Appendix 22 Passage Models and Reach Survival Data

Wilson,P., Schaller, H., Petrosky, C., Bouwes, N. and P. Budy

In this document we discuss issues of relating passage model output to research survival study estimates, provide a critique of the TURB hypotheses, respond to criticisms of the FLUSH model, and lastly comment on Williams et al. 1998 attachment 4.

I. Model Calibration

Understanding the performance of the passage models in predicting survival estimates requires analysis of *how* the models fit the survivals; not just how well they match the values. Judgement of how well passage models fit survival estimates requires to consideration a few key issues: 1) how many parameters are in a particular model, 2) what year-specific inputs are needed, and 3) what parameter adjustments are made to fit the survival data. Models that best fit survival data with the lowest number of each of the above elements should provide more reliable estimates of future stock response.

Anderson (1998b) states: "Comparison of the models to survival data given in Table 4.4 of the Weight of Evidence report (July 3 1998) is irrelevant because the FLUSH model was calibrated to this. Note CRiSP, which fits the data equally well as FLUSH did in its calibration, was calibrated independent of the data used to calibrate FLUSH." (p. 4)

It appears that CRiSP is calibrated to survival data. The method of "evaluating" CRiSP with 1966-83 survival estimates in the spring/summer chinook retrospective modeling process was described by Zabel (1996d): "To fit CRiSP to the survival estimates, the strategy was first to configure the model according to all relevant information on dam operations and passage conditions. Second, model dam passage parameters were adjusted to reflect reported dam passage information. Finally, the model was fit to the survival information… For all years except 1973 and 1977, the two lowest flow years, a fit within a few days was obtained by adjusting the model parameter T_{seasn}. This parameter characterizes how quickly the flow dependent component of migration increase over [the] season...In 1973 and 1977 fish moved significantly slower than in all other years and the arrival time at dams could not be fit by adjusting T_{seasn} only. In both those years a good fit to the data was obtained by adjusting the maximum flow independent migration rate β_{max}… For the years 1969 through 1983 no additional model parameters were altered. To fit survival between release site and the upper dam for 1966-69 the predator density above Ice Harbor dam was set at 200 predator/sq km. This represents a density 50% lower than is estimated for the present levels in Lower Granite reservoir, reflecting the difference between a regulated pool and a free-flowing stream…The observed and predicted survivals in these years [1973 and 1977] were closely fit by adjusting the seasonal flow independent component of migration." Different parameters have been varied to fit PIT tag survival estimates: "The procedure involved adjusting FGEs in the CRiSP model to make the predicted PIT tag recoveries at each dam equal to the observed recoveries. In this way FGE [was the] only parameter that was adjusted to fit the observed PIT tag recoveries at the dams…The fit involved a step wise process in which FGE was adjusted sequentially from Lower Granite dam down to McNary dam. In this approach the FGE at a specific dam was dependent on the FGEs of the dams upstream." (CRiSP.1.5 Manual).

In addition to daily flow, spills, and reservoir volumes CRiSP requires input of daily temperatures to predict survival rates. The requirement for daily temperatures introduces another level of error in prospective modeling. This information is not required for the FLUSH model.

CRiSP has many more parameters than FLUSH to determine their relationship of reservoir survival to FTT. These include parameters for a function relating mortality to percent nitrogen supersaturation, prediction of fish density as a function of depth to determine TDG levels experienced during the migration, reservoir- and tailrace-specific predator densities for three major predators, and a consumption rate function for each predator as a function of temperature. As (Zabel 1996c) notes, "The relationship between flow and survival, however, is a result of several interacting submodels."

"CRiSP.1 version 5 (CRiSP.1.5) is a complex model, with hundreds of parameters. It is impossible to examine the potential interactions of all of these parameters" (Zabel 1996c). Some of the parameters which have been examined for sensitivity include (Zabel 1996c): reach predator density, forebay predator density, tailrace predator density, predator activity exponent, velocity variance, migration rate parameters. These are in addition to parameters which are required in both CRiSP and FLUSH, such as fish guidance efficiencies (FGEs), turbine mortality, spill mortality, and bypass mortality.

Travel time in CRiSP is modeled by two submodels: a reach model, which "[m]oves groups of fish through individudal reaches according to specified migration rate and rate of population spreading" and a migration rate model which "[d]etermines migration rate as a function of river flow, date in season, and duration of migration time. Migration rate varies on a per reach and per time-step basis" (Zabel 1996a). Zabel (1996c) notes that "Much of the behavior of the model depends on the parameters used to describe fish movement within the system, and many different stock behaviors can be produced by varying one or several of these parameters." Among the parameters in the equations used to predict fish migration are (Zabel 1996c): initial flow-independent velocity of smolts, final flow-independent velocity of smolts, date of the inflection point in the flow related term of the migration rate equation (T_{seasn}), and smolt start/stop date. There is also an "α" parameter which "determines how quickly the fish mature from early season behavior to later season behavior" (Zabel 1998a), and a "velocity variance" parameter, which sets the variability in the migration velocity, and represents variability from all causes including water velocity and fish behavior. This variance can vary on a daily basis (CRiSP.1.5. Manual). From manual: "The model-predicted average travel times are then compared to observed average travel times. Migration rate parameters are selected that give the best model fit to the data."

