Re-Analysis of Statistical Models used in California Department of Fish and Game's (CDFG) San Joaquin River Fall-run Chinook Salmon Population Model

September 15, 2008

1 Introduction

The purpose of this technical report is to re-analyze the original statistical models used in the CDFG San Joaquin River Fall-run Chinook Salmon Population Model (referred to subsequently as SJRModel). The analyses address a critique of the original models by fitting so-called proper models, which insure the outcome will be predicted to be in the appropriate range (i.e., a model that guarantees that survival probability is predicted to be between 0 and 1). This report also presents exploratory analyses, using smooth regression, used to empirically examine relationships of, for instance, flow and survival of smolts. Detailed descriptions of the data as well as how these statistical models are incorporated in the overall model can be read in the original description of the model, "FINAL DRAFT (11-28-05) San Joaquin River Fall-run Chinook Salmon Population Model".

2 Cohort Abundance

Cohort abundance is determined from a regression relationship between the annual calculated number of smolts arriving at Chipps and the estimated production year cohort (data for years 1988 through 2000 with 1989 being removed for reasons described in our response to the SJRGA report). We start using a generalized additive model (GAM) smooth of the estimated proportion of Chipp Smolts of a cohort that return to spawn versus the estimated number of smolts at Chipps ([Hastie & Tibshirani(1990)]). The automatic bandwidth selection available within the *gam* function in *R* ([R Development Core Team(2005)]) is used, which governs the smoothness of the curve. As one can see from figure 1, the logit model is not at all (logit)linear, but looks very much log-linear in the logit. Thus, we re-fit the smooth using the log(smolts) as the predictor.

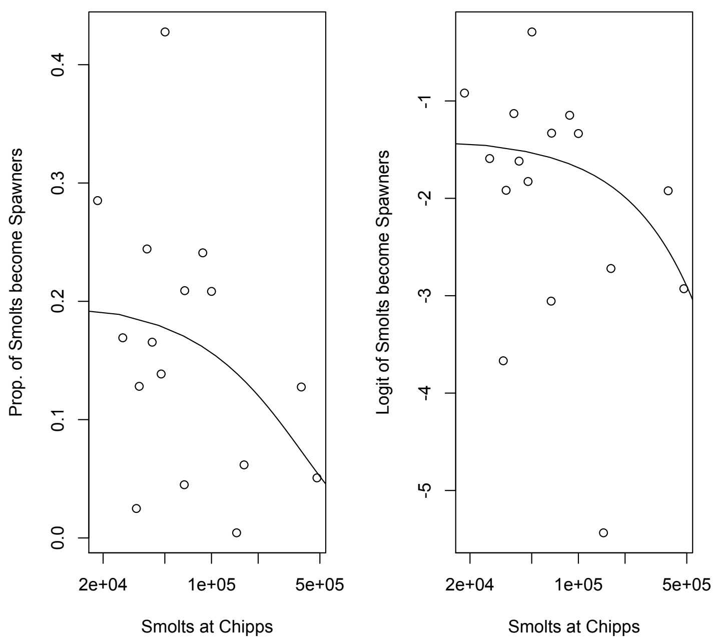

Figure 1: Original Data and GAM smooth on proportion (left) and logit scale (right) of Escapement vs. Chipps Smolts

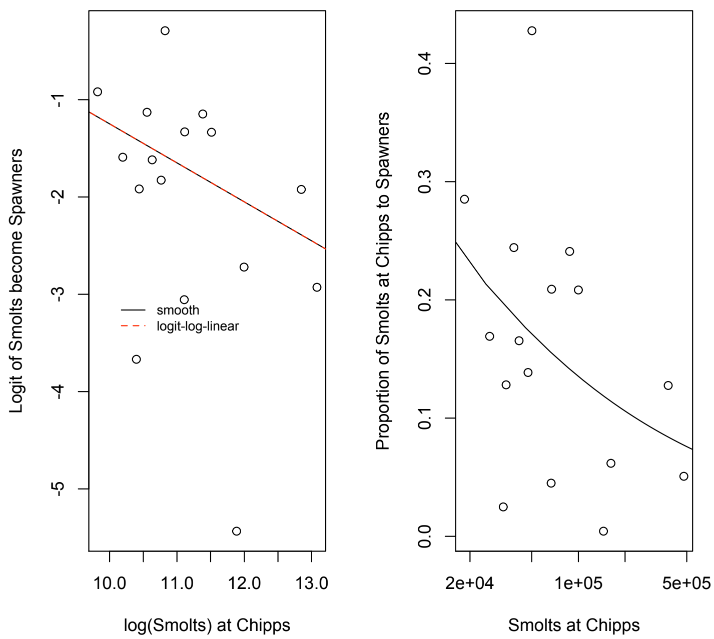

Figure 2: A) GAM smooth and GLM logit-linear model (with original data) on logit scale of Escapement vs. log(Chipps Smolts). B) Same GAM model on probability scale.

As seen on the plot (Figure 2), both the smooth and the logit-linear model fit exactly the same trend, which results from the automatic bandwidth selection procedure that results in a line. This suggest, given the amount of data, the logit-linear is a relatively good fit. We now summarize this fit.

```
> summary(glm.1)

Call:
glm(formula = prop ~ logsmolts, family = binomial(), na.action = na.omit)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.52417  -0.17584  -0.07036   0.14396   0.60002

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   2.7578     9.5381   0.289    0.772
logsmolts    -0.4007     0.8682  -0.461    0.644

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1.4860  on 15  degrees of freedom
Residual deviance: 1.2528  on 14  degrees of freedom
AIC: 9.5521
```

This implies that for relatively low numbers of smolts at Chipps, high percentages on average return ultimately to spawn, whereas for high numbers, that percentage drops. At this point, this is just a black-box model, and there is no obvious biological interpretation for the negative association of the estimated number of Chipp smolts and the ultimate proportion (spawners/smolts) of returning spawners. In our Version 2.0 model, we will add much detail to the ocean component to try to estimate how changes in both ocean conditions and intensity of sports and commercial fishing will impact future populations relative to inland environmental factors.

