

## Inseason Forecasting of Southeastern Alaska Pink Salmon Abundance Based on Sex Ratios and Commercial Catch and Effort Data

JIE ZHENG\*

Alaska Department of Fish and Game  
Commercial Fisheries Division  
Post Office Box 25526, Juneau, Alaska 99802-5526, USA

OLE A. MATHISEN

University of Alaska Fairbanks, Juneau Center for Fisheries and Ocean Sciences  
11120 Glacier Highway, Juneau, Alaska 99801, USA

**Abstract.**—We developed a sex ratio index that, together with cumulative catch of all gears or cumulative catch per unit effort of the seine fishery, estimated abundance and catch of pink salmon *Oncorhynchus gorbuscha* in southern Southeast Alaska during a fishing season. We evaluated three inseason forecast models—linear, nonlinear, and combined—using data from 1983 to 1997. Based on a cross-validation evaluation of forecast accuracy, the nonlinear model generally outperformed the linear and combined models. Cumulative catch per unit effort was a better predictor than cumulative catch in the first 3 weeks (statistical weeks 28–30) of a fishing season, but the relation was reversed for the remaining 5 weeks. Inseason abundance estimations greatly outperformed the preseason forecasts. Incorporating sex ratios into inseason forecast models correctly adjusted the run timings during a large majority of years and thus improved overall forecasts starting in the second week. In the second through fifth weeks (weeks 29–32), the best performing model with sex ratios improved forecasts more than 30% over the best model without sex ratios; improvements included averages of absolute percentages of relative forecast errors, absolute deviations, and squared residuals. Average absolute percentages of relative forecast errors from the best model were less than 20% before the midpoint of the run and less than 14% at and after the midpoint. Sex ratios improved inseason forecasts more in the recent years (1992–1997) than in the earlier years (1983–1991).

Pink salmon *Oncorhynchus gorbuscha* is the most abundant salmon species in southeastern Alaska and supports an important commercial fishery that takes place during a short period from late June to mid-September, although most catches occur from the late July to late August, or statistical weeks 31–34 (the first week starts with the first Sunday of the year). Like most salmon fisheries in Alaska, pink salmon fisheries in southeastern Alaska are managed by a fixed escapement policy that aims to maximize harvests without jeopardizing the population recruitment. To achieve a targeted escapement, managers must know the magnitude of the incoming spawning run. The accuracy of abundance information acquired inseason substantially affects the manager's ability to achieve management objectives.

Pink salmon abundance in southeastern Alaska fluctuates greatly from year to year and is notoriously difficult to predict. The observed maximum abundance variation was greater than 10-fold dur-

ing the last three decades. Based on the information of the size of parent escapements and subsequent environmental conditions or on time series analysis of catches (Quinn and Marshall 1989), preseason forecasts of pink salmon runs and catches lack sufficient accuracy for management purposes, with relative forecast errors in some years exceeding 200% (Zheng 1988). Consequently, inseason forecasts were initiated to update abundance estimates over the fishing season (Zheng 1988).

Pink salmon migrate from the ocean to the streams of southeastern Alaska through four principal portals: Cross Sound, lower Chatham Strait, Sumner Strait, and Dixon Entrance (Figure 1). The runs spread out from these entrances into branching straits, channels, and fjords. Numerous tagging experiments have demonstrated that northern (NSE) and southern Southeast (SSE) Alaska pink salmon stocks are geographically separated by Sumner Strait (e.g., Rich 1926; Davidson and Christey 1938; Elling and Macy 1955; Nakatani et al. 1975; Hoffman 1982). Thus, pink salmon

---

\* Corresponding author: jiez@fishgame.state.ak.us

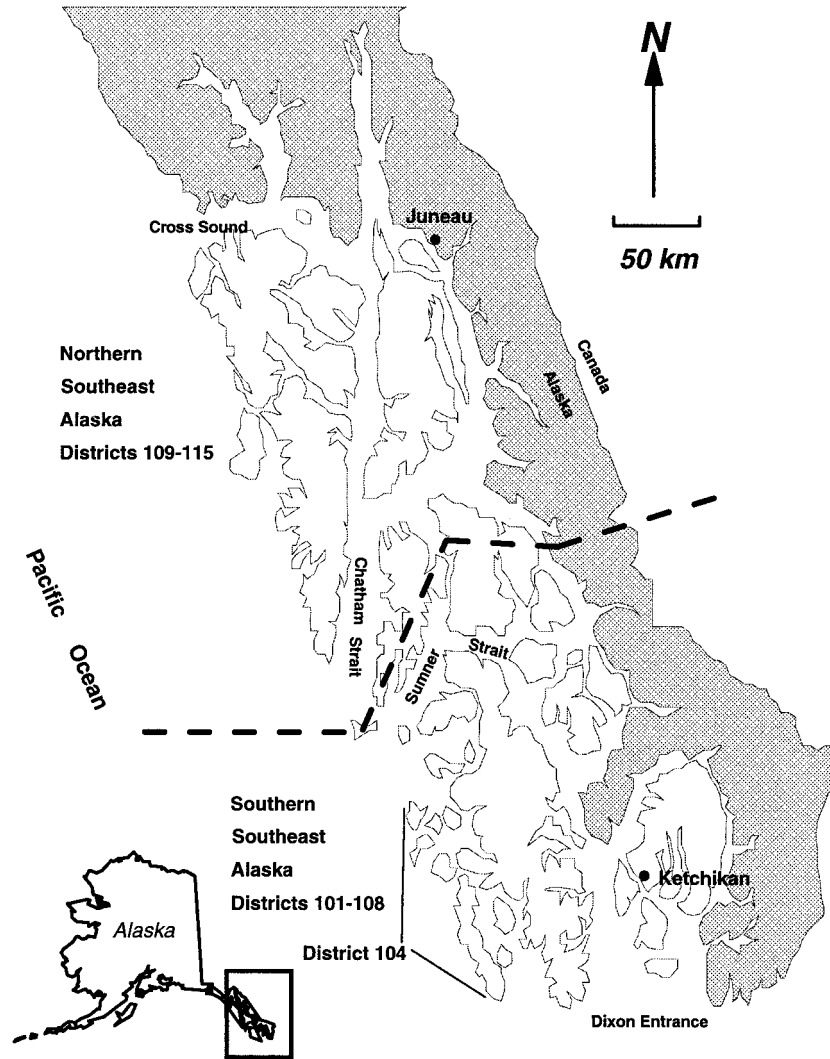


FIGURE 1.—Map of southeastern Alaska showing the four major entrance portals of pink salmon: Icy Strait/Cross Sound, lower Chatham Strait, Sumner Strait, and Dixon Entrance. Note District 104.

abundances are predicted separately for the NSE and SSE Alaska stocks. The pink runs to southeastern Alaska can be characterized by location and time of spawning as early, middle, and late runs (Sheridan 1962). However, these three run segments tend to overlap, so we treated them as a single unit.

