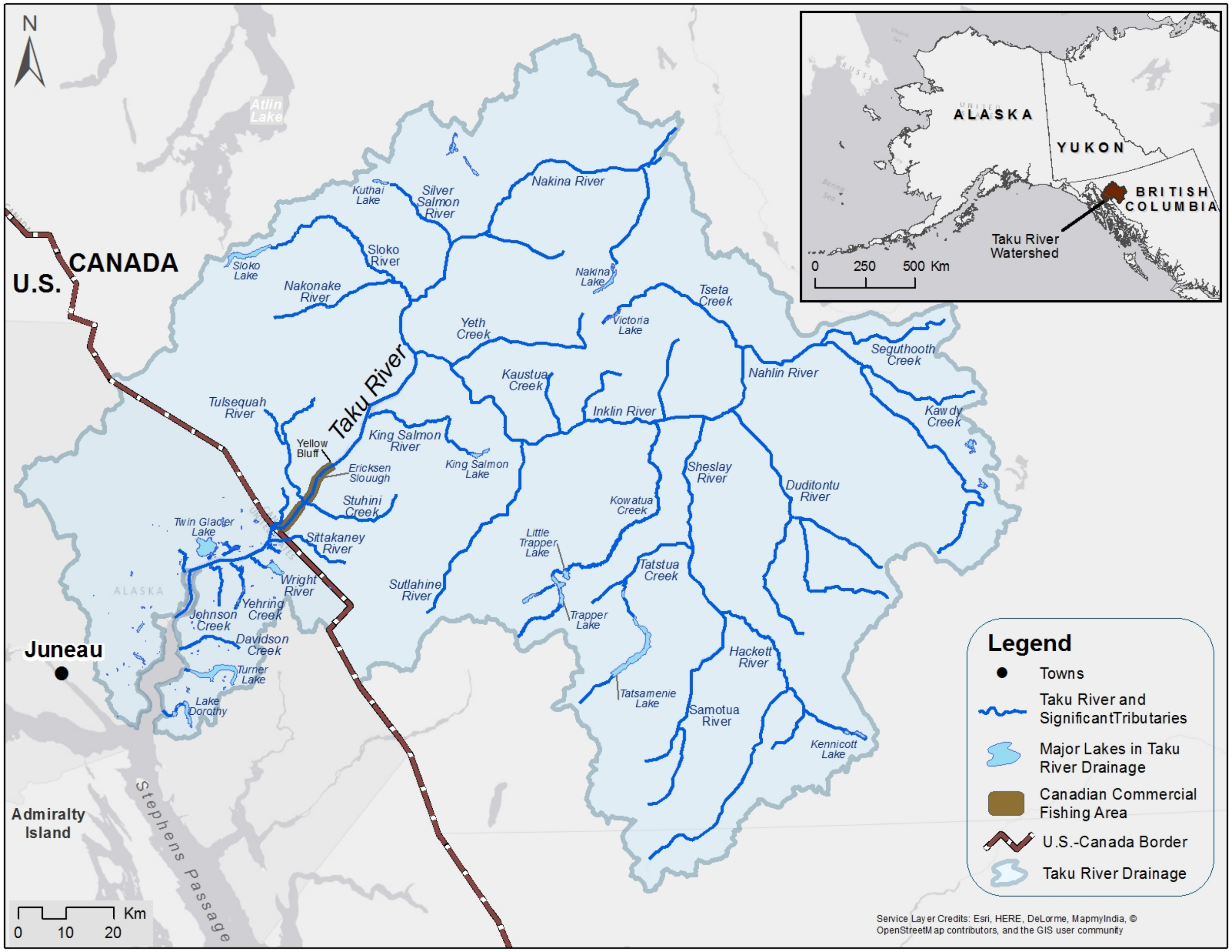

Figure 1.–Taku River watershed and significant tributary systems in northern Southeast Alaska and northwestern British Columbia.

