Plan for Analyzing and Testing Hypotheses (PATH)

Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook

Author List

David R. Marmorek and Calvin Peters, ESSA (editors)

Contributors:

James J. Anderson, CBR
Larry Barnthouse, McH
Lou Botsford, UC (Davis)
Tom Cooney, WDFW
Rick Deriso, IATTC
Jim Geiselman, BPA
Al Giorgi, BioAnalysts Inc.
Rich Hinrichsen, CBR
Josh Hayes, CBR
Mike Jones, MSU
Lyne Krasnow, NMFS
O.P. Langness, WDFW
Danny Lee, USFS
Chip McConnaha, NPPC
Charles M. Paulsen, PER
Randall Peterman, SFU
C.E. Petrosky, IDFG
Chris Pinney, CORPS
Miki Promislow, ESSA
Howard Schaller, ODFW
Steve Smith, NMFS
Chris Toole, NMFS
Earl Weber, CRITFC
John Williams, NMFS
Paul Wilson, CBFWA
Richard W. Zabel, UW

Draft for Scientific Review Panel

Plan for Analyzing and Testing Hypotheses (PATH)

Preliminary Decision Analysis Report on Snake River Spring/Summer Chinook

Prepared for

Implementation Team
and
PATH Scientific Review Panel

Compiled by

ESSA Technologies Ltd.
3rd Floor, 1765 West 8th Avenue
Vancouver, BC V6J 5C6

March 6, 1998

Citation: **Marmorek, D.R. and C.N. Peters (editors)**. 1998. Plan for Analyzing and Testing Hypotheses (PATH): Preliminary decision analysis report on Snake River spring/summer chinook. Draft report compiled and edited by ESSA Technologies Ltd., Vancouver, BC 92 pp. and appendices.

Executive Summary

Introduction

The Plan for Analyzing and Testing Hypotheses (PATH) is a formal and rigorous program of formulating and testing hypotheses. It is intended to identify, address and (to the maximum extent possible) resolve uncertainties in the fundamental biological issues surrounding recovery of endangered spring/summer chinook, fall chinook, and steelhead stocks in the Columbia River Basin. This process grew out of previous efforts by various power regulatory and fisheries agencies to compare and improve the models used to evaluate management options intended to enhance recovery of these stocks.

The objectives of PATH are to:

1. determine the overall level of support for key alternative hypotheses from existing information, and propose other hypotheses and/or model improvements that are more consistent with these data;
2. assess the ability to distinguish among competing hypotheses from future information, and advise institutions on research, monitoring and adaptive management experiments that would maximize learning; and
3. advise regulatory agencies on management actions to restore endangered salmon stocks to self-sustaining levels of abundance.

This report describes the methods and results of the decision analysis framework we have used to address the third objective for Snake River spring/summer chinook salmon. The specific purposes of this preliminary decision analysis report are to: 1) test the methods of decision analysis we have formulated over the last two years; 2) provide decision makers with our **preliminary** insights into the range of potential responses of Snake River spring/summer chinook to alternative management decisions; and 3) characterize the magnitude of various uncertainties and demonstrate their relative importance in affecting the outcomes of alternative management decisions.

The preliminary decision analysis builds on the “retrospective” analyses completed to date by PATH under our first objective. PATH retrospective analyses attempt to identify the major spatial and temporal patterns in abundance, productivity, and survival of these stocks over the last 30 to 40 years and to determine the relative contribution of Habitat, Harvest, Hatchery, Hydro, and Climatic influences to these patterns. Results of these analyses were published in the peer-reviewed PATH FY96 Retrospective Analysis report, and summarized in “Conclusions of FY96 Retrospective Analyses”, a consensus document written by PATH scientists in December, 1996. Other retrospective analyses were completed in FY97, and will soon be published. All of the retrospective analyses completed to date are considered in this report.

PATH retrospective analyses have helped to bring a substantial set of empirical information to bear on alternative hypotheses to explain recent declines and have led to considerable improvements in both our understanding and modeling approaches. In addition, there has been considerable convergence on the historical data sets to use in calibrating and testing models, and on many of the assumptions to be made when projecting future population changes.