Timing of SSE Alaska pink salmon runs varies greatly from year to year (Mathisen and Zheng 1994), and inseason information for run timing each year is needed to accurately forecast run sizes. Typically, male pink salmon predominate the early segment of the run, but this shifts to female predominance toward the end (Mathisen

and Zheng 1994). The changes in sex ratios can be indexed to the temporal progression of the run (McKinstry 1993; Mathisen and Zheng 1994). Using a logistic regression of sex ratios against time to estimate run timing, McKinstry (1993) showed that adjusting run timing improved inseason forecasts, although big improvement generally occurred only following the middle of the run. Because of the small amount of data on sex ratios during a fishing season and nonlinear relationships between sex ratios and time, the logistic regression used by McKinstry (1993) does not provide an accurate estimate of run timing until after the mid-run date. Another drawback of McKinstry's ap-

proach is that the same shape of run-timing curves was assumed for all years, and only the timing of the mid-run changes each year. In reality, the shape of a run-timing curve is affected by the relative strengths of early, middle, and late-run segments; these relative strengths change each year. Because of this assumption and inaccurate run-timing estimates before the mid-run date, inseason forecasts that incorporated sex ratio information did not greatly improve forecast accuracy.

In our study, we incorporated sex ratio information into inseason forecast models to annually adjust timing and shape of the run-timing curves for SSE Alaska pink salmon to improve accuracy of inseason forecasts. First, we defined an index for sex ratio information from deviations of weekly sex ratios and shapes of the sex ratio curves. Then, we used this index to evaluate three inseason forecast models through both cross-validation and 1-year-ahead forecast procedures using available data from 1983 to 1997.

### Methods

Annual escapement data from 1983 to 1997 for SSE Alaska pink salmon and time series of catch, effort, and sex ratio data by statistical week were obtained from the Alaska Department of Fish and Game. Pink salmon caught by the seine fishery were sampled at landing ports to determine what proportion of fish was male; harvest location and date were also recorded. Sample sizes were typically equal to or greater than 100 fish, and several such samples were normally taken for District 104 during an opening that typically lasted from 6 h to 2 d. The male proportion for a statistical week was estimated as total males sampled divided by total sample size. To reduce sampling bias, sex ratio data were discarded for any statistical week if total sample size was fewer than 100 fish. Catch and effort data were reported from fish tickets (i.e., fish sale receipts). Catch was defined as number of pink salmon caught by all gear types; catch per unit effort (CPUE) was computed as total weekly number of pink salmon caught by seine gear divided by total weekly seine effort measured in boat hours. Only CPUE from seine gear was used because more than 90% of pink salmon catch was caught by seiners and because other gear types generally target other species. Total run (or return) was equal to total catch plus 2.5 times peak escapement counts (Geiger and Savikko 1992).

The sex ratio index was derived from deviations of weekly male proportions to the corresponding mean values and the deviation of the sex ratio

curve to its mean curve for a given year. An early run is represented by a lower than expected proportion of males or a sharper than expected decrease in male proportions over time; a late run is represented by the opposite proportions. Accumulating these deviations over time reflects the run timing for a given year. Because most SSE Alaska pink salmon pass through District 104 before entering other districts, we used the sex ratios from District 104 to represent the run timing for all SSE Alaska. A sex ratio index,  $SR_{t,w}$ , in year  $t$  and week  $w$  was defined as

$$\begin{aligned} SR_{t,w} &= P_{t,w} - \bar{P}_w, \quad w = 28; \\ SR_{t,w} &= (P_{t,w} - \bar{P}_w) \\ &\quad + [(P_{t,w} - P_{t,w-1}) - (\bar{P}_w - \bar{P}_{w-1})], \\ &\quad w = 29; \\ SR_{t,w} &= SR_{t,w-1} + (P_{t,w} - \bar{P}_w) \\ &\quad + [(P_{t,w} - P_{t,w-1}) - (\bar{P}_w - \bar{P}_{w-1})], \\ &\quad w \geq 30; \end{aligned} \quad (1)$$

where  $P_{t,w}$  is the proportion of males in year  $t$  and week  $w$ , and  $\bar{P}_w$  is mean proportion of males for week  $w$  from all years, except the forecasted one, for the cross-validation procedure and from all years before the forecasted one for the 1-year-ahead forecast procedure. The seine fishery for SSE Alaska pink salmon usually starts in statistical week 28 each year, so shape information on sex ratio curves was not available in week 28. To avoid overweighting a deviation, starting in week 29, the deviation of a week's male proportion to its corresponding mean value was reduced by multiplying by 0.5 if that week's shape information for the sex ratio curves was unavailable due to missing data.

Three inseason forecast models were evaluated in our study: (1) a linear model,

$$R_t = (a + b CU_{t,w})(1 + cSR_{t,w}), \quad (2)$$

where  $R_t$  is total annual run in year  $t$ ,  $CU_{t,w}$  is either cumulative catch or cumulative CPUE in year  $t$  and week  $w$ , and  $a$ ,  $b$ , and  $c$  are parameters; (2) a nonlinear model,

$$R_t = [a(CU_{t,w} - b)^d](1 + cSR_{t,w}), \quad (3)$$

and (3) a combined model,

$$R_t = (aC_{t,w} + b U_{t,w}^d)(1 + cSR_{t,w}), \quad (4)$$

where  $C_{t,w}$  is cumulative catch in year  $t$  and week

$w$ ,  $U_{t,w}$  is cumulative CPUE in year  $t$  and week  $w$ , and  $d$  is a parameter. The second part of these three models represents an adjustment for run timing by the sex ratio index. The commonly used inseason forecast models, average-timing and linear regression models (e.g., Walters and Buckingham 1975; Barth 1984; Zheng 1988), are special cases of these three models. If sex ratio information is not used, parameter  $c$  is assumed to be 0.

Measurement errors were assumed to be log-normally distributed, and parameters were estimated by a nonlinear least-squares approach using the Microsoft Excel Solver. The cross-validation procedure was used to select the best models (i.e., leaving 1 year's data out of the parameter estimation and then forecasting the run for that year using the parameter estimates; Hjorth 1994; Chen and Shelton 1996). Accuracy of model forecasts of total runs for 1992–1997 was also evaluated with the 1-year-ahead forecast procedure (Quinn and Marshall 1989); thus, data from 1983 to 1991 were used to estimate parameters and forecast the 1992 run, and so on. Different initial parameter values were used to find the best parameter estimates for the full data set for each model. These estimates then served as initial parameter values for cross-validation and 1-year-ahead forecasts. Parameter  $b$  in equation (3) was constrained so that  $CU_{t,w} - b$  is equal to or greater than one for all data. Among 1,050 estimations for cross-validation, only one (the nonlinear model with cumulative catch and sex ratio index as independent variables in year 1988 and week 29) failed to converge because  $b$  and  $d$  tended to be highly confounded for this case;  $b$  was fixed at zero to make the estimation converge. All 420 estimations for 1-year-ahead forecasts converged.

Averages and medians of absolute percentages of relative errors, absolute deviations, and squared residuals (Stocker and Noakes 1988) were used to evaluate model performance. Relative error was defined as the difference between the forecasted and observed runs divided by the observed run. The improvement or deterioration, in terms of evaluation criteria from the best model with sex ratios over the best model without sex ratios, was computed.