3 Mossdale Smolt Production

The model presented in this section are used to predict the total number of smolts that will arrive at Mossdale as a function of number of SJR salmon escaping into east-side SJR tributaries in the previous fall-run escapement coupled with current year spring Vernalis Spring out-flow. The year 1989 was removed, believing this point to be unrepresentative of the data-generating distribution for which we are attempting to estimate relationships of annual smolts at Mossdale versus environmental factors and numbers of spawners the previous fall (see responses to SJRGA review for justification). We employ the same exploratory procedure to look at smolt production as we did for cohort production above. First, we examine smolt production (a count) as a function of spring Vernalis flow and previous year escapement, separately, using GAM smooths with log-link (Poisson). Note, by including the log(Escapement) as a predictor, the class of models will include the possibility of density dependence (rate of smolts/spawner negatively associated with number of spawners). We

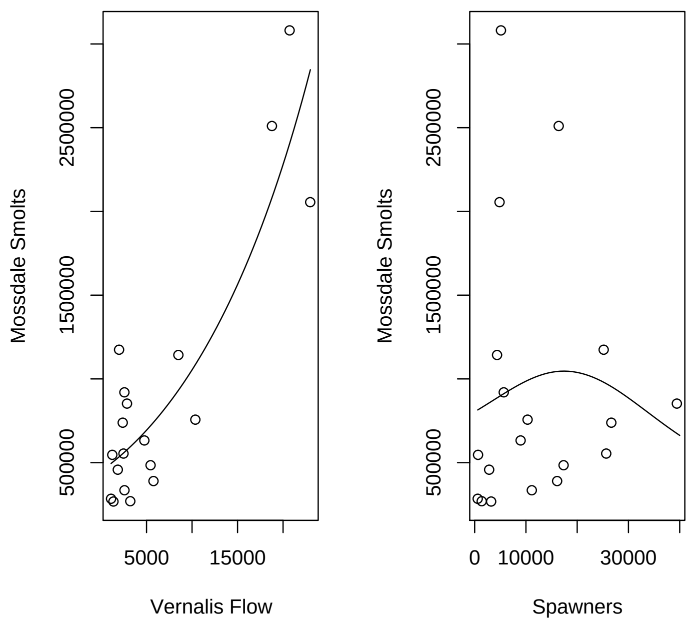

Figure 3: GAM smooth for flow (left) and number Spawners in Fall (right) done separately.

will fit a more standard density-dependent model form as a follow-up to this analysis.

First, we examine separately the relationship of Vernalis flow and escapement in the fall to the estimated total numbers of smolts migrating out past Mossdale in the spring. Given the bandwidth chosen (based upon an algorithm attempting to balance bias and variance) the relationship with Vernalis flow (figure 3) looks almost perfectly log-linear, whereas that with previous fall's spawners looks quadratic (implying a density dependence). However, when we estimate a quadratic relationship and perform a permutation test to derive exact inference of the test of independence of Mossdale smolts and number of estimated spawners, the p-value is .85, suggesting no evidence of any bivariate statistical relationship of Mossdale smolts and estimated numbers of fall spawners. Next, we re-fit with GAM smooths that include both variables and examine whether this lack of relationship of spawners and smolts persists when both variables are in the model. Now, the pattern is concave up (still quadratic, but in the other direction) suggesting that number of spawners in fall and Vernalis flow confound one another in this model; there does appear to be a slight negative

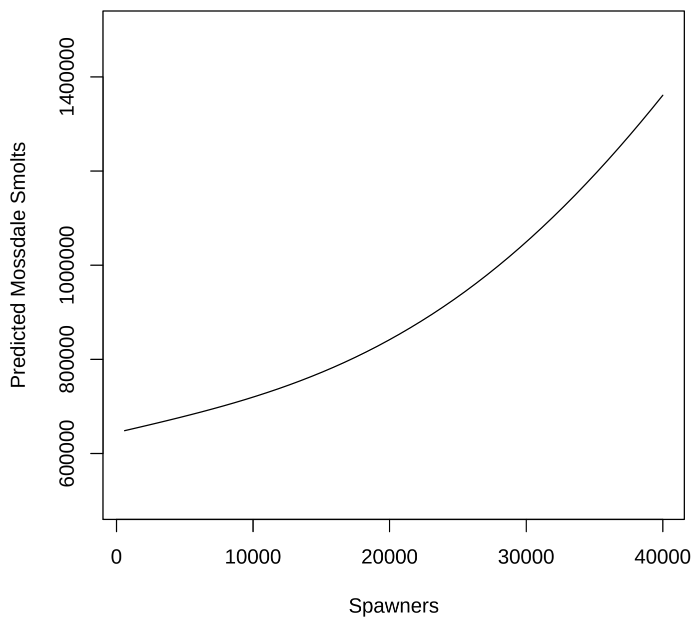

Figure 4: Results of GAM smooth when both Spawners and Vernalis flow are in the model. Plot shows the predicted number of smolts by Spawners when Vernalis flow is set at the average.

relationship of Vernalis Flow in the Spring and the number of spawners in the previous Fall - although certainly not causal, this empricial confounding could explain the difference in the relationship of fall spawners and spring smolts in the unadjusted (without flow) and adjusted (with flow) models. If the number of smolts at Mossdale is modeled as a quadratic (concave up) versus the number of fall spawners, it will blow-up the number of smolts produced if number of spawners gets very large (extrapolating beyond data). To avoid this, we fit a log-linear Poisson regression model of smolts versus log(spawners).

```
Call:
glm(formula = smolts ~ flow + logspawn, family = poisson)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-519.83  -324.37   -64.45   222.23   600.21

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) 1.170e+01  2.410e-03  4857.0   <2e-16 ***
flow        8.082e-05  2.817e-08  2868.6   <2e-16 ***
logspawn    1.481e-01  2.595e-04   570.7   <2e-16 ***
---

    Null deviance: 9958794  on 18  degrees of freedom
Residual deviance: 2127786  on 16  degrees of freedom
AIC: 2128082

Number of Fisher Scoring iterations: 4
```

We then plot the results of the fitted model of predicted smolts versus Vernalis flow at different numbers of fall spawners (see figure 5).