The relationship between spill at dams and smolt mortality in CRiSP is complex: "CRiSP.1.5 models mortality due to gas saturation as a function not only of the saturation but also the depth distribution of fish: the deeper fish swim, the less affected by gas bubble trauma they are" (Zabel 1996c). To use CRiSP prospectively, not only must the rate of mortality of fish exposed to a given level of total dissolved gas (TDG) be modeled, the level of TDG experienced at different depths throughout the hydrosystem for given flow and spill conditions must be modeled, and the depth at which fish are found in different projects must be predicted.

The mechanistic structure of CRiSP is dependent on a level of detail for many categories of temporal and spatial data that is sparse or does not exist. This mechanistic modeling approach coupled with the limitations on available data from experiments and monitoring leads to making guesses about a number of key parameters. There are a number of key assumptions that are made

implicitly and explicitly in order to construct the underlying function which relates mortality rate to TDG levels. These key assumptions in CRiSP 1.5 (Anderson et al. 1995) are: a) laboratory experiments performed at relatively shallow depth (2.5 m) provide better estimates of cumulative survival than studies performed where fish are able to move volitionally to greater depths in riverine conditions; b) there is a mortality-length relation which is linear and independent of species or life history characteristics; c) the length of migrating fish is known and constant when computing the ratio for adjusting the mortality rates of fall chinook which are used to construct the Mortality rate - TDG function; d) the length adjustment and fall chinook mortality rates will have predictive capability for spring chinook; e) the underlying form of the mortality rate vs TDG is that of a broken stick; and f) the parameter determining the inflection point of the broken stick function is selected and is apparently assumed to be known without uncertainty. This function was indirectly developed from cumulative survival vs nitrogen supersaturation estimated in laboratory experiments where fish were confined in relatively shallow tanks of 2.5 meters. The experiments included only buttoned up fry fall chinook from Spring Creek hatchery and yearling steelhead captured during migration in the Snake River. There were no data for spring chinook in these experiments. The CRiSP modelers ignored any studies conducted in deep cages and or in situ when deriving the function. This function can be expected to predict higher mortality rates than one using more relevant deep cage studies.

Mortality/day is indirectly estimated by regression of cumulative survival and days of exposure for 6 levels of nitrogen saturation for fall chinook (105, 110, 115, 120, 124, and 127) and 4 levels for steelhead (110, 115, 120, and 127). The fall chinook data was fit for up to 40 days of exposure where steelhead was for 7 days of exposure.

The assumption is that TDG effects are proportional to body length independent of species or life history type. A length-mortality relationship was developed with the intention of extrapolating laboratory experimental results using one fish length to field conditions where it is assumed fall chinook are larger by a constant proportion. Also, CRiSP claims to extrapolate results of fall chinook experimental results to spring chinook in Lower Granite reservoir using fish lengths. The relation of Mortality/day to length is developed from a regression of regression coefficients (CRiSP 1.5 pgs 5.75-5.77) for cumulative mortality vs days of exposure. Although no mechanism can be determined for justifying a linear relationship (CRiSP 1.5 p 5.75) of mortality/day to length, one is fit. The goodness of fit, for this relation, expressed by the r^2 in CRiSP is misleading because it is a regression of regression coefficients which appear to show a poor fit to the data on the graph (CRiSP 1.5 fig 5.27). The linearity of the mortality/day vs length relation is driven by one point, which is the estimate derived from the steelhead results. The assumption that there is a linear relation between mortality rate and fish length is speculative at best.

The CRiSP computed mortality rates/day at % saturation level are adjusted by the ratio of a CRiSP assumed length of fish in the environment to length of fish in nitrogen mortality experiments. This is justified by the CRiSP assumption of linearity in the relation of mortality rate to fish length. This length adjustment to the computed mortality rates increase the mortality rate values considerably. For fall chinook the adjustment increases the mortality rate by 180% (2.8 * mort. rate) from the CRiSP fitted fall chinook mortality rate. For Spring chinook the adjustment increases the mortality rate by 225% (3.25 * mort. rate) from the CRiSP fitted fall chinook mortality rate. For steelhead the adjustment increases the mortality rate by 27% (1.27 * Mort. rate) from the CRiSP fitted steelhead mortality rate. These adjusted mortality rates significantly increase the mortality rates (CRiSP 1.5 manual table 5.43), which are primarily justified by a

regression driven by one point estimated from steelhead data. In addition, this highly influential adjustment is from a ratio which assumes constant species specific length of migrating fish.