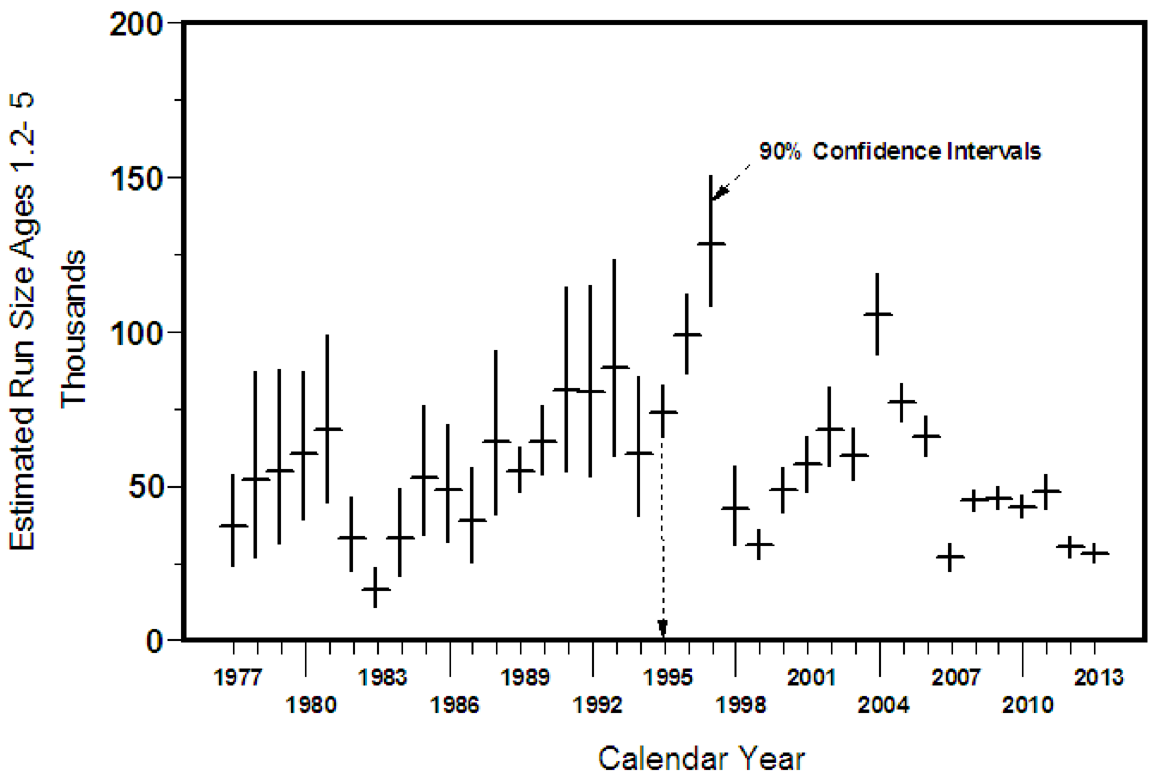

Figure 2.–Estimated annual run size for age-1.2–1.5 Chinook salmon of the Taku River, 1977-2013. Vertical bars represent 90% confidence intervals with positions of horizontal bars representing point estimates. Only estimates from 1995 (flagged by the arrow) through 2013 were used in our analysis.