As a comparison, relative forecast errors were computed from actual preseason forecasts made by the Alaska Department of Fish and Game for 1983–1992 (Geiger and Savikko 1992) and by the University of Alaska Fairbanks for 1993–1997. The escapement indices used in 1983–1990 were converted to total escapement by multiplying by

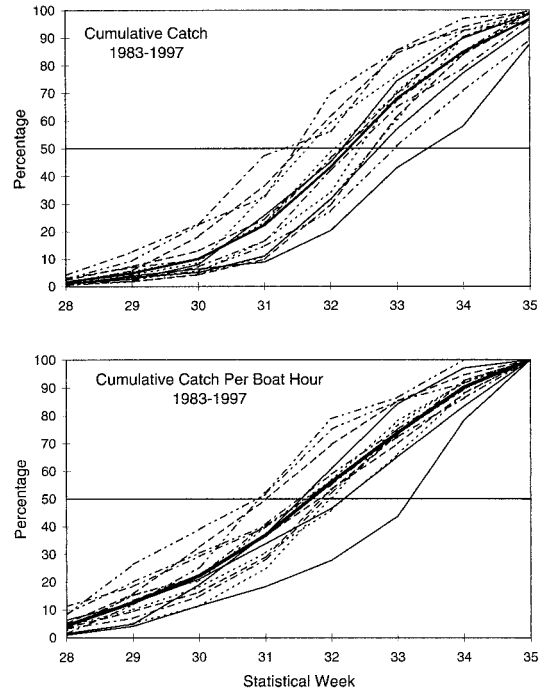


FIGURE 2.—Cumulative percentages of catch and catch per seine boat hour (CPUE) from the pink salmon fishery in southern Southeast Alaska, 1983–1997.

2.5, and the total forecasted run in 1995 was estimated from the total forecasted catch times 1.712, which is the average ratio in 1993, 1994, 1996, and 1997, when both runs and catches were forecasted.

## Results

To illustrate the pink salmon run-timing variation from year to year, we plotted percentages of cumulative catch and cumulative CPUE from 1983 to 1997. At the 50% point, or mid-run point, the difference between the earliest and latest runs exceeded 2 weeks, and the difference between the cumulative percentages at the same week for an early and late run can exceed 50 percentage points (Figure 2).

Sex ratio data in District 104 generally reflected timing patterns of pink salmon runs to SSE Alaska; for example, a lower than expected proportion of males indicates an earlier than expected run. The mid-run point generally fell between the 50% cumulative catch and CPUE and was closer to the 50% cumulative catch than to the cumulative CPUE for most years (Figure 3). Similar to percentages of cumulative catch and CPUE, the percentage of males crossed 50% in week 31 in 1997

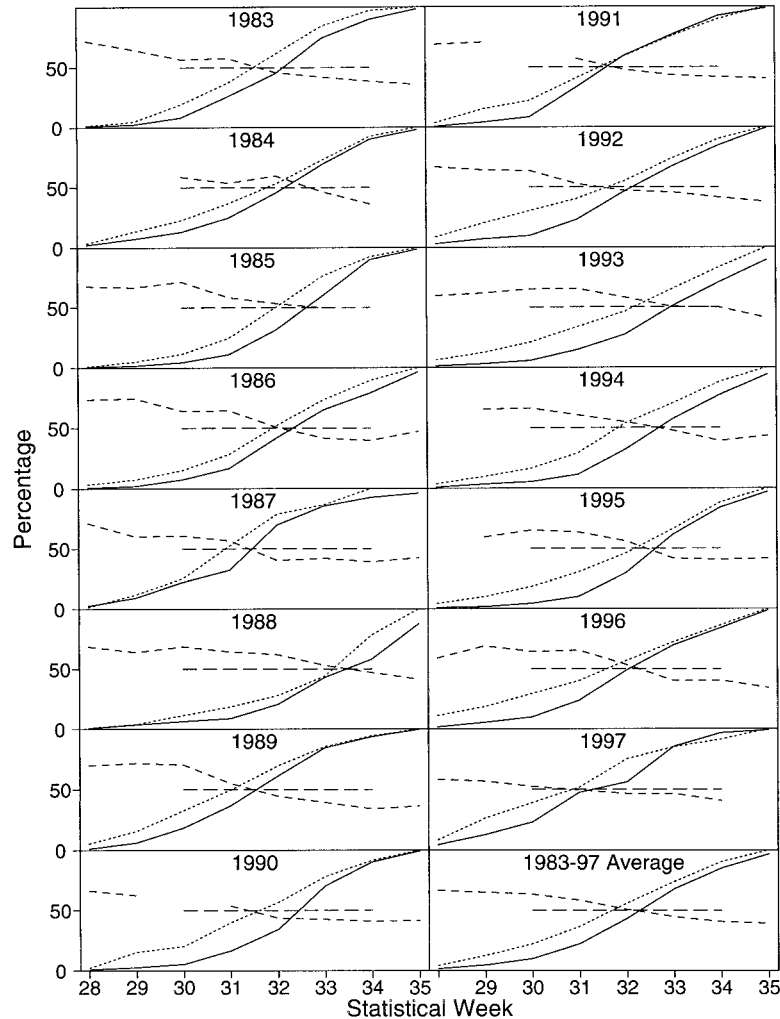


FIGURE 3.—Cumulative percentages of catch (solid line) and catch per seine boat hour (CPUE, dotted line), and percentages of male pink salmon (stippled line). Broken horizontal lines from weeks 30–34 represent 50% lines.

and between weeks 33 and 34 in 1988 (Figure 3), the earliest and latest runs observed during the last 15 years.

The relationships between total run size and cumulative catch or CPUE changed over a fishing season. These relationships were nonlinear early in the season and resembled power functions. Late in the season, the relationships became linear; this was especially the case for cumulative catch because it makes up a large portion of the total run (Figure 4). The lowest curves capture these relationships. In the first week of the seine fishery (week 28), total run as a function of cumulative catch or CPUE varied greatly; as expected, this variation decreased over time. In addition, this

variation decreased faster over time as a function of cumulative catch than as a function of cumulative CPUE (Figure 4).

The decreases of averages and medians of relative forecast errors, absolute deviations, and squared residuals over time and the improvements of these evaluation criteria with the addition of sex ratios depended on the model and whether cumulative catch or CPUE was used as the independent variable. Forecast improvements over time were much less with CPUE than with catch as the independent variable because cumulative CPUE did not predict total run well during the second half of a fishing season (Tables 1–3). Overall, with the cross-validation procedure, the nonlinear mod-