The function used to relate fish mortality rate to nitrogen saturation level is constructed from the CRiSP fitted mortality rate/day estimates adjusted for experimental length to assumed length in the environment. The form of the CRiSP mortality function is birectilinear (broken stick) (CRiSP 1.5 manual equ. 5.32 pg 5.78). That is to say, there is a component with a shallow slope of increasing mortality rate and a component with a steep slope of increasing mortality rate over TDG levels (CRiSP 1.5 figure 5.29 pg 5.79). This function was fit one segment at a time, which means the steep slope of the function was fit to one to two points per species. In addition, the steep slope portion of the function was fit by constraining the first point. The mortality rates predicted by this function are hypersensitive to the selection of the inflection point between the shallow slope and steep slope part of the function. The variable that creates the inflection point in CRiSP is N_{crit}. The documentation is confusing as to what N_{crit} value was selected in the analyses. The N_{crit} value presented in Zabel 1996b table 1 is 10.9%. However, figure 2 in Zabel 1996b and the figure 5.29 in CRiSP 1.5 graphically presents a function that would have had to use a N_{crit} value of 20%. If an N_{crit} value of 10% was used the function would not fit closely to the points presented in CRiSP 1.5 fig. 5.29. Knowledge of the value selected for N_{crit} variable is essential to understand the underlying driving function which would be used in any analyses of spill. The prediction of mortality rate per day, for spring chinook at a TDG of 125%, would triple from 3% to 9% (200% increase) by using a N_{crit} value of 20% vs 10% respectively (see attached graphs 1 and 2). When computing the cumulative mortality associated with TDG levels, this difference becomes very significant. For example, if spring chinook travel time through the Snake and Columbia rivers was 25 days and at a TDG of 125% the cumulative mortality due to nitrogen saturation in CRiSP would be approximately 53% (47% survival) for an N_{crit} of 20% vs 91% (9% survival) for an N_{crit} of 10%. The fish survival predicted by CRiSP appears to be extremely dependent on the selection of N_{crit}, due to the broken stick structure of the mortality/day function.

The survival results appear to be highly sensitive to a number of these key assumptions as described in the above comments and further work will be needed to assess the model sensitivity to others, particularly the use of only laboratory study survival results for the basic data.

In contrast, FLUSH fits reservoir survival estimates to fish travel time using a model containing two parameters (PATH Retrospective Report on spring/summer chinook 1996 Chp. 6). A relationship between fish travel and water travel time is derived for three sections of the juvenile migration corridor. These relationships are used to estimate fish travel time. Each relationship consists of two parameters, for a total of six. These are all the parameters that are used to predict prospective reservoir survival in the FLUSH model. FLUSH achieves similar values in the measure of goodness of fit to survival data (Table 4.4 in WOE report), with far fewer parameters than CRiSP. FLUSH TURB5 had the best fit (lowest AIC/BIC scores) compared to spawner-recruit data within the alpha model, and CRiSP TURB4 had the poorest fit (Table 4-2 WOE report). It is also important that *the AIC/BIC scores did not take into account the number of parameters in the passage models*.

Anderson (1998b) provided survival estimates from Lower Granite Dam to Bonneville Dam for 1997 and 1998, to compare to the PIT tag survival estimates reported in Williams et al. (1998). The 1997 and 1998 (Williams et al. 1998) survival estimates are using new procedures which have not been reviewed and are critiqued below. The CRiSP estimate in 1998 was higher than that in 1997 (55% to 48.5%). Since 1997 was a higher flow and spill year, and there is a strong relationship of FTT to WTT (Decision Analysis Appendix A, pg. 12) in CRiSP, the lower survival in 1997 has to be due to the spill mortality function or release distribution, or a combination of both, unless predator activity coefficients or other parameters were adjusted. Note that CRiSP reservoir survival vs. FTT relationship with and without gas mortality comparison called for (in the Decision Analysis pg. 28 and 29 Appenix A) was never provided. As we've seen, it appears the release date of fish above the first dam is varied to fit the survival data. Even if CRiSP plugged in PIT tag release distribution from studies, rather than adjusting it, they have no way of prospectively estimating the survival of the general migration without taking a guess at the distribution.

II. Response to Anderson (1998b)

1. "Temperature and total dissolved gas levels in CRiSP combine into an exponential coefficient C" (p.2)

Response:

CRiSP reservoir survival vs. FTT relationship with and without gas mortality comparison called for (pg. 28 and 29 in Appendix A) was never provided. Therefore it is impossible to tell how much impact flexibility in TDG (total dissolved gas) - induced mortality has on relationship. This is discussed in more detail above. CRiSP has many more parameters than FLUSH to determine their relationship of reservoir survival to FTT. These include parameters for a function relating mortality to percent nitrogen supersaturation, prediction of fish density as a function of depth to determine TDG levels experienced during the migration, reservoir- and tailrace-specific predator densities for three major predators, and a consumption rate function for each predator as a function of temperature.

2. "The resulting [upside]-down logistic equation is unique to ecology. No similar relationships or bioligical evidence have been cited." (p. 3)

Response:

In the FLUSH model, impacts such as descaling and energy depletion from dam and reservoir passage are hypothesized not only to instantaneously kill some smolts; the cumulative effects are also posited to weaken some of the survivors and decrease their probability of surviving a given day as they migrate lower in the hydrosystem. Survival as a function of time in the FLUSH model is of the form commonly referred to in ecological texts as 'Type 1' (e.g., Fig. 2.4 in Whittaker 1975). Taking the complement of the survival vs. FTT relationship gives a sigmoidal cumulative mortality vs. time curve. This is exactly the form one would expect if there were sub-lethal impacts on the surviving fish from each project passed. Lethal exposure times (i.e. LE_{100}) to different stressors are ubiquitous in the biological literature. Cumulative mortality in these studies is usually represented as a logistic (sigmoidal) curve rather than a line; i.e, the rate of mortality changes with time.