3.1 Model based on Ricker Density Dependence

To examine empirical evidence of density dependence in this portion of the model, we re-fit the model discussed above with a form that practically guarantees some density-dependence (roughly equivalent to a regression of $log(Y/X)$ vs. X will result in a negative association even if Y and X are independent). We use a modified form of the Ricker density dependence (where this dependence can depend on flow) presented in [Speed(1993)], page 280:

$$log(S_t) = \alpha log(F_t) + log(E_{t-1}) - \beta E_{t-1}$$

where S_t is the total number of smolts surviving to Mossdale in year t, E_{t-1} is the number escaping the previous spring, F_t is Vernalis flow, and α and β are parameters. This model can be fit using Poisson regression (log-linear regression) of S_t versus and F and E_{t-1} and entering an offset of $log(E_{t-1})$. We fit this model and derived the following results:

```
> summary(glm.1)

Call:
glm(formula = smolts ~ flow + spawners + offset(logspawn), family = poisson)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
```

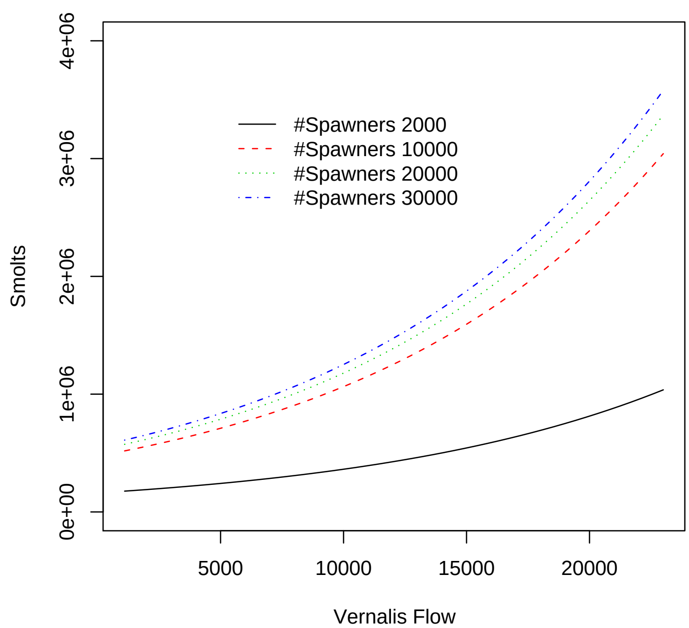

Figure 5: Poisson model showing predicted smolts versus flow at different numbers of fall spawners.

```
-752.30  -278.74    95.66   429.31   989.56

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  4.969e+00  6.548e-04    7589   <2e-16 ***
flow         6.919e-05  3.143e-08    2201   <2e-16 ***
spawners    -7.304e-05  3.139e-08   -2327   <2e-16 ***
---

    Null deviance: 23077076  on 18  degrees of freedom
Residual deviance:  4567802  on 16  degrees of freedom
AIC: 4568099

Number of Fisher Scoring iterations: 5
```

which suggests a significant negative relationship of spawner abundance and smolts at Mossdale, or evidence of density dependence. We now plot predicted numbers of smolts versus flow at different numbers of fall escapement (figure 6).

Thus, though there is no strong bivariate relationship at all of spawners and smolts (see figure above), when a Ricker-type of model is used and flow is in the model, there is what appears to be a significant density-dependence relationship. Thus, whether or not there is density-dependence depends on what model is chosen to fit the data (this result appears to contradicts the model shown in figure 5). However, we follow-up by examining which of these two models fits the data better, providing evidence for or against strong density dependence. To do so, we use model selection criteria, Aikake's Information Criteria (or AIC), where the bigger the statistic, the worse the fit. The results suggest a much, much better fit to the data for the simple log-log linear model shown in figure 5 ($AIC = 2.12x10^6$) relative to the Ricker-type model shown in figure 6 ($AIC = 4.57x10^6$). Thus, the data suggests there is little evidence of density dependence being a driving factor given the recent historic numbers of spawners and flows, which make up the current data set.

4 Delta Survival

Once the annual smolt abundance is apportioned on a daily basis in each year (e.g., 1967 through 2000), using either HORB-in or HORB-out, a Delta smolt survival relationship is applied. The number of smolts arriving at Mossdale, combined with Vernalis flow level, are associated with the number of smolts reaching Chipps Island each day via a statistical model. In this case, we use the data provided by Ken Newman, when used VAMP-flow and combined both the release experiments at Durham Ferry and at Mossdale and adjust the survival estimates relative to releases at Jersey Point. Using Dr. Newman's notation, we have the survival estimate as:

$$\hat{S}_{DF \to CI} = \frac{(Y_{DF \to Ant} + Y_{DF \to CI} + Y_{DF \to Oc})/R_{DF}}{(Y_{JP \to CI} + Y_{JP \to Oc})/R_{JP}}$$