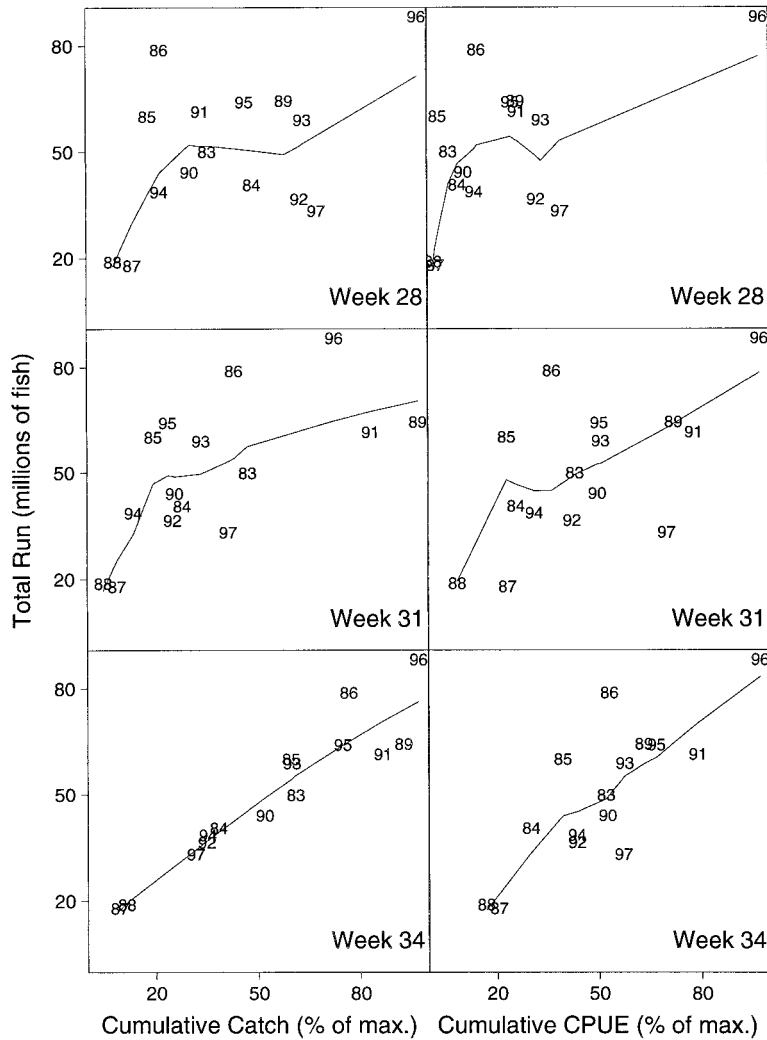


FIGURE 4.—Plots of total run against cumulative catch in number and cumulative catch per seine boat hour (CPUE) for weeks 28, 31, and 34. Numbers within plots are years, and solid lines are lowest smoothed curves (bandwidth  $f = 2/3$ ).

el performed much better than the linear model from statistical weeks 29–32, whether sex ratios were used or not (Tables 1–3). When sex ratios were used, the forecast improvements were also generally greater with the nonlinear model than with the linear model. The poor performance of the linear model partially resulted from extremely poor predictions of weak runs in 1987 and 1988. As expected, the combined model that used both cumulative catch and CPUE produced forecasts between those made by the two linear models that used cumulative catch and CPUE separately (Tables 1–3). Medians were generally smaller than averages for all models and independent variables,

especially for the squared residuals (Tables 1–3), indicating long right tails in the distributions for these criteria measurements. As expected, the forecast improvements with sex ratios were also much smaller with median as a criterion than with average in weeks 28–32 because sex ratios were primarily used to adjust abnormal run timing. Overall, all six criteria indicated forecast improvements with sex ratios from weeks 29–33, except for the median of absolute deviations in week 30. In weeks 29–32, the best performing model with sex ratios reduced averages of relative forecast errors more than 32%, averages of absolute deviations more than 30%, and averages of squared residuals

TABLE 1.—Summary of absolute percentages of relative inseason forecast errors of pink salmon runs to southern Southeast Alaska with a cross-validation procedure, 1983–1997. The results are classified by week and model, with catch or catch per seine boat hour (CPUE) and with sex ratios (WSR) or without sex ratios (XSR). “Difference” is the results from the best performing model without sex ratios minus those with sex ratios. “Improve” is “Difference” divided by the results from the best model without sex ratios.

Week of year	Linear model				Nonlinear model				Combined model		Best model	
	Catch		CPUE		Catch		CPUE		Catch and CPUE		Difference (XSR – WSR)	Improve (%)
	XSR	WSR	XSR	WSR	XSR	WSR	XSR	WSR	XSR	WSR		
<b>Average</b>												
28	43.4	44.9	41.6	38.6	36.3	43.0	33.8	36.5	38.8	35.9	-2.1	-6.2
29	42.2	23.3	45.8	24.0	34.2	24.8	37.4	18.7	42.2	25.3	15.5	45.5
30	41.1	43.3	39.9	36.7	37.5	28.4	38.2	23.7	40.0	39.0	13.8	36.9
31	35.4	30.4	34.6	30.5	34.2	19.5	33.4	20.8	34.6	27.2	14.0	41.8
32	21.9	22.4	28.8	27.1	20.7	14.0	36.3	21.5	23.4	19.5	6.7	32.3
33	13.2	12.1	23.7	23.0	13.0	10.9	29.7	19.2	14.1	13.1	2.2	16.6
34	9.0	9.7	20.5	24.9	9.3	9.7	25.6	20.7	9.3	10.1	-0.6	-7.0
<b>Median</b>												
28	26.6	39.4	26.0	21.0	23.5	40.1	18.6	29.8	30.8	24.9	-2.4	-12.6
29	27.8	15.0	28.8	17.2	32.2	21.6	21.5	17.7	27.6	15.9	6.5	30.1
30	26.5	29.9	17.4	25.8	30.0	22.8	21.5	15.0	30.3	37.1	2.4	13.8
31	35.8	16.9	15.7	24.3	24.3	14.9	17.6	16.1	33.2	23.1	0.9	5.6
32	16.7	14.0	16.8	26.7	20.8	13.3	20.8	20.3	19.0	16.4	3.5	20.8
33	11.0	5.2	15.8	18.5	14.4	10.4	17.3	15.0	13.8	6.6	5.8	52.5
34	9.5	6.6	14.4	17.9	10.5	7.5	17.9	20.3	9.5	6.6	2.9	30.2

more than 47% (Tables 1–3). Averages of relative forecast errors for the best performing model with sex ratios were less than 24% in weeks 29–34, compared with less than 38% for the best model without sex ratios and 51% for preseason forecasts.

With the cross-validation procedure and averages of relative forecast errors as a criterion, the best models that used sex ratios were the combined model in week 28, nonlinear CPUE model in weeks 29–30, nonlinear catch model in weeks 31–

TABLE 2.—Summary of absolute deviations (millions of fish) between forecasted and observed pink salmon runs to southern Southeast Alaska with a cross-validation procedure, 1983–1997. The results are classified by week and model, with catch or catch per seine boat hour (CPUE) and with sex ratios (WSR) or without sex ratios (XSR). “Difference” is the results from the best performing model without sex ratios minus those with sex ratios. “Improve” is “Difference” divided by the results from the best model without sex ratios.

Week of year	Linear model				Nonlinear model				Combined model		Best model	
	Catch		CPUE		Catch		CPUE		Catch and CPUE		Difference (XSR – WSR)	Improve (%)
	XSR	WSR	XSR	WSR	XSR	WSR	XSR	WSR	XSR	WSR		
<b>Average</b>												
28	17.2	18.9	16.1	14.7	15.7	17.9	14.7	15.3	16.0	15.5	0.01	0.1
29	16.7	9.6	18.1	10.0	14.8	10.8	16.1	9.3	16.8	10.5	5.6	37.5
30	17.9	19.7	15.3	14.6	15.2	13.0	15.6	10.6	17.4	18.0	4.6	30.3
31	16.1	13.4	13.7	13.3	14.4	8.0	13.9	9.0	15.5	12.6	5.7	41.5
32	11.1	10.6	12.1	12.1	10.1	6.9	14.0	9.5	11.6	9.6	3.2	32.0
33	7.4	5.5	10.4	10.5	7.3	5.2	11.9	9.4	7.7	5.7	2.1	29.0
34	5.3	4.7	9.2	11.2	5.5	4.8	10.9	10.5	5.5	4.9	0.6	11.4
<b>Median</b>												
28	17.8	15.5	16.0	10.5	13.2	12.3	10.8	11.5	13.6	9.3	1.5	13.9
29	13.8	8.1	15.5	9.5	15.2	11.4	12.2	9.9	17.3	10.0	4.1	33.3
30	16.3	11.3	10.9	12.2	13.0	9.8	8.0	10.5	12.6	13.6	-1.8	-22.7
31	15.7	7.5	9.4	9.7	13.6	8.8	7.8	7.7	16.7	9.0	0.3	4.1
32	7.3	9.0	8.4	10.3	10.8	5.7	10.4	8.6	9.9	7.5	1.7	23.0
33	5.3	2.2	8.5	9.9	7.3	3.9	9.2	9.2	5.7	2.9	3.0	57.7
34	4.3	2.4	6.3	10.6	4.8	2.8	7.8	10.1	4.2	2.4	1.9	44.6

TABLE 3.—Summary of squared residuals between logarithms of forecasted and observed pink salmon runs to southern Southeast Alaska with a cross-validation procedure, 1983–1997. The results are classified by week and model, with catch or catch per seine boat hour (CPUE) and with sex ratios (WSR) or without sex ratio (XSR). “Difference” is the results from the best performing model without sex ratios minus those with sex ratios. “Improve” is “Difference” divided by the results from the best model without sex ratios.