3. “In FLUSH fish traveling together will have different rates of mortality if they were release[d] at different locations and times. This produces strong differences between fish and indeterminancy for fish released at the top of the hydrosystem” (and more text on p. 4).

Response:

See point 2 and Appendix 23 under **Influence of control survival rate estimates on Ds under the two hypotheses.** In brief: In Anderson’s V_{stress} equation, descaling mortality of non-transported fish is “resolved” in 6 days (Anderson 1998a). Since no change in migration rate of these fish is posited or modeled, control fish thus stressed should complete the migration from point of release to the end of the hydrosystem in the same amount of time as run-of-the-river fish. Therefore, under this hypothesis, the mortality rate per day of smolts increases with cumulative hydrosystem passage effects. This effect is similar, in theory to that which Anderson has criticized FLUSH, which is implicitly assumed in fitting the reservoir survival function to empirical estimates of survival over reaches of different length. In Spring CRiSP, and both Fall FLUSH and Fall CRiSP, temperature and predation are explicitly modeled, and mortality rate increases with time over the migration season, since consumption rate and predation mortality is assumed to increase with temperature.

4. “Using these two years [1973 and 1977] is problematic because they are the only two significant outliers from suite of survival estimates dating back to 1966” (p.5)

Response:

In a statistical sense, outliers are data points which have independent variable values which overlap with those of other data points, but have anomalous dependent variable values. The fish travel times for 1973 and 1977 study reaches are the two longest, by a substantial amount (PATH Decision Analysis Report, Figure A.2.1-18a). There are no other data points at those fish travel times to contradict the dependent variable values. They are therefore not outliers, and in fact provide much of the contrast in the available data. In explaining them away, hypothesis TURB4 requires that 2.5 times as many fish died as were descaled due to debris impacts in 1973, and over twice as many times died as were descaled in 1977 (Decision Analysis Report, Table A.2.3-6).

5. “The essential difference is that CRiSP seeks to explain the D by indexing it to the experience of the transported fish, while FLUSH seeks to explain the D data by indexing it to the experience of the non-transported fish” (p. 9) and “A relationship between D and the conditions of the transported fish is inferred but is not needed to estimate the D values [in CRiSP]” (p. 10) and “FLUSH transportation hypothesis produces a biologically unrealistic connection between post Bonneville survival of the transported fish [and] in-river fish” (p. 10) and “For example, a decreased (sic) in passage survival due to adverse passage conditions at McNary dam, would decrease transported fish ocean survival” (p. 11) and “D is neutral to T/C evidence but since D used passage model results the hypothesis is supported by the strength of the CRiSP passage model.” (p. 11)

Response: Problems with CRiSP D's and Vc's are explained in detail in Appendix 22. D's are sensitive to Vc estimates, which are underestimated by CRiSP for years of multiple studies (esp. 1976) and are dependent on drastic TURB4 assumptions. D's used in life-cycle modeling of CRiSP results are also sensitive to the use of wrong measure of central tendency and mis-aggregation of time periods. The 1986 and 1989 D estimates in CRiSP assumed no effect of descaling, as TURB4 includes the assumption that negative impacts from descaling at dams vanished after 1980. Since no plausible reason has been given why descaling was so deadly prior to 1981 but innocuous afterward, the V_c estimates used to calculate 86-95 D should include mortality represented by the V_{stress} equation of Anderson (1998a). FLUSH T/C vs. Vc relationship uses observed reach survivals, not model estimates, prior to 1986. Under either hypothesis, D is calculated from the T/C data and an estimate of control (in-river survival). Every D estimate, in both models, is dependent on the survival rate of non-transported fish. It is unknown which factors affect T/C, but under any hypothesis, it must vary with in-river survival, since avoidance of the direct mortality of the hydrosystem below the collection point is presumed to be the primary benefit of transportation, especially since both models posit D values always or mostly less than 1. Because the FLUSH model T/C is not fixed, but varies inversely with Vc, a decrease in non-transported fish survival rate below collector project results in an increase in T/C. Since in FLUSH model T/C increases with decreasing Vc, underestimates of Vc will be mitigated to some extent by the increase in T/C, since D = T/C * Vc/Vt. There is no such effect to mitigate underestimates (or overestimates) of Vc in CRiSP method, since T/C does not vary with Vc.