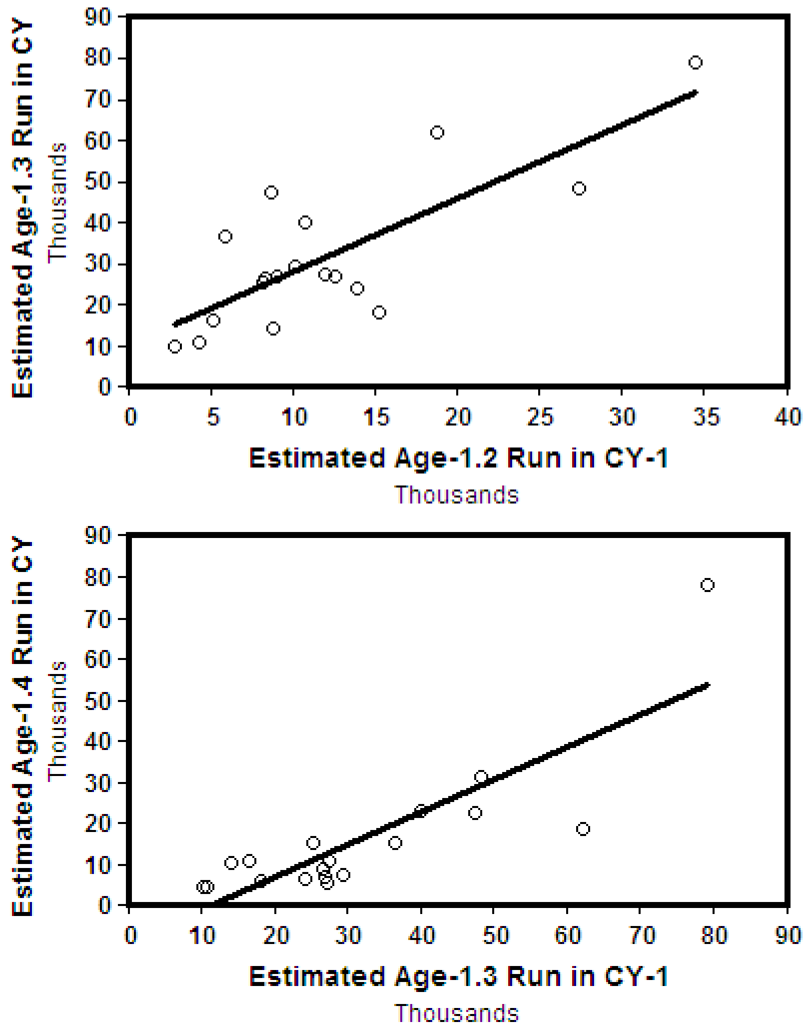

Figure 3.–Regressions (solid line) and data (circles) between run size of age-1.3 and age-1.2 Chinook salmon within the same brood year (top panel) and between run size of age-1.4 and age-1.3 Chinook salmon within the same brood year (bottom panel). Both regressions were based on estimated run sizes from 1995 through 2013 for the Taku stock. These regressions form the basis of the current method. CY = calendar year.