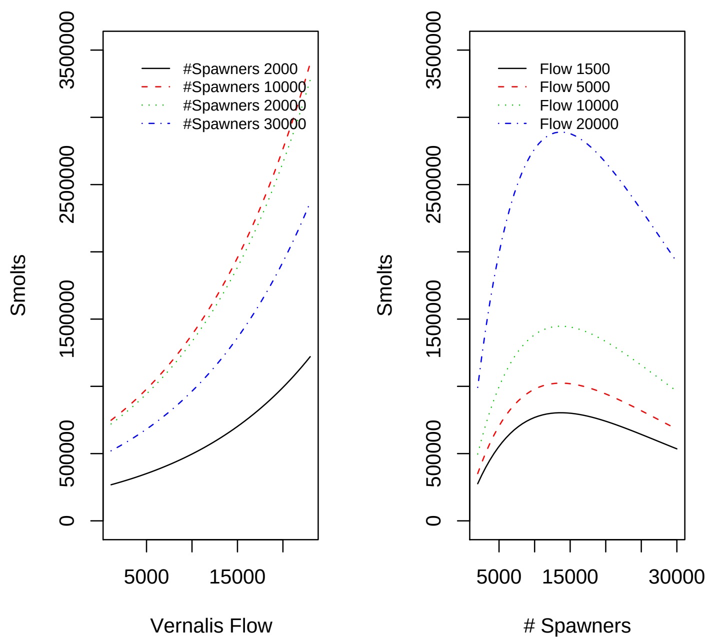

Figure 6: Poisson model showing A) predicted smolts versus Vernalis flow at different numbers of fall spawners for Ricker type of Model and for the same model B) predicted smolts versus number of spawners at different Vernalis flows.

with MD in place of DF when Mossdale is used. This resulted in the following data set used for this analysis.

Table 1: Survival Estimates using Ken Newman's approach for Estimating Survival

	VAMP.Year	HORB	surv.md.chipps	surv.DF.chipps	MD.flow.raw
1	1985.00	0.00			2475.00
2	1986.00	0.00			7140.00
3	1987.00	0.00			2480.00
4	1989.00	0.00			2500.00
5	1989.00	0.00			1945.00
6	1990.00	0.00			1400.00
7	1990.00	0.00			1400.00
8	1991.00	0.00			
9	1994.00	0.00	0.14		1580.00
10	1994.00	1.00	0.13		3115.00
11	1995.00	0.00	0.79		18700.00
12	1995.00	0.00			21250.00
13	1995.00	0.00			23100.00
14	1996.00	0.00	0.14		6665.00
15	1996.00	0.00	0.04		6565.00
16	1996.00	0.00			
17	1997.00	1.00	0.46		6135.00
18	1997.00	1.00			
19	1997.00	1.00			
20	1998.00	0.00	0.41		24950.00
21	1998.00	0.00	0.21		20250.00
22	1999.00	0.00	0.36		6905.00
23	2000.00	1.00	0.31	0.35	6995.00
24	2000.00	1.00		0.18	5969.00
25	2001.00	1.00	0.22	0.24	4170.00
26	2001.00	1.00	0.11	0.13	4145.00
27	2002.00	1.00	0.16	0.14	3255.00
28	2002.00	1.00	0.07	0.06	3356.00
29	2003.00	1.00	0.02	0.02	3345.00
30	2003.00	1.00	0.01	0.01	3370.00
31	2004.00	1.00	0.01	0.01	3160.00
32	2005.00	0.00		0.07	8195.00
33	2005.00	0.00		0.05	9085.00
34	2006.00	0.00	0.11		29350.00
35	2006.00	0.00	0.01		24650.00

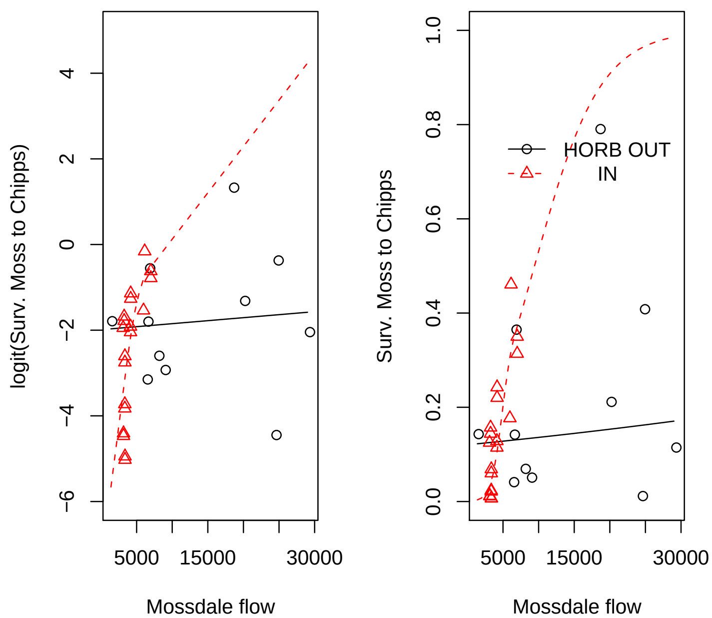

Figure 7: Results of GAM smooth for survival from Mossdale/Durham Ferry to Chipps Island versus Mossdale flow separately by HORB in and out. The plot on the left is on logit scale, the right on probability scale.

As above, we first fit *gam* smooths (logit link) to examine the smooths by HORB both in and out. Figure 7 shows that a logit-linear fit for the HORB out is suggested by the smooths, as well as a logit-linear fit for the HORB In (note, the bend at the end is pure extrapolation), so we used a linear regression model on the logit scale to derive the coefficients by HORB -status.

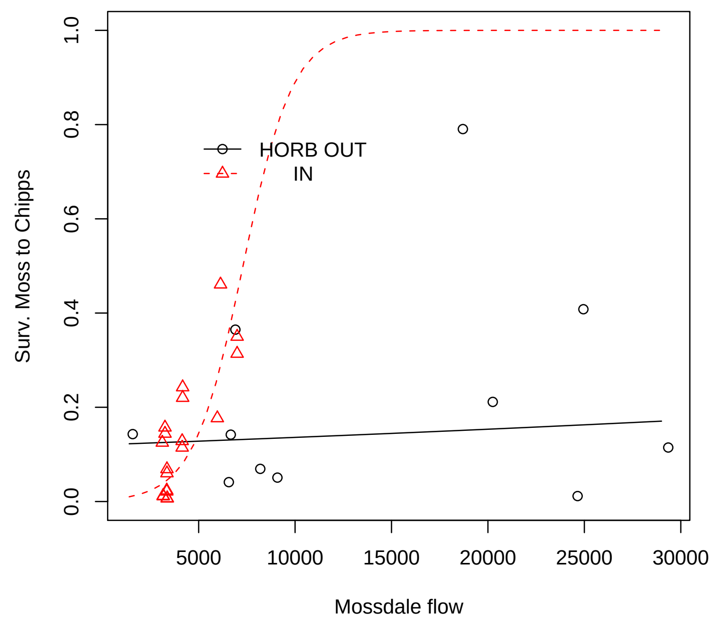

Figure 8: Results of LM fit for survival from Mossdale to Chipps Island versus Mossdale flow separately by HORB in and out.