6. "The [FLUSH] relationship ignores changes that have occurred to the transport program over two decades of operation." (p. 10)

Response:

There is no evidence of improved SAR of Snake River fish since 1980. SARs of transported fish have been consistently less than the 2% to 6% interim goal defined in PATH (Marmorek and Peters 1998). Not only have SARs remained extremely low, there is no indication that the gap has narrowed between performance of Snake River stocks and down river stocks, as might be expected if transportation and hydrosystem improvements were merely masked by generally poor ocean conditions for all stocks. The differential mortality between Snake River and downriver stocks ("mu") did not decrease over time: the mean of mu by period was 1.5, 1.5, 0.8, 1.5 and 2.1 for 1972-1974, 1975-1979, 1980-1984, 1985-1989, and 1990-1992 brood years, respectively (Deriso et al. 1996; Fig. 5-5). The differential mortality increased significantly (ibid.) as water velocities decreased during the smolt migration. Examination of these data does not support a hypothesis that migration conditions, including transportation survival, were continually improving for Snake River stocks compared to downriver stocks.

III. TURB assumptions

The WOE document contrasts the most disparate aggregate hypotheses (hydro and non-hydro) to highlight results of model runs under different management options. The hydro hypothesis includes the worst case passage scenario while the non-hydro hypothesis includes the best-case passage scenario. We reviewed the assumptions of the non-hydro hypothesis and suggest that these optimistic assumptions are not realistic. The non-hydro passage aggregate hypothesis includes TURB4 assumptions that adjust survivorship of juveniles in the passage models pre-1980. The mechanism used to justifying this adjustment to only pre-1980 years is descaling due to trash in the forebay and by-pass routes. After 1979, hydro project modifications have been suggested to

remedy the problems responsible for descaling. We first discuss the problems of using descaling data to estimate mortality rates. Second, we reveal inconsistencies in the implementation of these descaling. Finally, we illustrate the effects TURB4 survivorship adjustments have on retrospective and prospective survival estimates and ultimately the influence this has on assessing alternative management options.

Anderson (Decision Analysis Report Appendix A) related descaling estimates to mortality estimates from several studies conducted over different time periods. Studies estimating mortality of juvenile salmonids were conducted over a 6 hr, 2 day and 6 day period. To convert these estimates to a common currency, Anderson related descaling to instantaneous mortality. Other attributes of these studies, however, do not allow for direct comparisons simply through converting to instantaneous mortality. The 0.25 and 2 day experiments were conducted in holding tanks where the proportion of dead juvenile salmon was recorded after the experimental period. In contrast, the 6 day studies were conducted by comparing release of juveniles above and below Lower Granite dam and calculated survivorship through the difference in detections between the release sites (Raymond 1979). Therefore, estimation of survivorship from the 6 day studies includes losses due to other factors such as predation. Also, the time mortality occurred in these studies can not be inferred from these studies. This also suggests that the 6 day time interval in which mortality is assumed to occur is arbitrary.

The 2 day experiments were conducted before or after fish were trucked to a site below Bonneville. Thus, delayed mortality due to stress of transportation is inconsistently included in these studies reviewed by Williams and Matthews (1995). Also, these studies used marked fish, which do not include descaled or otherwise unhealthy fish. As Williams and Matthews (1995) do not cite the origin of these data, we were not able to review other potential problems with methods used in these studies. These are a few examples of several of the potential problems of the descaling data used in Anderson's descaling/mortality relationship. As Anderson's descaling relationship produces mortality estimates in excess of 250% of descaled fish, these problems appear to have very large ramifications.

Assuming the data used in this relationship are correct, we find inconsistencies in how this relationship is implemented. The descaling relationship adds mortality to pre-1981 estimates of survivorship and assumes that after this period problems responsible for the descaling no longer occur. However, descaling estimates after this period suggest that descaling problems continue (Figure 1). If descaling is a causal or correlative measure of mortality associated with the dams then this descaling adjustment should consistently be applied to all years rather than to arbitrarily selected periods.

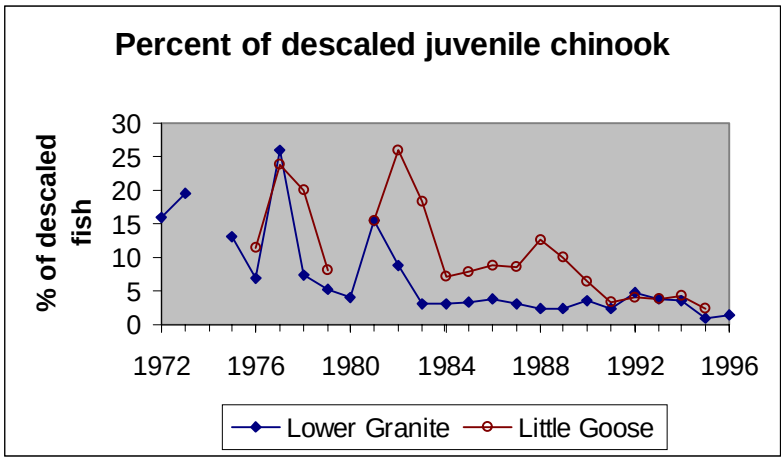

Figure 1. Estimates of the percent of juvenile chinook that are >10% descaled at Lower Granite and Little Goose dams between 1972 and 1996.