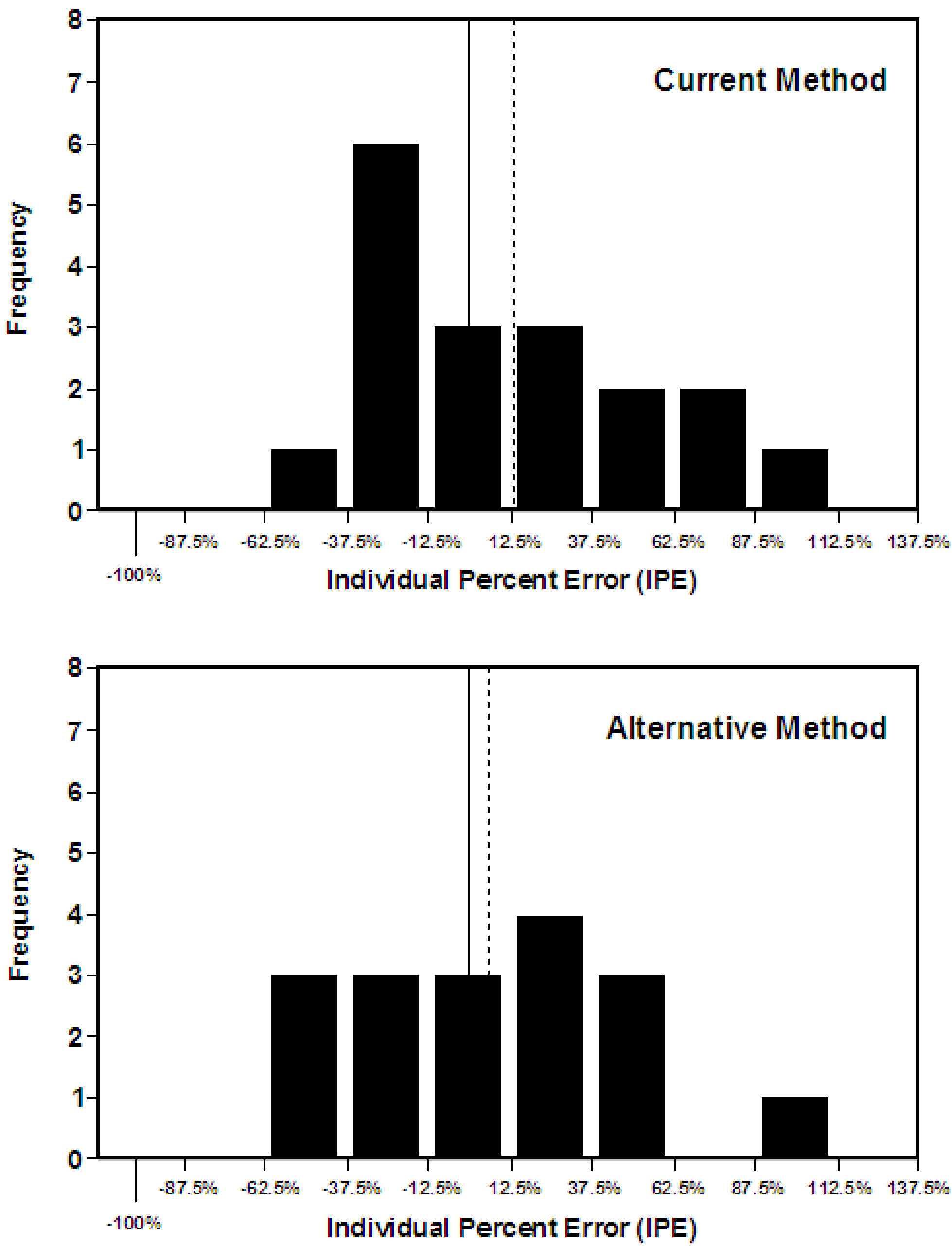

Figure 4.–Comparisons of individual percent error (IPE) in hindcast jackknifed forecasts of the annual run size of large (age-1.3–1.5) Chinook salmon of the Taku River from the current method (top panel) and from the alternative method (lower panel). Solid vertical lines designate IPE = 0, and dashed vertical lines MPE (13% for the current method and 7% for the alternative). Results in both panels are from methods that do not address measurement error in estimates of terminal run size.