Figure 8 suggests a strikingly different survival function depending on the HORB status, although there is not data for the HORB in at high flows, so the curve beyond a flow of around 7,000 is pure extrapolation.

5 Hatchery

The conceptual model for hatchery augmentation includes: 1) estimate the fraction of inriver escaping salmon that would migrate into the hatchery; 2) estimate the female fraction of total hatchery escapement ratio ; 3) estimate the number of smolts that would be produced by the number of salmon migrating into the hatchery ; 4) estimate salmon smolt survival as a function of spring flow in each SJR east-side tributary; 5) estimate hatchery smolt survival through the South Delta ; 6) estimate the adult salmon production cohort for each brood year; and 7) add hatchery cohort production to wild cohort production; 8) reconstruct combined wild and hatchery produced SJR salmon escapement; and 9) subtract hatchery escapement from wild escapement for future year cohort production and escapement prediction.

5.1 Fraction of Escapement that goes to Hatchery

Following the pattern of the above model fitting procedures, we first fit a logistic smooth of proportion of spawners entering hatchery in Merced River versus the total of escapement in that river.

Table 2: Data used to Estimate Proportion of Escapement Into Hatchery

	Year	X	MRH	Female	Male	In.River	Total
1	1970		100			4700	4800
2	1971		200			3451	3651
3	1972		120			2528	2648
4	1973		375			797	1172
5	1974		1000			1000	2000
6	1975		700			1700	2400
7	1976		700			1200	1900
8	1977		661			350	1011
9	1978		100			525	625
10	1979		227			1920	2147
11	1980		157			2849	3006
12	1981		924			9491	10415
13	1982		189			3074	3263
14	1983		1795			16453	18248
15	1984		2109			27640	29749
16	1985		1211			14841	16052
17	1986		650			6789	7439
18	1987		958	156	802	3168	4126
19	1988		457	206	251	4135	4592
20	1989		82	32	50	345	427
21	1990		46	14	32	36	82
22	1991		41	9	32	78	119
23	1992		368	41	327	618	986
24	1993		409	153	256	1269	1678
25	1994		943	282	661	2646	3589
26	1995		602	196	406	2320	2922
27	1996		1141	361	780	3291	4432
28	1997		946	397	549	2714	3660
29	1998		799	304	495	3292	4091
30	1999		1637	383	1254	3129	4766
31	2000		1946	937	1009	11130	13076
32	2001		1663	703	960	9181	10844
33	2002		1838	797	1041	8800	10638
34	2003		549	248	301	4110	4659
35	2004		1050			3000	4050

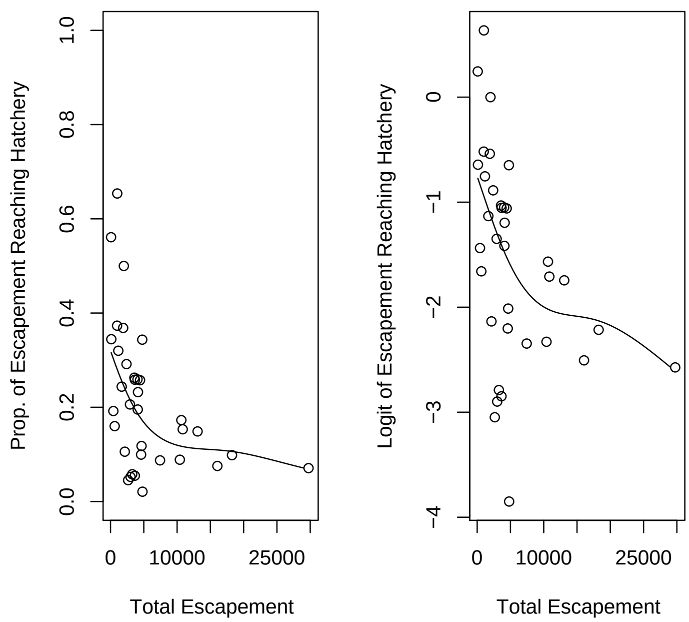

Figure 9: GAM smooth of Proportion of Escapement into Hatchery as function of total Escapement in Merced River in both probability (left) and logit (right) scale

Table 2 has the data used in the following analyses. Using this data and a generalized additive model approach, we get the following fits on both the probability and logit scale. Figure 9 indicates that something more quadratic than logit-linear might fit the data better, and so we fit the data with a quadratic model as follows:

$$Pr(Hatchery \mid TotalEscp) = \frac{1}{1 + exp(-(b_0 + b_1 TotalEscp + b_2 TotalEscp^2))}$$

Figure 10 shows that the quadratic curve is a good fit to the data, which is further supported by the following results of the model fit:

```
> summary(glm.1)

Call:
glm(formula = prop ~ wt + wt2, family = binomial(), weights = wt,
    na.action = na.omit)
```