Week of year	Linear model				Nonlinear model				Combined model		Best model	
	Catch		CPUE		Catch		CPUE		Catch and CPUE		Difference (XSR – WSR)	Improve (%)
	XSR	WSR	XSR	WSR	XSR	WSR	XSR	WSR	XSR	WSR		
<b>Average</b>												
28	0.23	0.23	0.21	0.20	0.17	0.25	0.15	0.20	0.19	0.19	-0.03	-21.5
29	0.21	0.08	0.24	0.09	0.16	0.07	0.17	0.05	0.22	0.09	0.11	69.1
30	0.19	0.21	0.22	0.19	0.23	0.11	0.20	0.10	0.19	0.20	0.09	47.6
31	0.15	0.12	0.17	0.13	0.33	0.11	0.17	0.07	0.15	0.10	0.08	55.0
32	0.07	0.07	0.13	0.10	0.05	0.03	0.26	0.07	0.07	0.05	0.03	51.6
33	0.02	0.03	0.09	0.07	0.02	0.02	0.23	0.05	0.03	0.03	0.002	7.2
34	0.01	0.02	0.07	0.08	0.01	0.02	0.11	0.06	0.01	0.02	-0.003	-24.0
<b>Median</b>												
28	0.08	0.11	0.09	0.06	0.07	0.11	0.04	0.07	0.07	0.08	-0.01	-30.5
29	0.07	0.03	0.12	0.03	0.10	0.04	0.06	0.03	0.10	0.02	0.04	62.9
30	0.10	0.09	0.04	0.05	0.08	0.05	0.06	0.02	0.07	0.10	0.01	37.4
31	0.10	0.03	0.03	0.06	0.05	0.03	0.03	0.03	0.08	0.04	0.001	4.0
32	0.03	0.02	0.02	0.06	0.06	0.02	0.04	0.03	0.04	0.03	0.004	16.0
33	0.01	0.00	0.02	0.04	0.02	0.01	0.03	0.02	0.02	0.01	0.01	78.8
34	0.01	0.00	0.02	0.04	0.01	0.01	0.03	0.04	0.01	0.00	0.004	51.1

33, and linear catch model in week 34 (Table 1). The best models without sex ratios were the nonlinear CPUE model in week 28, the nonlinear catch model in weeks 29–33, and the linear catch model in week 34. Forecasted runs with and without use of sex ratios plotted against observed runs further confirmed the improvements from using sex ratios; the correlation coefficients between forecasted and observed runs were always higher with the use of sex ratios than without them (Figure 5). Without sex ratios, the forecast model did not fit the data well in week 29. The decrease in  $r^2$  from week 29 to week 30 with the sex ratio model was primarily due to deteriorated forecasts in 1983, 1988, and 1989. The sex ratios in District 104 greatly overadjusted the run timing in 1983 during weeks 29–31. Overall, the best models with sex ratios outperformed those without sex ratios in 10 out of 15 years in weeks 29–30 and 11 or 12 out of 15 years in weeks 31–34 (Figure 5).

Box plots are an alternative way to compare relative forecast errors. The median of the absolute percentages of relative forecast errors with pre-season forecasts was much larger than those with inseason forecasts with the cross-validation procedure, whether sex ratios were used or not (Figure 6). In week 28, sex ratios did not improve forecasts, but during weeks 29–33, incorporating sex ratios into the forecast models resulted in much

smaller extreme and median forecast errors than without sex ratios. In week 34, the median error was smaller with sex ratios, but the extreme error was larger than it was without sex ratios. When using sex ratios, distributions of forecast errors were skewed to the left in 5 out of 7 weeks, and none were skewed to the right; without using sex ratios, skewness was to the right in 4 out of 7 weeks and to the left in 2 out of 7 weeks (Figure 6). Apparently, use of sex ratios correctly adjusted many abnormal run timings and thus reduced the right-side skewness of the forecast errors.

Starting in statistical week 29, sex ratios generally improved inseason forecasts greater with the 1-year-ahead forecast procedure for 1992–1997 than with the cross-validation procedure for 1983–1997 (Tables 1–4). The different forecast improvement was largely due to different time periods rather than different forecast procedures. The improved forecasts with sex ratio data generally occurred every year during 1992–1997; the greatest improvement was for 1997. Overall, with the 1-year-ahead forecast procedure, the nonlinear model still outperformed the linear model when using sex ratios, but the linear model generally outperformed the nonlinear model without sex ratios (Table 4). Generally, forecast results were similar with both linear and nonlinear models for run sizes equal to or greater than average abundance,



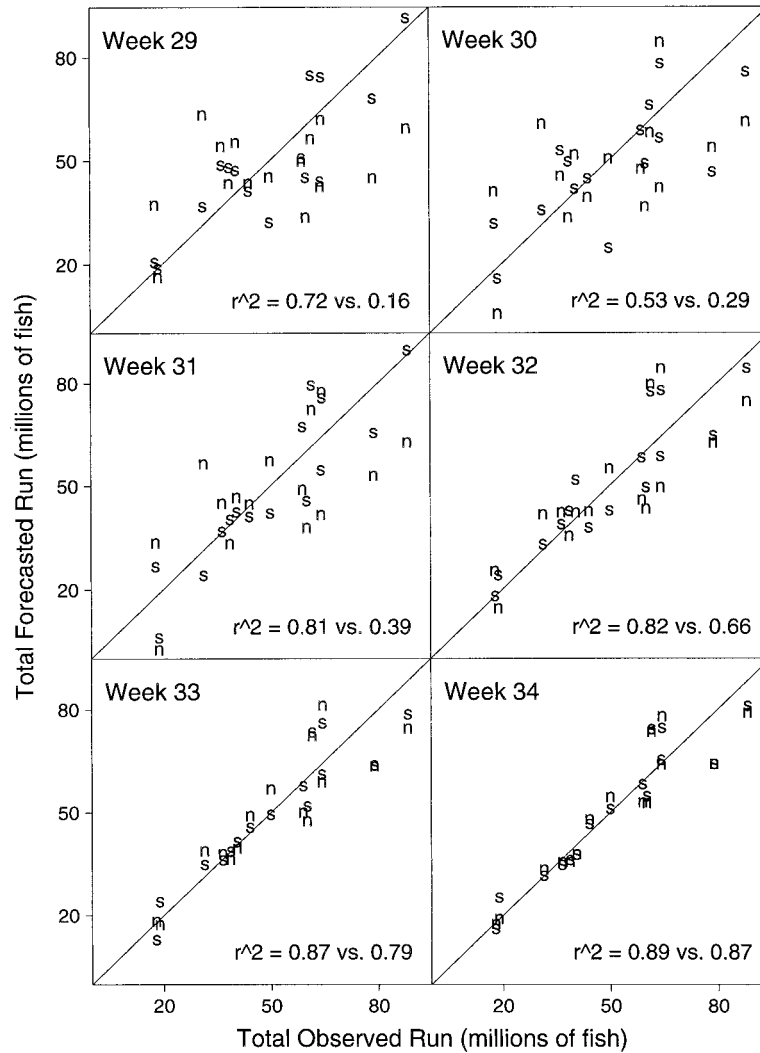


FIGURE 5.—Plots of forecasted runs from the best performing models with and without sex ratios against observed runs from weeks 29–34 with a cross-validation procedure. The first  $r^2$  is the coefficient of determination for the results with sex ratios and the second  $r^2$  is for the results without sex ratios. Letters “n” and “s” represent the results without sex ratios and with sex ratios, respectively.