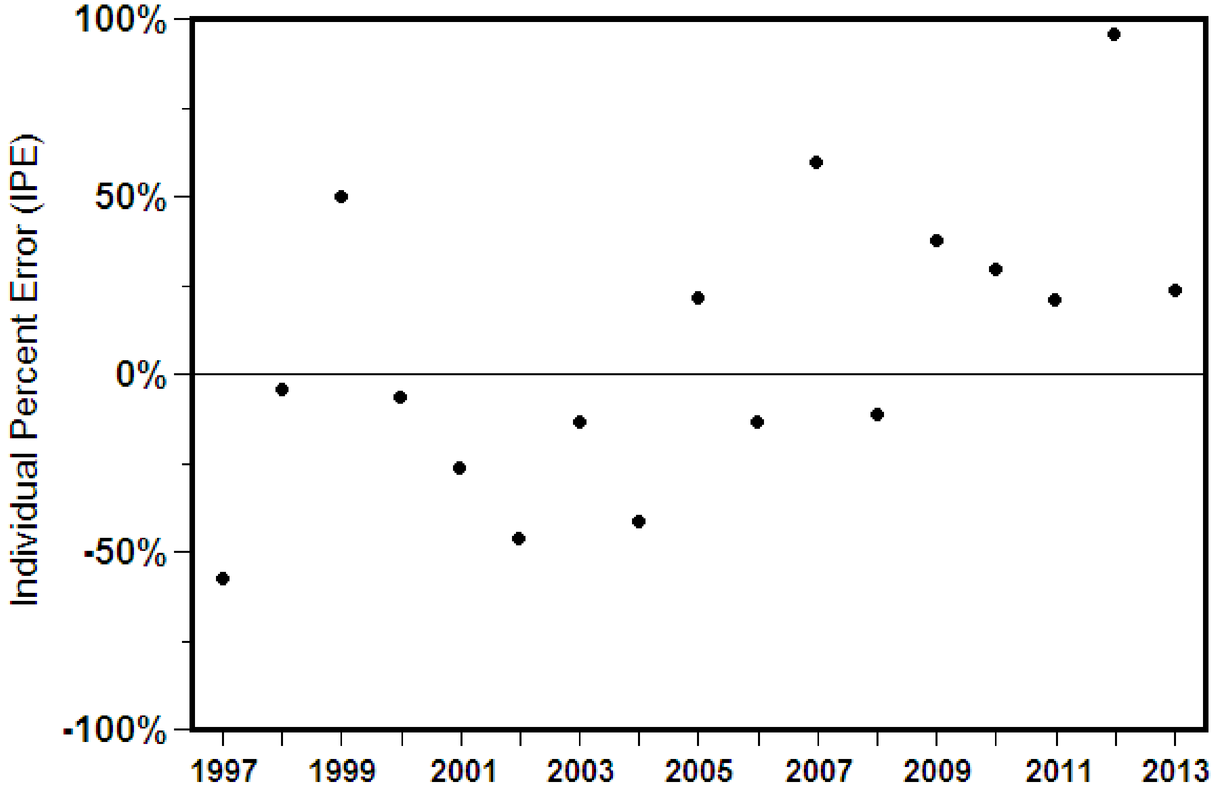

Figure 5.–Individual percent errors (IPEs) for hindcast, jackknifed forecasts of terminal run size to the Taku River of large Chinook salmon from the alternative method. Each forecast was the sum of median projections from two sibling regressions, each using a power function to forecast age-1.4 or age-1.3 salmon returning in the excluded year.