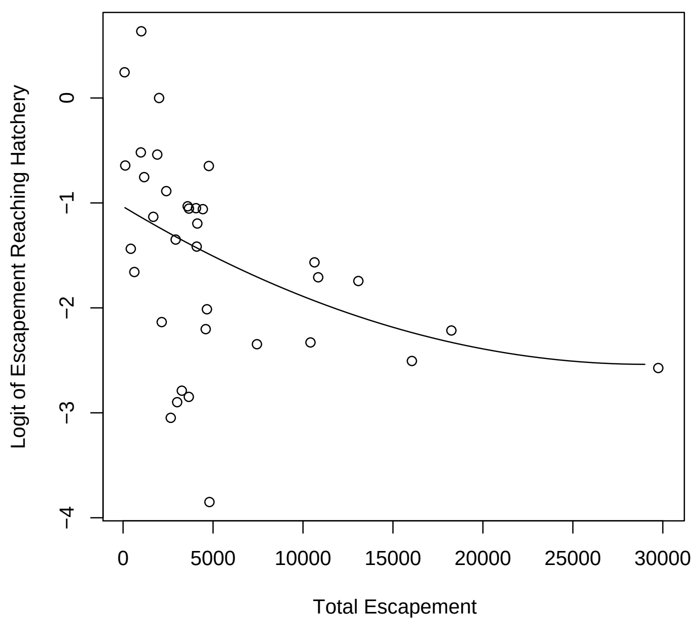

Figure 10: Quadratic GLM fit of Proportion of Escapement into Hatchery as function of total Escapement in Merced River in logit scale

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-35.741  -12.621    2.046    9.027   27.601

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.035e+00  1.558e-02  -66.43   <2e-16 ***
wt          -1.033e-04  2.906e-06  -35.54   <2e-16 ***
wt2          1.773e-09  9.442e-11   18.78   <2e-16 ***
---

    Null deviance: 12765.6  on 34  degrees of freedom
Residual deviance:  8356.7  on 32  degrees of freedom
AIC: 8634.7

Number of Fisher Scoring iterations: 4
```

which shows a significant quadratic term (wt2). We note that there is no easy biological explanation for this quadratic pattern, but for now we retain this functional form as it is a much better fit to the data than say the logit-linear model and given our philosophy surrounding Version 1.5 is to err on the side of empiricism. In Version 2.0, we will concentrate more on biological interpretation of the constituent models.

5.2 Proportion of Females versus Total Escapement into Hatchery

Going straight to the conclusion, the results here are identical to above - a quadratic logistic-linear model where the probability of being female is quadratically related to the total number of fish (males+females) escapement into the hatchery. The data used is precisely the same as shown in the table, for those years with observed numbers of females. Figure 11 indicates that something more quadratic than logit-linear might fit the data better, and so we fit the data with a quadratic model as follows:

$$Pr(Female \mid TotalHatchery) = \frac{1}{1 + exp(-(b_0 + b_1 TotalHatchery + b_2 TotalHatchery^2))}$$

Figure 12 shows that the quadratic curve is a good fit to the data, which is further supported by the following results of the model fit:

```
> summary(glm.1)

Call:
glm(formula = prop ~ wt + wt2, family = binomial(), weights = wt,
    na.action = na.omit)
```

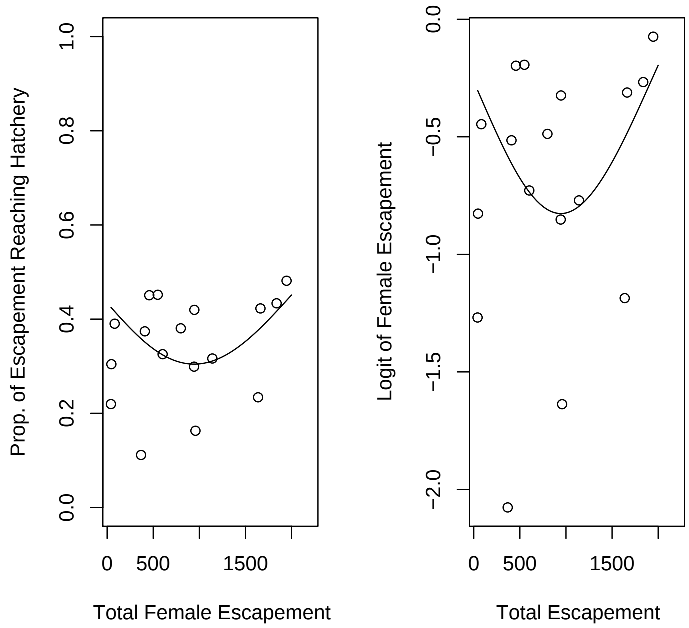

Figure 11: GAM smooth of Proportion of Females as function of total Escapement into Merced River Hatchery in both probability (left) and logit (right) scale

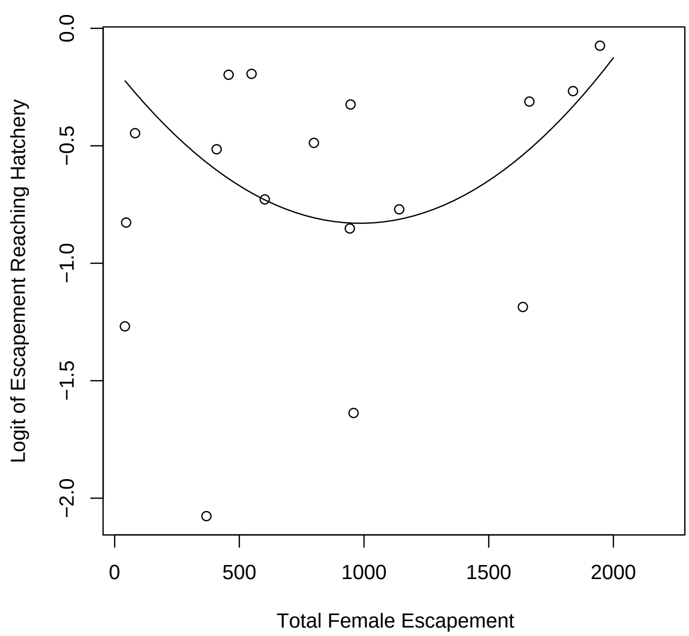

Figure 12: Quadratic GLM fit of Proportion of Females as function of total Escapement into Merced River Hatchery in logit scale

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-11.6746   -1.9276    0.6683    4.0782    7.5215

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.686e-01  9.741e-02  -1.730   0.0835 .
wt          -1.344e-03  1.843e-04  -7.293 3.03e-13 ***
wt2          6.827e-07  7.544e-08   9.050  < 2e-16 ***
---

    Null deviance: 696.83  on 16  degrees of freedom
Residual deviance: 528.32  on 14  degrees of freedom
  (18 observations deleted due to missingness)
AIC: 645.83

Number of Fisher Scoring iterations: 4
```

which shows a significant quadratic term (wt2).