The PATH retrospective analyses have also highlighted some major uncertainties in past and current conditions that have yet to be resolved because of incomplete data and differences in interpretation. These uncertainties, along with uncertainties in projecting future conditions, imply that a single management action can have a number of possible outcomes, depending on what is assumed about past, present, and future conditions. This range of possible future outcomes of management actions is best captured by modeling salmon populations under a set of alternative hypotheses about uncertain components of the system.

The preliminary decision analysis described in this report looks systematically at the outcomes of management actions under several alternative hypotheses about biological mechanisms that link actions to possible outcomes. This approach was recommended by the SRP and by independent scientists within PATH as a tool for explicitly considering uncertainties in the decision-making process, in recognition that decisions cannot wait for all uncertainties to be resolved. A variety of management objectives can be used to evaluate alternative actions.

We anticipate that review of this preliminary analysis will lead to refinements in methods and consideration of additional alternative hypotheses. These improvements will be incorporated in the final decision analysis report for spring/summer chinook. The final report will also present analyses of additional management actions to those evaluated in this report (including drawdown of John Day Dam), and will endeavor to reach consensus to the maximum extent possible on the relative weights assigned to alternative hypotheses based on the strength of supporting evidence and our professional judgements. We anticipate, however, that lack of evidence will constrain our ability to reach consensus on the relative likelihood of some alternative hypotheses.

For the next four months, PATH intends to focus on completing analyses for fall chinook before returning to spring/summer chinook. We are distributing this preliminary report now rather than wait until the above refinements are made to show what we have been doing and where we are headed. The final report for spring/summer chinook (which will incorporate the above revisions) will be completed by the fall of 1998.

Decision Options

Although many agencies have drafted some very broad goals to help direct decision making, this decision analysis is focused on a narrower question: **To what extent can alternative hydrosystem actions prevent extinction and lead to recovery of stocks either listed or proposed for listing, including wild spring/summer chinook, fall chinook and steelhead stocks in the Snake River and mid-Columbia region?** This preliminary decision analysis considers three alternative hydrosystem actions: A1 (current operations), A2 (maximize transportation without surface collectors), and A3 (drawdown to natural river level of the four Lower Snake dams). We chose this restricted set of options to allow a thorough evaluation of our biological decision analysis and modeling tools by both PATH scientists and decision-makers. We believe that the next options to be evaluated should be B1 (natural river drawdown of both the four Lower Snake dams and John Day Dam), maximizing transportation with surface bypass collectors (A2'), and the in-river option (A6), so as to bracket the potential range of responses of fish populations.

While PATH is only looking at hydrosystem decisions explicitly, the effects of habitat and harvest management actions are being considered in sensitivity analyses. We are also developing approaches to including uncertainties with respect to management of hatcheries, to be added to our final report. The approaches used for all four H's (hydro, habitat, harvest, hatcheries) will be re-examined following peer review of this report.

We also intend to explore options for an experimental management approach, which varies management actions over time and space in a deliberate attempt to test key hypotheses. An experimental management approach has been recommended by some members of the PATH Scientific Review Panel because some of the major uncertainties are difficult to resolve with current information. Though experimentation may pose risks to these stocks, there is risk inherent in any actions, including continuing present operations, as these populations are at dangerously low levels.

How to Assess the Outcomes of the Options

Outcomes of the alternative actions will depend on what is assumed about past, present and future conditions experienced by fish in response to management actions. The previous PATH retrospective analyses have elucidated a great deal, and have also pointed out uncertainties in past conditions due to incomplete data and potentially confounding influences. These uncertainties generate a range of alternative assumptions about historical conditions, which are used in retrospective modeling analyses that generate quantitative estimates of parameters needed to run models into the future. Results from the retrospective analysis are passed to the prospective modeling analysis, which quantifies the range of possible futures. This set of possible futures depends not only on the understanding and parameter estimates gleaned from the retrospective analysis, but also on assumptions about future conditions (such as climate) and the response of stocks to new management actions (such as Snake River drawdown).