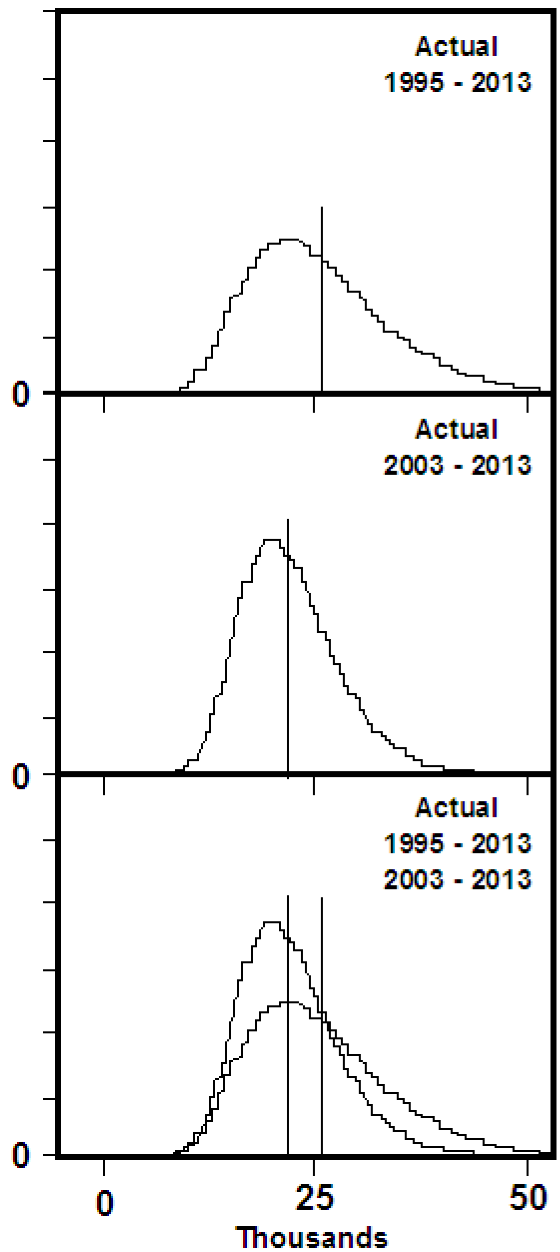

Figure 6.–Kernel densities for posterior probability distributions for the node Actual, the unknown variable that represents the estimated terminal run size in the year yet to come—here 2014. Top panel is the density from data collected from 1995 through 2013; middle panel is the density from data collected from 2003 through 2013; and the bottom panel contains both densities for comparison. Vertical lines represent the means for the longer data series (25,980 large, age-1.3 and age-1.4 Chinook salmon) and for the truncated series (22,090).

APPENDIX A

Appendix A1.–WinBUGS (Version 1.4.2) code to model sibling relationships under the alternative method to forecast terminal run size of large Chinook salmon to the transboundary Taku River of the US (Alaska) and Canada.

```
model {
# --------------------set single priors for regressions
lna1 ~ dunif(-10,10)
lna2 ~ dunif(-10,10)
b1 ~ dunif(0,2)
b2 ~ dunif(0,2)
tau.R12 ~ dgamma(c12,d12)      #  tau.R12 (and tau.R13) represent measurement
tau.R13 ~ dgamma(c13,d13)      #      error in estimates of abundance
tau.Y1 ~ dgamma(0.001,0.001)   # tau.Y1 (and tau.Y2) represents combined measurement and
tau.Y2 ~ dgamma(0.001,0.001)   #      process error in the dependent variable in a sibling regression
for(j in 1:n) {                # j represents CY 1995 thru 2013
# --------------------set priors for data series for actual abundance of independent variables
xR12[j] ~ dunif(0,20)
xR13[j] ~ dunif(0,20)
}
# --------------------regression to predict 1.3s from 1.2s
R12[1] ~ dlnorm(xR12[1],tau.R12)
R13[1] ~ dlnorm(xR13[1],tau.R13)                     # not needed, but completes array
for(j in 3:n) {                                      #  j represents CY 1997 thru 2013
R12[j-1] ~ dlnorm(xR12[j-1],tau.R12)
mu1[j-2] <- lna1+b1*log(R12[j-1])                    #  j-2 scales mu1 array to begin with 1
R13[j] ~ dlnorm(mu1[j-2],tau.Y1)
# --------------------regression to predict 1.4s from 1.2s and 1.3s
R13[j-1] ~ dlnorm(xR13[j-1],tau.R13)
mu2[j-2] <- lna2+b2*log(R12[j-2] + R13[j-1])         #  j-2 scales mu2 array to begin with 1
R14[j] ~ dlnorm(mu2[j-2],tau.Y2)
# --------------------calculation of IPEs
Prd[j-2] <-exp(mu1[j-2])+exp(mu2[j-2])
Est[j-2] <- R14[j]+R13[j]
IPE[j-2] <- (Prd[j-2]-Est[j-2])/Est[j-2]
}
# --------------------use observed IPEs (scaled) to get parameters for their lognormal distribution
MPE <- mean(IPE[])
MPE.1  <- MPE+1                                      #  Scale IPEs to begin at 0 instead of -1
IPE.sd <- sd(IPE[])
IPE.cv <- IPE.sd/MPE.1
sig2.IPE <- log(pow(IPE.cv,2)+1)
IPE.mu <- log(MPE.1)-sig2.IPE/2
tau.IPE <- 1/sig2.IPE
# --------------------generate posterior distribution for estimated run size 1 year beyond series
IPE.t ~ dlnorm(IPE.mu,tau.IPE)
R13.t <- exp(lna1+b1*log(R12[n]))
R14.t <- exp(lna2+b2*log(R12[n-1]+R13[n]))
Actual <- (R13.t + R14.t)/IPE.t                      #  +1 in denominator dropped to unscale IPEs
```