Anderson applied descaling adjustments of survival for all in-river fish. Park et al. (1978) indicated that fish sampled in the forebay, which had been swimming in and out of the trash racks, were descaled at approximately half the rate as by-passed fish. Consequently, fish passing through turbines should have descaling related mortality at half the rate of by-passed fish. During the pre-1981 period, nearly all fish passed Lower Granite or Little Goose dams through the turbines and spill; nearly all the fish that entered the bypass routes were diverted and collected for transportation. Therefore, the high rates of descaling for fish encountering screens and blocked orifices do not apply to survival estimates of majority of in-river fish.

The effects of the assumptions of descaling related mortality generally tends to overestimate mortality. The turbine/by-pass mortality associate with descaling is on average 270% of the descaled fish for the pre-1981 period. This large adjustment caused mortality due to descaling from dam passage to exceed mortality in one year of the reach studies. CRiSP adds mortality due to predation in the reservoirs and GBD to this overestimate of mortality due to descaling.

TURB5 applies descaling adjustments to turbine fish at half the rate of by-passed fish. TURB5 also assumes only 100% mortality of descaled fish. While TURB5 applies a much lower descaling mortality adjustment relative to TURB4, this increase in mortality to dam passage is quite liberal. TURB6 adjustments differ from TURB5 by assuming turbine mortality is fixed at 10%. TURB5 and TURB6 assumptions both acknowledge that problems due to dam by-pass routes may differ between years but provides an alternative to the TURB4 assumption of problems through all routes of passage.

Because only post 1980 years are used in the prospective model, applying descaling related mortality (TURB4, TURB5, and TURB6) adjustments arbitrarily to only the pre-1980 period provides over-optimistic expectations of current and future conditions in A1 and A2 management scenarios. In addition to this correction to in-river survivorship, the same descaling adjustment is applied again to V_c in the calculation of D (i.e. V_c fish are given descaling adjustment twice-see discussion on D). While TURB5 and TURB6 provide alternatives to TURB4, we believe these 3

assumptions lead to erroneous conclusions regarding the projected impacts of alternative management actions because of the inconsistent survivorship adjustments across different years.

IV. Response to Williams et al. 1998

Attachment 4. 1989-92 NMFS PIT tag issues:

NMFS has reintroduced into PATH the PIT tag data from 1989-1992, claiming that it provides an "out-of-sample data set" that can be used to validate reach survival estimates of different passage models. They have not adequately documented limitations of these data sets that serve to limit their applicability. In its present state, the data set is more of an "out-of-sample model" with its own unique assumptions.

In Smith and Williams (NMFS Memo, Attachment 4, July 27, 1998) Table 4, there are no true 'observed' survivals for 89-92. The estimates NMFS has provided are approximations not statistical estimates (eg. CJS). NMFS presented 2 methods for estimating survival with their preferred method being #2. It was impossible to reproduce their second alternative form of estimating survival because we could not assertain how they mixed and matched survival estimates from release groups for 94-95 to represent 89-92 survivals. Their estimation procedure is further problematic in that it relies heavily on FGE assumptions for the 3 projects.

We used NMFS numbers of released, detected, and returned to the river from Table 1 and their FGE's for LGR, LGO, & MCN to estimate survival directly from these marks. NMFS reported survival estimates for release to Lower Granite of 90% in 1991 and 77% in 1992 under Factor 2 in their discussion. These values were reproduced in our Table 1. NMFS did not report the survival estimates (release to LGR) with this method of 108% in 1989 and 116% in 1990, which seem optimistic. These reach survival estimates are inconsistent with those reported by NMFS in their Table 4. Our direct estimates of survival from release to McNary dam are much lower than those reported by NMFS (*see Table below*). After adjusting for spill, our direct estimates of survival from release to MCN range from 28.9% to 45.4%, compared to NMFS estimates, which range from 43.1% to 81.1%.

Thus, there appears to be a wide range in the survivals estimated using these different methods which confirms the problem with NMFS's description of these estimates as 'observed' values and also highlights the sensitivity of survival estimates to assumptions about factors like FGE, spill, and flow.

The fact that survival estimates exceed 1 for 'Release to LGR'(S_1) in 2 out of the four years (S_1 = 1.080 in 1989, S_1 = 1.161 in 1990) suggests that assumptions about FGEs are unreasonable (note that all methods for 1989-92 rely heavily on FGE). If FGE's are off and result in overestimates in the first reach, survival will then be underestimated in the second reach. This problem is demonstrated by the large drop in survival from S_1 to S_2.

Table 1. Direct estimates of survival using NMFS numbers of released, detected, and returned to the river from Table 1 and their FGE's for LGR, LGO, & MCN.