whereas the nonlinear model greatly outperformed the linear model for small run sizes. The run sizes during 1992–1997 were about average or greater.

Total catches can be forecasted in the same way as total runs or can be estimated from total forecasted runs through a linear model. For the latter approach, we illustrate the relationship between total catch and run size during 1983–1997 in Figure 7. Interestingly, the fixed escapement strategy has been implemented as a threshold strategy (i.e., no fishing is allowed when abundance is below a threshold and a constant harvest rate is used when

abundance is above the threshold; Quinn et al. 1990). The current practice of allowing a maximum 50% harvest time (i.e., 2-d opening followed by 2-d closure) when the run is strong is essentially a threshold approach.

### Discussion

Researchers have long recognized that migratory run timings of many fish populations vary greatly from year to year (Mundy 1982; Mathisen and Zheng 1994; Chen and Shelton 1996). In-season forecast models based on average migratory

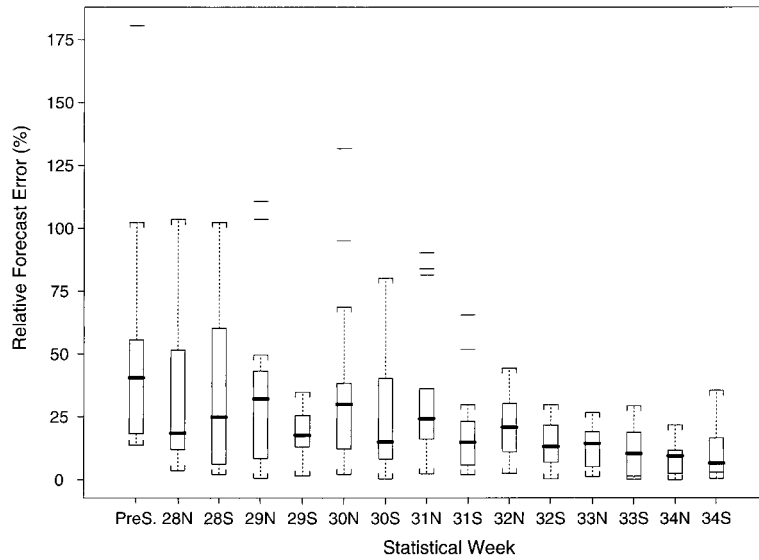


FIGURE 6.—Box plots of absolute percentages of relative forecast errors by pre-season forecasts (PreS.) and the best performing in-season models with (S) and without (N) sex ratios from week 28 to week 34 using a cross-validation procedure. The dark lines within boxes are medians; boxes represent distributions of the central 50% of the data points, the full distribution of forecast errors is bounded by brackets, and individual extreme forecast errors are depicted by single horizontal lines.

timing have underforecasted or overforecasted population abundances because of run timing variation among years (Walters and Buckingham 1975; Mundy 1982; Barth 1984). The run timing may be correlated with some environmental variables, (e.g., temperature; Mundy 1982; Chen and Shelton 1996); these variables can be used to adjust and improve in-season forecasts (Chen and Shelton

TABLE 4.—Summary of the best performing models (Mod) and their statistic values (Val) for pink salmon runs to southern Southeast Alaska with a 1-year-ahead forecast procedure, 1992–1997. “Improve” is difference of the results between the best models without sex ratio (SR) and with SR divided by the result from the best model without SR. Notations for models: L = linear model, N = nonlinear model, C = cumulative catch, U = cumulative catch per seine boat hour, and CU = combined model for cumulative catch and catch per seine boat hour.

Week of year	Absolute % of relative errors					Absolute deviations (million)					Squared residuals (logarithm)				
	No SR		With SR		Improve (%)	No SR		With SR		Improve (%)	No SR		With SR		Improve (%)
	Mod	Val	Mod	Val		Mod	Val	Mod	Val		Mod	Val	Mod	Val	
<b>Average</b>															
28	NC	39.3	NC	45.8	-16.6	LC	16.4	CU	20.6	-25.7	NC	0.16	NU	0.22	-36.4
29	LC	36.3	LC	18.4	49.2	LC	15.3	NU	10.1	34.0	NC	0.16	NU	0.06	65.5
30	LC	32.0	NC	15.5	51.7	CU	14.6	NC	8.5	42.0	NC	0.13	NC	0.03	74.2
31	LC	29.6	CU	10.4	64.8	LU	11.5	NC	6.0	47.9	NC	0.11	NC	0.02	84.8
32	LC	16.0	NC	7.0	56.2	LC	8.0	NC	3.4	57.5	NC	0.04	NC	0.01	86.8
33	LC	10.7	LC	5.1	52.3	LC	5.4	CU	2.3	57.9	LC	0.02	LC	0.00	79.6
34	LC	6.6	NC	4.1	38.0	LC	3.7	LC	2.4	35.1	LC	0.01	LC	0.00	62.3
<b>Median</b>															
28	LC	12.3	NU	36.2	-193.5	LC	8.4	LU	19.0	-125.9	LC	0.03	LU	0.11	-349.7
29	LU	18.6	LC	17.8	4.2	LU	11.6	NU	8.2	28.9	LU	0.05	LC	0.03	41.9
30	LU	16.3	NU	14.7	9.4	CU	12.1	NC	7.9	34.5	LU	0.03	NC	0.02	18.2
31	LU	14.0	CU	9.0	35.3	LU	8.7	CU	5.1	41.5	LU	0.02	CU	0.01	51.4
32	LU	15.2	NC	8.2	46.2	LC	6.6	NC	3.7	44.1	LU	0.03	NC	0.01	74.9
33	CU	8.9	CU	3.8	57.9	LC	5.0	CU	1.9	61.1	CU	0.01	CU	0.00	84.4
34	LC	7.4	NC	3.4	53.8	NC	2.6	NC	1.4	43.5	LC	0.01	NC	0.00	77.8

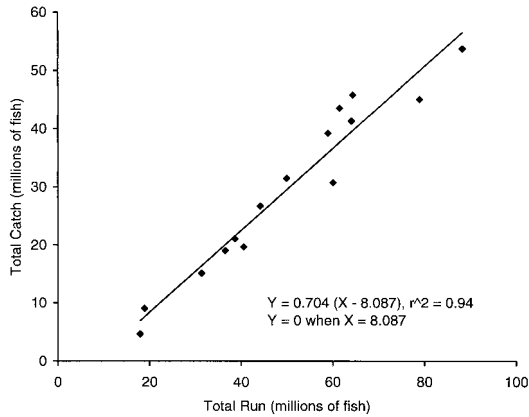


FIGURE 7.—The relationship between total catch and total run for southeastern Alaska pink salmon from 1983 to 1997.