5.3 Smolts per female

Table 3: Data used to Estimate Smolts/Female in Hatchery

	Year	totalescpape	mrhescape	Females	Total.Eggs	Eyed.Eggs	smoltsp
1	1987	3168	958	156	609133.00	445850.00	2286.41
2	1988	4135	457	206	1069258.00	790799.00	3071.06
3	1989	345	82	32	172053.00	103795.00	2594.88
4	1990	36	46	14	59919.00	23273.00	1329.89
5	1991	78	41	9	48075.00	19310.00	1716.44
6	1992	618	368	41	203454.00	121742.00	2375.45
7	1993	1269	409	153	740020.00	559721.00	2926.65
8	1994	2646	943	282	1569937.00	1047887.00	2972.73
9	1995	2320	602	196	977637.00	650031.00	2653.19
10	1996	3291	1141	361	1736391.00	1267974.00	2809.91
11	1997	2714	946	397	1985782.00	1661035.00	3347.17
12	1998	3292	799	304	1210055.00	1037789.00	2731.02
13	1999	3129	1637	383	1862840.00	1573540.00	3286.77
14	2000	11130	1946	937	5299480.00	3855560.00	3291.83
15	2001	9181	1663	703	2947812.00	1799565.00	2047.87
16	2002	8800	1838	797	3348581.50	2059304.70	2067.06
17	2003	4110	549	248	1249074.60	947082.00	3055.10

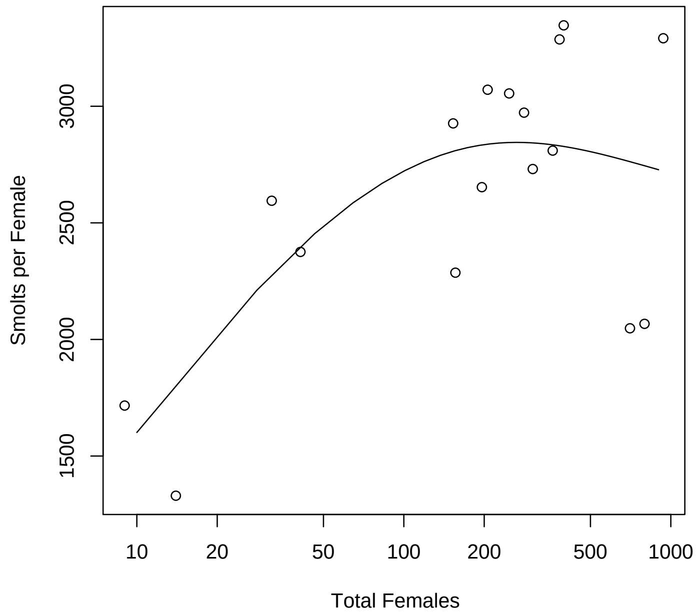

Figure 13: GAM smooth of smolts per female vs. log(females).

We use the data in table 3 to estimate the relationship of smolts per female and used the same sequence of analyses. Specifically, we look at the smolts per female as a function of the log(females) in the hatchery. Using a GAM approach, we see again that the curve looks somewhat quadratic (13).

Thus, we fit a quadratic curve:

$$E(Smolts|Females) = b_0 + b_1 Females + b_2 Females^2$$

resulting in the fit presented in figure 14. The resulting fit suggest that the quadratic effect fits significantly better than the linear model.

```
> summary(glm.1)

Call:
glm(formula = smoltsp ~ logfem + logfem2, data = smoltsper, na.action = na.omit)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-675.37  -209.34    86.95   273.57   659.25
```

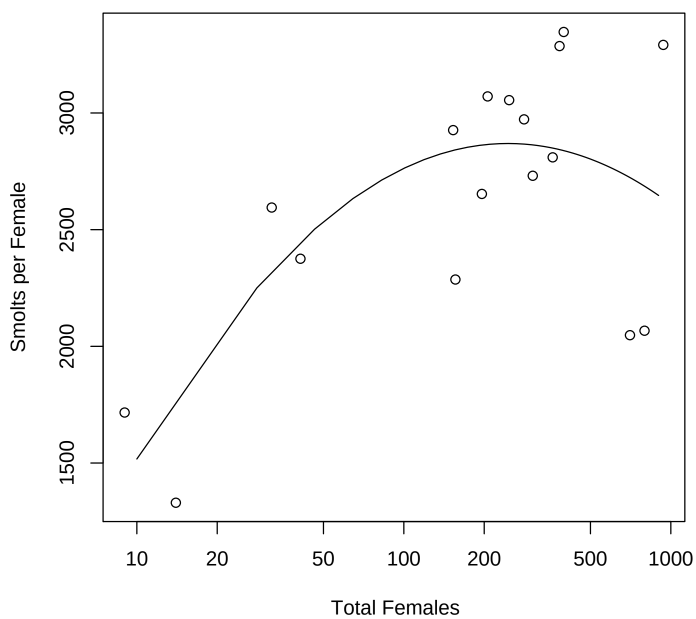

Figure 14: GLM fit of smolts per female vs. quadratic log(females).