1989-1992 PIT Tag Spring Chinook Survival Estimates				
	1989	**1990**	**1991**	**1992**
Release Dates				
Numbered Released	6163	1952	3262	1195
LGR Detected	2330	793	1349	426
% LGR Detected	37.8%	40.6%	41.4%	35.6%
# Detected &Returned LGR	110	100	249	0
# Detected @LGS	1070	253	575	194
# Detected &Returned LGS	173	174	198	125
# Detected @ McNary	451	174	198	125
% Detectetd @ McNary	7.3%	8.9%	6.1%	10.5%
FGE LGR	0.35	0.35	0.46	0.46
FGE LGS	0.58	0.58	0.58	0.58
FGE McNary	0.56	0.56	0.56	0.56
Trans				
LGR	2220	693	1100	426
LGS	897	79	377	69
LGR Pop	6657	2266	2933	926
LGS Pop	1845	436	991	334
McNary Pop	805	311	354	223
McNary Pop adjusted for spill	863	321	483	232
Release to LGR (S1)	**1.080**	**1.161**	**0.899**	**0.775**
LGR to LGS (S2)	0.416	0.277	0.541	0.669
LGS to McNary (S3)	0.850	0.870	0.575	0.841
LGS to McNary (S3) spill adjusted	0.911	0.898	0.786	0.876
Direct (FGE) Survival Rel to McNary	**38.2%**	**28.0%**	**28.0%**	**43.6%**
Direct Survival per Project	**82.5%**	**77.5%**	**77.5%**	**84.7%**
Direct (FGE) Survival Rel to McNary adjusted for spill	**40.9%**	**28.9%**	**38.2%**	**45.4%**
Direct Survival per Project Adj for spill	**83.6%**	**78.0%**	**82.5%**	**85.4%**
NMFS Table 4 Rel to McNary Survival Estimate	**47.6%**	**53.7%**	**43.1%**	**81.1%**
NMFS Table 4 Survival per Project	**86.2%**	**88.3%**	**84.5%**	**95.9%**

NMFS does not note in their analysis or comments that the authors, J.G. Williams (NMFS) and others, had previously proposed and abandoned an attempt to use these PIT tag data to estimate reach survival in 1994. Past reviews of the proposed methods and survival estimates using the 1989-1992 PIT tag data had pointed out many of these limitations. In a March 28, 1994 memo to M.H. Schiewe (NMFS), Williams states that "As you know, the PIT-tagged fish used in the [1989-1992] survival analysis were not originally released with a survival study in mind". Also, "[t]he analyses required that we use some best-guess survival estimates or ranges of estimates to Little Goose Dam." In addition to these limitations, more formal reach survival estimate methods had been resumed in 1993 (as a pilot) and 1994. Concern about limitations of this data set were also expressed by Fish Passage Center in an April 4, 1994 memo from M. DeHart to U. Varanasi (NMFS). The memo stated: "Since the data were collected as part of our Smolt Monitoring Program and since we have been cautioning regional groups against making assumptions and expanding the application of these data beyond their designated use, we are happy to note that the paper was not further pursued."

A major limitation of the data was that PIT tag releases from the Snake River trap were not made throughout the smolt migration season, and that trap catches were flow-dependent (i.e., the trap is inefficient or unfishable in low flows and high flows—E. Buettner, IDFG, pers. comm.). We have previously commented on the 1992 PIT tag data in the Preliminary Decision Analysis Document (Marmorek and Peters 1998; Appendix A, p. 42) in a response to J.G. Williams. Using the 1992 PIT tag data, Williams argued that survival in 1992 was greater than in 1973, that flow years were similar, and that, therefore, FLUSH was not a "good predictor of survivals that will occur under low-flow conditions".

> Response by Schaller et al. The comparison of 1973 and 1992 migration years is not as straight forward as implied. The statement that 1992 and 1973 flow years were similar is incorrect particularly for the peak yearling migration period. In 1992, 90% of the smolts were trapped before May 6 at the Lewiston Trap. This is the population that was PIT tagged. In 1992 these fish would have experienced flows of 54 kcfs (4/15-5/6 average), compared to only 38 kcfs in 1973. During this period in 1992 water temperatures were in the low 50F range. In contrast the 1973 survival was measured throughout the season and the temperatures in the Snake and Columbia rivers reached 70F in the later part of that season.

If we were willing to accept that the 1989-1992 PIT tag studies provided an index of survival, given the large estimation problems identified above, the 1992 Snake River trap catches had severe sampling problems that were ignored by Smith and Williams (1998; Attachment 4). It is very apparent that the Snake River trap catch did not randomly sample the spring/summer chinook smolt migration in 1992. Buettner and Brimmer (1993) summarize sampling problems at the Snake River trap in 1992:

> Annual chinook catch at the Snake River trap was the second lowest since the beginning of this project. The low trap catch was due to extremely poor trap efficiency associated with severe low flows (p. 1).
>
> There appears to be a threshold velocity required within the trap to collect chinook salmon

> effectively. Below this threshold velocity, which is about 1.6 to 1.8 feet per second, trap efficiency is very low and chinook salmon trap catch will not be representative of the chinook salmon population passing the trap. The threshold velocity is generally exceeded when discharge is above 30 to 33 kcfs (p. 9).

Based on Figure 3 in Buettner and Brimmer (1993), this threshold (for the Snake River above the Clearwater River confluence) was exceeded for only about 2 weeks at the Snake River trap starting in late April. Therefore, the 1992 data would tend to overestimate survival relative to the population at large, given these problems of inefficient sampling during low flow periods.

Smith and Williams also report that only PIT tag groups released before May 20 were used for survival estimates in 1992. There appears to be a pattern of decreasing detection percentages later in the 1992 season based on examination of Table 10 in (Buettner and Brimmer 1993). Of the small number (52) of PIT tagged smolts released after May 10, only 2 were subsequently detected at McNary Dam. After May 20, none of the 16 PIT tagged smolts were detected at McNary.