-continued-

```
# --------------------probabilities of management  actions given posterior for estimated run size 1 year beyond series
Prob.GT.BTR <- step(Actual-33901)                    # counter for probability having directed fishing
Prob.GT.LGBLC <- step(Actual-25401)                  # counter for probability of not compromising base level catch
Prob.LT.LG <- step(19000-Actual)                     # counter for probability of not having fishing at all
}
```

data:
⇨
```
list(n =19, c12 = 2.23051, d12 =114.90877, c13 = 0.95114, d13 = 48.93510,
R12 = c(34471, 8685, 2880, 8763, 12009, 10778, 5918, 8221, 18842, 27462, 10123, 4363, 8348, 13888, 12627,
9098, 15216, 5165, 5646),
R13 = c(16559, 79109, 47381, 10156, 14139, 27455, 40142, 36649, 25388, 62156, 48274, 29536,10842, 26737,
24175, 27123, 27028, 18162, 16379),
R14 = c(22105, 10969, 78242, 22721, 4538, 10300, 10729, 23332, 15565, 15562, 18654, 31482, 7434, 4851, 9115,
6712, 5700, 6999, 6060))
```
⇦
```
# ----------------------------To add another year, just tack the
#                               relevant statistics on the end of
#                               R12, R13, and R14, and increment
#                               n by +1 in the data section.
#
#                               If desired (but not needed), update
#                               the gamma distribution shape
#                               parameters to produce informative
#                               priors for measurement error:
#                               c12 and d12 for age-1.2 estimates,
#                               c13 and d13 for age-1.3 estimates.
#
#                               In the initial values section, just
#                               extend the sequence in xR12 and xR13
#                               by an additional element in each vector
#                               of initial values.
# ---------------------------------------------------------------------------
```
initial values:

inits:⇨
```
list(lna1=1.0, lna2=1.0, b1=1.0, b2=1.0, tau.Y1=4.0, tau.Y2=4.0, tau.R12=0.02, tau.R13=0.02,IPE.t=1.0,
xR12 = c(10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10.),
xR13 = c(10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10., 10.))

list(lna1=2.0, lna2=2.0, b1=1.5, b2=1.5, tau.Y1=3.0, tau.Y2=3.0, tau.R12=0.03, tau.R13=0.03,IPE.t=0.5,
xR12 = c(11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11.),
xR13 = c(11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11., 11.))

list(lna1=0.5, lna2=0.5, b1=0.5, b2=0.5, tau.Y1=2.0, tau.Y2=2.0, tau.R12=0.01, tau.R13=0.01,IPE.t=1.5,
xR12 = c(10., 9., 10., 9., 10., 9., 10., 9., 10., 9., 10., 9., 10., 9.,10., 9., 10., 9., 10.),
xR13 = c(10., 9., 10., 9., 10., 9., 10., 9., 10., 9., 10., 9., 10., 9.,10., 9., 10., 9., 10.))
```
⇦

Note: WinBUGS code developed by D.R. Bernard of D.R. Bernard Consulting LLC, 17 November, 2014. Prior distributions, likelihoods, and data in the code are described in Table 6.

Note: Prior distributions, likelihoods, and data in the code are described in Table 6. Nodes describing probability of management actions (Prob.) are described in the Discussion Section.

Appendix A2.–Computer data file used in the analyses described in this report.

File name	Description
Taku Chin Production_73-13.xls	Computer file containing the estimated production by year and age of Chinook salmon bound for the Taku River, 1973 to 2013

Note: Electronic data file is available upon request from Research and Technical Services (RTS) Publication archives, ADF&G, Sport Fish Division, 333 Raspberry Road, Anchorage, AK 99518-1565.

Appendix A2.–Computer data file used in the analyses described in this report.

File name	Description
Taku Chin Production_73-13.xls	Computer file containing the estimated production by year and age of Chinook salmon bound for the Taku River, 1973 to 2013

Note: Electronic data file is available upon request from Research and Technical Services (RTS) Publication archives, ADF&G, Sport Fish Division, 333 Raspberry Road, Anchorage, AK 99518-1565.