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1127.35    1185.90  -0.951   0.3579
logfem       1452.38     550.09   2.640   0.0194 *
logfem2      -131.95      59.92  -2.202   0.0449 *
---

    Null deviance: 5371011  on 16  degrees of freedom
Residual deviance: 2569205  on 14  degrees of freedom
AIC: 258.98

Number of Fisher Scoring iterations: 2
```

One factor will make the number of smolts per female decline with increasing number of females and that is that the maximum egg retention is two million. Thus, after two million eggs, adding more females will just drop the rate of smolts per female as the number of eggs is no longer increasing. This is probably contributing to the curve starting to descend at higher numbers of females.

5.4 Survival of Smolts to Confluence with Main Stem

The final model for the hatchery is to migrate the hatchery smolts out of the tributary in the main stem and in this case we can use release-capture experiments to estimate the survival. In this case, we used data that includes calculated survival estimates for release experiments in the 3 tributaries and corresponding flow, shown in the following table:

Table 4: Data used for Estimating Survival of Smolts in Tributaries to confluence iwth main stem of SJR

	River	Year	Date	Flow	FlowIndexBankFull	Temperature	Surv
1	TR	2002	4/25/06	1274	0.42	15.90	0.53
2	TR	2001	4/23/05	635	0.21	17.30	0.18
3	TR	2000	4/14/04	2982	0.99	13.10	0.28
4	TR	1999	4/18/03	1960	0.65	14.20	0.19
5	TR	1998	4/16/02	4050	1.34	12.10	1.03
6	TR	1997	4/23/01	1436	0.48	14.70	0.44
7	TR	1996	4/27/00	2664	0.88	13.40	0.32
8	TR	1995	5/5/99	8217	2.72	11.30	0.79
9	TR	1990	5/1/94	241	0.08	19.40	0.30
10	TR	1987	4/17/91	563	0.19	17.60	0.42
11	SR	2003	4/26/07	1300	0.71	15.00	0.57
12	SR	2002	5/2/06	825	0.45	18.00	0.41
13	SR	2000	5/19/04	1500	0.82	16.10	0.57
14	SR	1989	4/21/93	900	0.49	17.80	0.37
15	SR	1988	4/27/92	900	0.49	15.60	0.54
16	SR	1986	4/29/90	1200	0.65	16.70	0.59
17	MR	2004	5/10/08	1600	1.20	18.70	0.36
18	MR	2004	4/20/08	480	0.36	13.00	0.16
19	MR	2004	4/28/08	846	0.63	15.00	0.12
20	MR	2003	4/26/07	570	0.43	18.00	0.26
21	MR	2003	5/5/07	1380	1.03	17.00	0.25
22	MR	2003	4/14/07	650	0.49	12.00	0.20
23	MR	2002	4/22/06	1800	1.35	16.90	0.18
24	MR	2002	4/1/06	400	0.30	16.40	0.01
25	MR	2001	5/9/05	1099	0.82	18.10	0.34
26	MR	2001	4/22/05	1165	0.87	16.40	0.32
27	MR	2000	4/28/04	1556	1.16	13.30	0.30
28	MR	2000	4/13/04	364	0.27	16.10	0.22
29	MR	1999	4/15/03	1700	1.27	14.40	0.70
30	MR	1999	5/6/03	1500	1.12	13.90	0.17
31	MR	1998	4/13/02	2600	1.94	10.00	1.02
32	MR	1998	5/4/02	2500	1.87	12.80	0.69
33	MR	1997	4/21/01	900	0.67	13.90	0.33
34	MR	1997	5/14/01	600	0.45	16.10	0.00
35	MR	1996	4/26/00	1300	0.97	14.40	0.82
36	MR	1995	5/4/99	3700	2.77	12.80	0.58
37	MR	1994	4/23/98	700	0.52	13.90	0.34

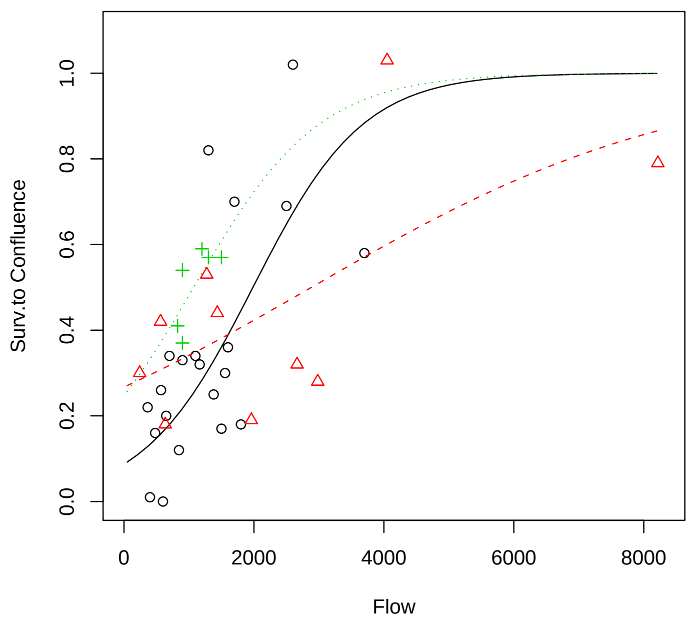

Figure 15: Results of least-squares logit fit for survival in tributaries to confluence with SJR.

We have very few data points per tributary, thus we have limited power to do exploratory analyses. Thus, we use a simple approach fitting logit-linear models of survival versus flow of the form:

$$Pr(Survive \mid flow, Trib = t) = \frac{1}{1 + exp(-(b_{t,0} + b_{t,1} flow))},$$

so for each tributary it is a 2 parameter model. As one can see, 2 of the data points have values with undefined logit transform (either 0 or > 1) - we used an arbitrary cut-off for these "outliers", truncating points at 0.05 and 0.95, respectively. The following shows the raw data (including these outliers at their original values) and the resulting fits by tributary.

References

[Hastie & Tibshirani(1990)] Hastie, T. & Tibshirani, R. (1990). *Generalized additive models*. New York: Chapman and Hall.

[R Development Core Team(2005)] R Development Core Team (2005). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org`.

[Speed(1993)] Speed, T. (1993). Modelling and managing a salmon population. In V. Barnett & K. Turkman, eds., *Statistics for the Environment*. New York: Wiley, 267–93.