We conclude that further review of the 1989-1992 PIT tag data set is needed to determine how comparable the data are to passage model survival or detection probabilities. However, it seems that any such comparison will be assumption-dependent, and would not constitute a strong test with "out-of-sample-data" due to inherent data limitations.

1997 and 1998 Survival Estimates.

The 1997 and 1998 survival estimates to Bonneville Dam are inconsistent with methods agreed to in the PATH Hydrogroup. In the hydro group, we determined that the relevant reach for the 1996 study survvial was from LGR to MCN tailrace. This was because of the extremely limited number of detections lower in the river, and the consequently large uncertainty in the estimate for John Day reach. For the same reasons, NMFS recommended the 1997 estimate extend only as far as MCN. In an August 25, 1997 NMFS memo to William Stelle, M.H. Schiewe writes "Because of insufficient sampling at John Day Dam, it was not possible estimate survival to John Day Dam tailrace in 1997." We have similar concerns about the extension of the 1998 estimate to Bonneville.

Care has been taken to separate wild from hatchery chinook in other PIT tags studies because of the potential differences in survivorship related to differential experiences in early life stages. Applying coho detection probabilities as a surrogate to chinook detections ignores biological and ecological differences between the two species. In additon, this is no longer a CJS type estimate. Caution should be applied to using estimates of 1997 chinook survivorship estimates based on coho PIT tag studies.

However, assuming these NMFS methods for calculating survivorship are valid, the error surrounding the estimates preclude strong inference. In 1998 NMFS estimates survival between 32.5 and 84.1 (mean of 58.3 with 95% CI). We can infer through 1998 estimates, errors surrounding 1997 estimates are also high (errors are likely much higher given data constraints). Suggesting that CRiSP and FLUSH are slightly lower than observed estimates when the mean survivorship differs by 2-5% from these model estimates is not reasonable given the surrounding error of this estimate.

The main memo (item 4.8.1) and Attachment 6.

Williams et al. (1998) express concern that overestimation of survival rate in 1971 by FLUSH leads to a mis-estimation of D, potentially introducing significant error in the prospective modeling results. The 1971 survival estimate was not used in calibrating FLUSH. Predictions from that year are not used in the prospective life cycle modeling. No survival estimate of any kind from FLUSH that year (1971) has any other significant influence on the T/C vs. V_c relationship or D values used in prospective modeling.

Williams et al. (1998) also raise doubts about FLUSH control survival rate in other retrospective years. They state that FLUSH estimates appear "quite high relative to Raymond (1979) estimates." They use FLUSH results calibrated to TURB4 assumptions about pre-1981 turbine and bypass survivals. The STFA analytical team has repeatedly criticized and questioned the credibility of TURB4; more detailed critiques are provided in Appendix __. TURB4 was modeled with FLUSH not because the STFA analytical team believes it is likely to be correct, but because it is one of the several dam survival hypotheses that were supposed to be run with both passage models in this process. The assumption of TURB4 mortality during the 1970s has a large effect on the FLUSH reservoir survival relationship (PATH Decision Analysis Report, Figures A.2.1-8 and A.2.1-9). It also results in a poorer retrospective fit to the empirical reach survival data than TURB1 and TURB5 (WOE report Table 4-4), and a poorer fit to the S-R data with the Alpha model (WOE report, Table 4-2). Comparing FLUSH TURB4 survival estimates to the Raymond estimates is not a proper test of the STFA hypotheses.

References:

Anderson, J.J., et al. 1995. The CRiSP 1.5 mainstem passage model manual. University of Washington.

Anderson, J. J. 1998a. Critique on transport and extra mortality hypothesis. PATH document January 12, 1998 (Included in February 5, 1998 submission to SRP).

Anderson, J. J. 1998b. Weight of evidence,passage, transport, extra mortality, and aggregate hypothesis. July 27, 1998.

Buettner, E.W. and A.F. Brimmer. 1993. Smolt monitoring at the head of Lower Granite Reservoir and Lower Granite Dam. Annual Report 1992. Idaho Department of Fish and Game. Prepared for U.S. Department of Energy, Bonneville Power Administration. Project Number 83-323B.

Williams, J., S. Smith, T. Cooney, and C. Toole. 1998. Compilation of comments on Draft PATH weight of evidence report. Memo to D. Marmorek and C. Peters. July 27, 1998.

Whittaker, R.H. 1975. Communities and ecosystems. Second Edition. MacMillan, New York. 385 p.

Zabel, R. 1996a. A summary of the flow/travel time relationship in CRiSP. Draft document for PATH. March 5, 1996.

Zabel, R. 1996b. Discussion of CRiSP model parameters for nitrogen mortality. Draft document for PATH. March 5, 1996.

Zabel, R. 1996c. Sensitivity analysis (extracted from CRiSP manual). Draft document for PATH. March 5, 1996.

Zabel, R. 1996d. Snake River spring chinook survival 1966-1983. Draft document for PATH. February 4, 1996.