The outcomes of alternative hydro management actions are evaluated in terms of various performance measures. These measures are used to rank alternative actions according to how well they meet specified management goals. A variety of performance measures have been developed to assess the biological implications of different management actions. Because our primary goals are to determine the hydrosystem actions that should be taken to prevent extinction and lead to recovery of endangered stocks, we focus here on the National Marine Fisheries Service (NMFS) jeopardy standards that account for each of these goals. These standards are a measure of the ability of actions to increase the spawning abundance of stocks to levels associated with long-term persistence and stability. Survival standards are based on projected probabilities that the spawning abundance will exceed a pre-defined “survival” threshold over a 24 or 100 year simulation period; survival standards are met when that probability is 0.7 or greater. Recovery standards are based on probabilities of exceeding a “recovery” threshold in the last eight years of a 48-year simulation period; this standard is met when the probability is 0.5 or greater.

The standards are applied to the sixth best stock out of the seven Snake River “index” stocks of spring/summer chinook (Imnaha, Minam, Bear Valley/Elk, Sulphur Creek, Marsh Creek, Johnson Creek, and Poverty Flats) to ensure that most of the stocks are able to meet the survival and recovery goals. These seven index stocks are the only ones for which sufficient historical data exist to develop spawner-recruit relationships, required for generating projections of future stock sizes. Further work is required to generalize results from these stocks to all wild chinook populations of the Snake River basin.

Uncertainties in the Response of Populations to Management Actions

There are many uncertainties that can potentially affect the responses of fish populations to management actions. We have focused on twelve of the most important of these uncertainties, and have laid out a range of alternative hypotheses for each. The uncertainties are of two types: uncertainty regarding the future environment, and uncertainty regarding how the system works (i.e., the survival changes caused by management actions). Although the future environment may be beyond human control (e.g., future climate), the uncertainty inherent in projecting it is of potential significance in determining future population sizes. Alternative hypotheses to describe how the system works often hinge on the interpretation of historical information, because the functional relationships in models are based on both general principles and historical data. However, as past information is incomplete, there are differing interpretations of the relative importance of different factors in causing recent declines of Snake River spring-summer chinook.

The twelve uncertainties considered in the preliminary decision analysis were:

1. *Passage assumptions* – uncertainty in direct survival of in-river fish, and the partitioning of in-river survival between dam and reservoir survival.
2. *Fish guidance efficiency (FGE)* – uncertainty in the effectiveness of extended-length screens in diverting fish away from the turbines, relative to standard-length screens.
3. *Turbine/Bypass Mortality* – uncertainty in historical estimates of bypass and turbine mortality for some projects prior to 1980.
4. *Predator Removal Effectiveness* – uncertainty in the effect of the predator removal program (i.e., removal of squawfish for bounties) on survival of salmon smolts in reservoirs.
5. *Transportation assumptions* – uncertainty in the relative survival of transported and non-transported fish after the fish have exited the migration corridor (i.e., below Bonneville Dam).
6. *Stock productivity* – uncertainty in the extent to which Snake River and lower Columbia stocks share common mortality effects.
7. *Extra mortality* – uncertainty in the mortality of both transported and non-transported fish occuring beyond Bonneville Dam.
8. *Future climate* – uncertainty in future patterns in climatic conditions.
9. *Habitat effects* – uncertainty in the biological effects of future habitat management actions.

We also considered the following three uncertainties when projecting the effects of drawdown to natural river of the four lower Snake River dams (option A3):

10. *Length of Pre-Removal Period* – the duration of time between a decision to proceed with drawdown and actual removal of dams (pre-removal period) due to uncertainty in the Congressional appropriations process and the possibility of litigation.
11. *Length of Transition Period* – duration of period between completion of dam removal and establishment of equilibrium conditions in the drawndown section of the river (transition period), reflecting uncertainty in the physical and biological responses to drawdown (e.g., short-term response of predators, release of sediment).
12. *Juvenile survival rate once river has reached equilibrium conditions after drawdown* – uncertainty in the long-term physical and ecological effects of drawdown (e.g., change in density of predators).