1996). A consistent pattern in the sex ratios of SSE Alaska pink salmon runs provides mileposts for migratory run timing (McKinstry 1993; Mathisen and Zheng 1994) that can be incorporated into in-season forecast models to greatly reduce forecast errors during years with abnormal run timing and thus improve overall forecasts. The best performing model, based on the cross-validation procedure, yielded mean forecast errors less than 20% for 1983–1997 run sizes before the midpoint of the run and less than 14% at and after the midpoint.

Although forecast improvements through incorporating sex ratios occurred during a large majority of years, sex ratios can wrongly adjust run timings during some years. For example, the run timing was about average in 1983, the first year sex ratios were collected throughout a season, but sex ratios collected in District 104 indicated an earlier than expected run. As a result, the sex ratios in 1983 greatly overadjusted the run timing in weeks 29–31. However, the sex ratios in District 101, the second strongest run district, suggested a slightly later than expected run. Combined sex ratios in Districts 104 and 101 may have resulted in prediction of close to average run timing in 1983. Because of change in relative run strengths among districts from year to year and missing data points, it is difficult to combine sex ratios from all districts as a single index. Fortunately, during most years, sex ratios from major districts reveal similar run timings, and sex ratios from District 104 appear adequate to represent run timings of the total SSE Alaska pink salmon run. In a year when sex ratios from District 104 suggest a large abnormal run timing, it is prudent to examine the geographic

distribution of samples for a potential sampling bias and to check sex ratios from other districts for confirmation.

Sex ratios improved inseason forecasts more in the recent years (1992–1997) than in the earlier years (1983–1991). Whether this is due to better sampling of sex ratios, temporal changes in biological relationships, or other factors is not clear. In the recent years, fishing time for District 104 has been reduced during the early part of each season because of concerns about bycatch of sockeye salmon *Onchorynchus nerka*, but the current practice of allowing fishing no more than 50% of the available time has resulted in more steady fishing effort through the season. The catch resulting from steady fishing effort is definitely a better predictor of run strengths than that resulting from fishing efforts that differ sharply from week to week. Forecast improvements incorporating sex ratios were enhanced when the model with catch or CPUE data alone was appropriate to fit the actual run; sex ratios can only adjust abnormal run timings and cannot overcome forecasting problems such as catch or CPUE fluctuations caused by irregular fishing patterns.

Our forecast errors were much smaller than those reported by McKinstry (1993). The difference is due to different forecast models, methods used to incorporate sex ratio data into forecast models, stock definitions, and periods in which the forecasts were conducted. First, the logistic regression used by McKinstry (1993) to incorporate sex ratio data into forecast models did not perform well due to the nonlinear relationship of percentages of male pink salmon over time and the small amount of data. The sex ratio index we developed not only reflects the difference in percentages of males but also derives information on the difference in the shapes of sex ratio curves between a given year and the average. The shape of a sex ratio curve contains information on the relative strengths of different run segments, and the sex ratio index provides crucial information on the run timing. Second, as demonstrated in this study, the nonlinear model outperformed the linear model. The average-timing model used by McKinstry (1993) is a special case of the linear model. Finally, we forecasted pink salmon abundances for all of SSE Alaska from 1983 to 1997, whereas McKinstry (1993) predicted abundances only in Districts 101 and 104 from 1985 to 1992. The abundance in a single district is much more variable than all of SSE Alaska because pink salmon migrate in and out of a district during a fishing

season. Although the forecast improvements by sex ratios were greater in the recent years (1992–1997) than in the earlier years, McKinstry (1993) also did not include the data in 1983 when sex ratios had the worst impact on forecast performance among all years.

Inseason and preseason forecasts can be combined into a single forecast through variance weighting (Walters and Buckingham 1975; Chen and Shelton 1996) or a Bayesian approach (Fried and Hilborn 1988). A combined forecast may improve the inseason forecast in the early season for the SSE Alaska pink salmon run. However, because of large forecast errors, preseason forecasts may not contribute much to the combined forecast.

We assumed inseason catch was accurate. In reality, there are 2–3-week lags to process fish tickets; therefore, the accurate catch was not available in time for the inseason forecasts. If an inseason forecast is made by the end of a week, the catch for the week has to be estimated. The Alaska Department Fish and Game normally estimates catch and effort for each opening by using information from tender reports and observed fishing effort in each fishing district. The estimates during the early season in the past were usually about 0–20% below the final value from all fish tickets, which, if not adjusted, would result in underforecasts of run size up to a similar or less percentage, depending on the total run size and model used. Errors in estimating inseason catch during a season could increase the forecast errors reported in our study. However, inseason CPUE estimates from tender reports are fairly accurate. Furthermore, early in the season, a CPUE model generally performs better than a catch model based on the cross-validation procedure. Therefore, CPUE should be used for inseason forecasts early in the season during weeks 28–30.

Other potential sources of error in catch and escapement estimates are catch conversion from weight to number and peak escapement count. The current practice is to weigh all fish, estimate daily mean weight from subsamples, and convert total catch in weight to number. Subsample size was usually 100 fish, and each processor took its subsamples from each tender delivery. This error affects both the postseason “true” catch used to calibrate the models as well as the weekly catch used for the inseason forecasts. This sampling error is probably random and small because daily multiple subsamples were usually taken by different processors. Estimates of sex ratios are not affected by this weight–number conversion because sex ratios

were sampled separately through counting both males and females. The relationship between total escapement and peak escapement count may be a function of run size, with a stronger run having a higher conversion factor. The current constant conversion factor of 2.5 is inadequate. Establishing a conversion relationship between total escapement and peak escapement count would improve escapement estimates.

Unbiased sex ratio data are a key to adjusting migratory timing in an inseason forecast model to improve forecast accuracy of pink salmon abundance in SSE Alaska. Ideally, pink salmon are randomly sampled in multiple locations within District 104 in each opening, and weekly sample sizes must be of sufficient size. A single sample or a small sample size can be biased because male and female pink salmon are not uniformly distributed and sex ratios at a given time may be different between pink salmon spawning in different streams. To reduce biased sex ratios, we discarded data with weekly sample sizes of fewer than 100 fish, but we have not dealt with the problems of small numbers of samples and limited geographic distributions of samples. Generally, three or more samples in different subdistricts and a total sample size greater than 500 were taken to determine weekly sex ratios in District 104. A study is needed to evaluate the current sampling practice and to design a sampling approach that minimizes sampling biases.

Sex ratios may also be used to estimate run timing and forecast run sizes for other salmon species. Change in sex ratios over the season is common for salmon, but due to multiple age structure, it is more difficult to interpret sex ratio data for salmon species other than pink salmon. For those other species, both age structure and change in sex ratio may need to be considered simultaneously. Furthermore, although different catchabilities of males and females and sex-selective fishing are not a likely problem for the southeastern Alaska pink salmon purse seine fishery, they may exist for other salmon fisheries, especially when one sex is more valuable than the other sex (e.g., a roe fishery). In such cases, test fisheries may be necessary to establish unbiased estimates of sex ratios.