We call a particular combination of hypotheses for these twelve uncertainties *a prospective aggregate hypothesis*. Each prospective aggregate hypothesis potentially yields a unique biological response to an action. We have explored 5,148 different aggregate hypotheses in this preliminary analysis. One of our objectives was to determine which uncertainties have limited effects on performance measures and the resulting decision, so that we can focus on the most critical alternative hypotheses. In the final report, we may also develop new variations or combinations of hypotheses that better reflect recent evidence.

We also consider alternative harvest schedules to assess the sensitivity of responses to hydro actions to variations in harvest rate. A number of potentially important factors were not explicitly quantified in the models, although some are considered implicitly in the models to some extent. These include several factors discussed by the Independent Scientific Group in the "Return to the River" report, such as the effects of genetic interactions between populations, and impacts of the hydropower system on conditions in the estuary.

Results

There were five objectives for the results of the preliminary analyses:

1. Explore ways to summarize complex analyses and results into graphs that are easy to understand, interpret, and explain to decision-makers.
2. Provide **preliminary** insights into the relative performance of alternative actions.
3. Identify key uncertainties that affect the results.
4. Test the sensitivity of decisions to the weights placed on key uncertainties, so as to focus the assessment of existing evidence, and the acquisition of additional evidence.
5. Summarize results for some other important performance measures.

Ways to Summarize Results

We have generated predicted outcomes for alternative management actions using each possible aggregate hypothesis. Since there are 5,148 unique aggregate hypotheses, there are 5,148 unique alternative futures that one could examine to evaluate alternative actions. We used two alternative ways to summarize these outcomes. The first was to show a frequency distribution of all outcomes for a single action. This shows the range of possible futures associated with the uncertainties in past, present, and future conditions (an example for Action A1 is shown in Figure E-1). We separated results generated with the CRiSP-T3 passage model and transportation assumptions from those generated with the FLUSH-T1/T2 model because these two models represent fundamentally different approaches to estimating mortality through the juvenile migration corridor and because they are each associated with different assumptions about the relative survival of transported and non-transported fish in the ocean (i.e. CRiSP is associated with transportation assumption T3, while FLUSH is associated with T1 and T2).

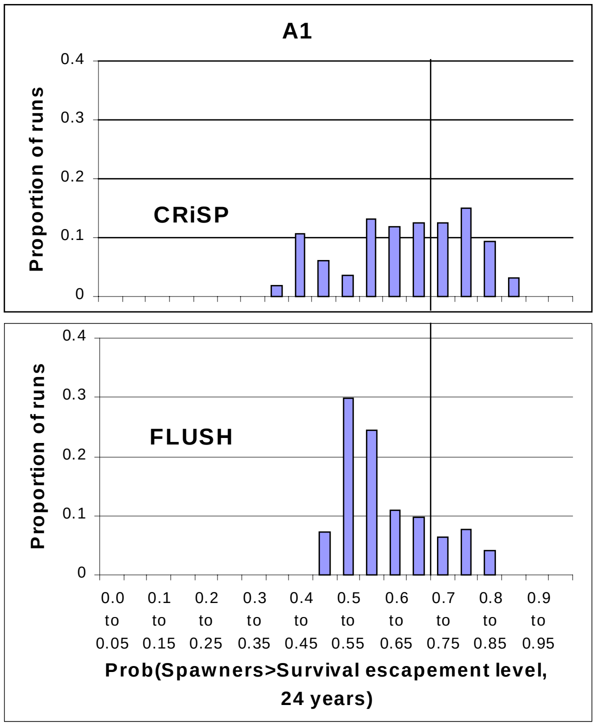


Figure E-1: Frequency distribution of possible future outcomes of Action A1 using CRiSP-T3 (top) and FLUSH-T1/T2 (bottom) passage models and transportation assumptions. Outcomes are measured as the probability of the spawning abundance of the sixth best stock exceeding the survival level of escapement in the first 24 years of the 100-year simulation period. The height of the bars reflects the relative frequency with which a particular outcome is projected. The vertical line at 0.7 represents the NMFS survival standard; outcomes to the right of that line are considered to have met the 24-year survival standard.

The second approach was to calculate the “expected ability” of an action to meet the NMFS survival and recovery goals. This is essentially the weighted fraction of the 5,148 outcomes that met the NMFS criteria for survival and recovery, where the weights reflect the relative degree of belief in one hypothesis over another. In the preliminary analysis, all hypotheses were given equal weights. An example of this type of output is shown in Figure E-2.

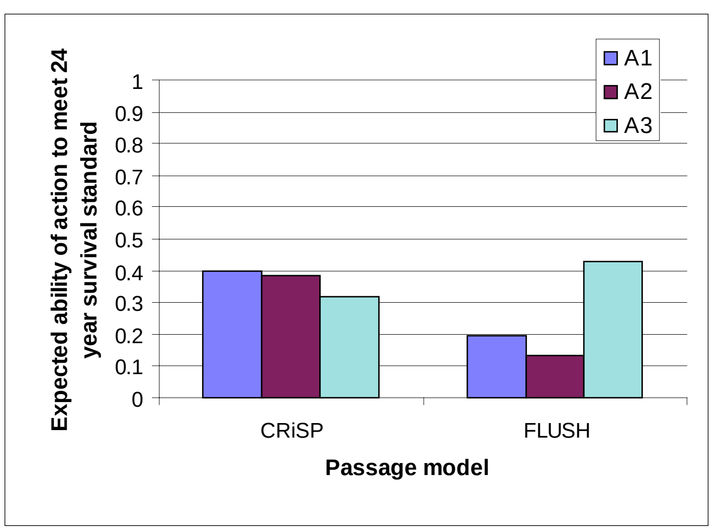


Figure E-2: Expected ability of A1, A2, and A3 to meet the 24-year survival standard. The standard is met when the spawning abundance of the sixth best index stock exceeds the survival escapement level an average of 70% of the time over the first 24 years in the 100-year simulation period.

Relative Performance of Alternative Actions (Preliminary results)

1. ***There is a large variation in outputs, even within models and actions.***

   There is considerable uncertainty in the outcomes of alternative management actions. Probabilities of spawning abundances exceeding survival and recovery escapement levels can range anywhere from very low to very high values, depending on the underlying aggregate hypothesis. For example, probabilities of being above the recovery escapement level generated with the FLUSH-T1/T2 passage model range from 0.15 to 0.85 under A1, while CRiSP-T3 probabilities range from 0.05 to 0.9. For both models, there is greater variation in probabilities associated with recovery escapement levels than in probabilities of exceeding survival escapement levels. CRiSP-T3 results generally have a greater range than FLUSH-T1/T2 runs, particularly under A3. Since this introduces considerable uncertainty into which decision should be made, it is important to identify the individual components of an aggregate hypothesis that have the greatest effect on decisions.

2. ***Relative performance of the management options depends on passage model assumptions.***

   Using CRiSP-T3 passage model and transportation assumptions, A1 or A2 had very similar expected abilities to meet the NMFS standards, while A3 always is the lowest. With FLUSH-T1/T2 passage model and transportation assumptions, A3 always has the highest expected ability, followed by A1 and then A2. Drawdown (A3) represents both improved in-river survival and a reduction in transprotation. Under the CRiSP transportation assumptions, A3 causes a net decline in survival relative to A2; with FLUSH the reverse occurs. This confirms our general expectations based on the structure and application of these models, but our result explicitly quantifies the differences. Such a quantification is extremely important.

3. ***Long-term standards are easier to meet than short-term standards.***

   With CRiSP-T3, the expected ability of action A2 was 0.38 for the 24-year survival standard, but was 0.65 for the 100-year survival standard, and 0.5 for the 48-year recovery standard. The expected ability of action A3 with FLUSH-T1/T2 was 0.42 for the 24-year survival standard, and 1.0 for both the 100-year survival and the 48-year recovery standards.