Change in sex ratios over a fishing season and forecast improvements by sex ratios have important implications for pink salmon fisheries management. First, improved forecasts will reduce management or implementation error and thus increase the ability to obtain optimal catch and optimal escapement levels. Without reliable inseason

forecasts, managers have to rely heavily on inseason escapement estimates to avoid underescapement, and the fishery is closed until "comfortable" numbers of fish have been observed in streams. This conservative approach will help to achieve escapement goals when the runs are weak but will also effect overescapement and loss of harvest opportunity due to the constricted fishing time when the runs are strong. This approach will also result in primarily harvesting middle and late segments of runs. On the other hand, reliable inseason forecasts will provide a lead time for managers to adjust harvest levels to improve the precision of harvest management. Because southeastern Alaska pink salmon comprise numerous individual stocks and inseason forecasts are more precise for the aggregate than for individual stocks, a buffer may be needed in the use of the inseason forecasts for the aggregate to increase the probability that no, or few, stocks will be underescaped. Such a buffer could be basing harvest on an inseason forecast that has an 80%, rather than a 50%, probability of being larger than the true run size. These buffers could be adjusted (or reduced) over the season to minimize loss of harvest opportunity as relative run strengths of individual stocks become increasingly verified from the inseason escapement surveys in streams. These relative run strengths can also supplement aggregate inseason forecasts used to manage subdistricts.

Second, it may be possible to implement a sex-selective harvest strategy (Reed 1982). If an individual pink salmon male can fertilize more than one female, as indicated by Mathisen (1962) for sockeye salmon when there is a preponderance of females on the spawning grounds, then escapement goal for males can be lower than for females, thus allowing a higher harvest rate for males than for females or a higher rate early in the season than late in the season. However, economical factors, such as different values of males versus females, and biological factors other than adequate fertilization (e.g., the role of competition among males in the population genetics or the importance of nutrient recycling from the excess males) need to be fully examined before a sex-selective harvest strategy should be implemented. Furthermore, a sex-selective harvest strategy may practically be difficult to implement on a small scale because of mixtures of fish from different streams in the early season and the difficulty of forecasting run strengths to individual streams.

Improved inseason forecasts also have important implications for pink salmon fishing and pro-

cessing operations. First, improved forecasts will improve fishing industry planning and preparation; this will reduce the cost of operation. Second, accurate forecasts will allow harvests to occur relatively early when fish quality is good rather than in the terminal area where fish quality deteriorates rapidly.

#### Acknowledgments

The data for this study were collected by the personnel in the Southeast Regional Office of the Alaska Department of Fish and Game. We thank Jim Blick, Phillip Mundy, and two anonymous reviewers for their comments on the manuscript. Suggestions from Jim Blick greatly improved this manuscript and are appreciated. This study was funded by the salmon processors of southeastern Alaska. This is contribution PP-168 of the Alaska Department of Fish and Game, Commercial Fisheries Division, Juneau.

#### References

- Barth, E. J. 1984. An intraseason yield forecasting system for commercial marine fisheries. Doctoral dissertation. Old Dominion University, Norfolk, Virginia.
- Chen, X. H., and P. A. Shelton. 1996. A nonparametric forecast model of inshore Atlantic cod (*Gadus morhua*) landings based on biomass, cumulative landings, and water temperature. *Canadian Journal of Fisheries and Aquatic Sciences* 53:558–562.
- Davidson, F. A., and L. S. Christey. 1938. The migration of pink salmon (*Oncorhynchus gorbuscha*) in the Clarence and Sumner straits region of Southeastern Alaska. *U.S. Bureau of Fisheries Bulletin* 48:643–666.
- Elling, G. H., and P. T. Macy. 1955. Pink salmon tagging experiments in Icy Strait and upper Chatham Strait, 1950. *Fishery Bulletin* 56:331–371.
- Fried, S. M., and R. Hilborn. 1988. Inseason forecasting of Bristol Bay, Alaska, sockeye salmon (*Oncorhynchus nerka*) abundance using Bayesian probability. *Canadian Journal of Fisheries and Aquatic Sciences* 45:850–855.
- Geiger, H. J., and H. Savikko. 1992. Preliminary forecasts and projections for 1992 Alaska salmon fisheries and review of the 1991 season. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J92-05, Juneau.
- Hjorth, U. 1994. Computer intensive statistical methods: validation model selection and bootstrap. Chapman and Hall, New York.
- Hoffman, S. H. 1982. Northern Southeastern Alaska pink salmon (*Oncorhynchus gorbuscha*) tagging investigations, 1977–1980. Alaska Department of Fish and Game, Division of Commercial Fisheries, Informational Leaflet 196, Juneau.
- Mathisen, O. A. 1962. The effect of altered sex ratios

- on spawning of red salmon. Pages 137–246 in T. K. Yoo, editor. *Studies of Alaska salmon*. University of Washington Press, Seattle.
- Mathisen, O. A., and J. Zheng. 1994. Changing sex ratios of pink salmon during the spawning migration. Pages 137–146 in *Proceedings of 16th north-east Pacific pink and chum salmon workshop*. University of Alaska Fairbanks, Sea Grant Report 94-02, Fairbanks.
- McKinstry, C. A. 1993. Forecasting migratory timing and abundance of pink salmon (*Oncorhynchus gorbuscha*) runs using sex-ratio information. Master's thesis. University of Washington, Seattle.
- Mundy, P. R. 1982. Computation of migratory timing statistics for adult Chinook salmon in the Yukon river, Alaska, and their relevance to fisheries management. *North American Journal of Fisheries Management* 2:359–370.
- Nakatani, R. E., G. J. Paulik, and R. van Cleve. 1975. Pink salmon (*Oncorhynchus gorbuscha*) tagging experiments in S.E. Alaska 1938–1942 and 1945. NOAA Technical Report NMFS-SSRF-686.
- Quinn, T. J., II, R. Fagen, and J. Zheng. 1990. Threshold management policies for exploited populations. *Canadian Journal of Fisheries and Aquatic Sciences* 47:2016–2029.
- Quinn, T. J., II, and R. P. Marshall. 1989. Time series analysis: quantifying variability and correlation in SE Alaska salmon catches and environmental data. Pages 67–80 in R. J. Beamish and G. A. McFarlane, editors. *Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models*. Canadian Special Publication of Fisheries Aquatic Sciences 108.
- Reed, W. J. 1982. Sex-selective harvesting of Pacific salmon: a theoretically optimal solution. *Ecological Modelling* 14:261–271.
- Rich, W. H. 1926. Salmon tagging experiments in Alaska, 1924 and 1925. U.S. Bureau of Fisheries Bulletin 42:109–146.
- Sheridan, W. L. 1962. Relation of stream temperatures to timing of pink salmon escapements in southeast Alaska. Pages 87–102 in N. J. Wilimovsky, editor. *Symposium on pink salmon*. H. R. Macmillan Lectures in Fisheries, University of British Columbia, Vancouver.
- Stocker, M., and D. J. Noakes. 1988. Evaluating forecasting procedures for predicting Pacific herring (*Clupea harengus pallasii*) recruitment in British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* 45:928–935.
- Walters, C. J., and I. S. Buckingham. 1975. A control system for intra-season salmon management. Pages 12–29 in *Proceedings of a workshop on salmon management*. International Institute for Applied Systems Analysis, Laxenburg, Austria.
- Zheng, J. 1988. Hydroacoustic information system for dynamic management of pink salmon fisheries in Southeastern Alaska. Master's thesis. University of Alaska Fairbanks, Fairbanks.

Received February 12, 1998

Accepted July 1, 1998