4. ***With this set of actions, there are few instances in which <u>all</u> of the survival and recovery standards are met with a high expected ability.***

   The highest expected ability to meet all survival and recovery standards using the CRiSP-T3 model is around 0.35 (obtained with action A2), and around 0.4 using FLUSH-T1/T2 (obtained with action A3). We would assume that decision-makers would want the expected ability to meet all of the recovery and survival standards to be high, since that implies a high degree of certainty that these standards will be met. These preliminary results suggest that significantly greater improvements in survival are required beyond those provided by the management actions analyzed here, since none of the current set of actions are able to meet all of the standards with any degree of certainty (at least when the aggregate hypotheses are weighted equally).

5. Alternative standards and harvest schedules affect the outcomes of management options, but not their relative ranking.

The ranking of actions was not affected when we applied weaker (i.e. easier to meet) and stronger (more difficult to meet) jeopardy standards than the informal NMFS definition (0.70 probability of exceeding survival escapement levels, 0.50 probability of exceeding recovery escapement levels), although the expected ability of actions was predictably lower for the stronger standard and higher for the weaker standard. The ranking of actions was also unaffected when we use two more conservative harvest rate schedules than the one based on current management. In one of these alternative schedules, harvest rates are reduced by one-third from their current values. This change had little or no effect on the expected ability of actions to meet survival and recovery standards. In the other alternative schedule, harvest rates of spring-summer chinook are set to 0. Here, the effects were greater than when harvest rates were reduced by one-third; the magnitude of these effects on outcomes depended on the action and passage model assumptions.

Sensitivity of Outcomes and Decisions to Effects of Uncertainties

To assess the sensitivity of outcomes to other uncertainties, we defined two possible criteria for decision-making, both based on the NMFS Jeopardy Standards. The first is a **relative criterion**, in which the preferred action is the one that simply maximizes the expected ability to meet all three NMFS survival and recovery standards. Because the transportation vs. drawdown question seems to be of most interest in the region, we are concerned primarily with the relative ranking of A2 and A3 in this sensitivity analysis. The second possible basis for decision-making is based on an **absolute criterion**. We assume that some minimum expected ability to meet survival and recovery standards is required for an action to be considered acceptable. Since it is not clear at the moment what the minimum expected ability should be, we use 0.7 for illustrative purposes. These criteria are admittedly difficult to meet, since they include the 24-year survival standard (see conclusion #3 above).

6. Within each model, very few uncertainties have significant effects on outcomes and decisions.

The only uncertainty that significantly affects the decision is the uncertainty about the source of extra mortality. Under the "BKD" hypothesis and the regime shift hypothesis, all of the actions fall short of our assumed criterion of 0.7 expected ability to meet all standards. In other words, if post-Bonneville extra mortality remains regardless of hydrosystem actions, the stocks will have a poor ability to recover. However, if extra mortality is related to the hydrosystem, both A1 and A2 (under CRiSP-T3) and A3 (under FLUSH-T1/T2) exceed this minimum level.

In terms of future analyses and monitoring, overall results suggest that many of the uncertainties could be ignored, since they appear to have relatively modest affects on the model results. The main uncertainties to resolve are those associated with passage model assumptions (i.e., estimates of direct in-river survival, and relative survival of transported and non-transported fish), and extra mortality. In some cases, experimental management actions may present the only opportunity for resolving these uncertainties. We plan to have a workshop in 1998 to explore the feasability, benefits, and risks of such experiments.

7. The ranking of actions is relatively insensitive to "best" and "worst" case combinations of hypotheses.

We looked at the effects of "best-case" and "worst-case" combinations of passage-related hypotheses, drawdown-related hypotheses, and stock productivity and future climate hypotheses. Best and worst case sets of have predictably large effects on results, but they do not significantly affect the relative ranking of actions. A1 or A2 is still always the best with CRiSP-T3, and A3 is still always the best under FLUSH-T1/T2. *In terms of meeting an absolute criterion, the expected ability of actions to meet the survival and recovery standard is below 0.7 for all cases, except for the best-case drawdown scenario.* Under this combination of hypotheses, the expected ability to meet survival and recovery standards for A3 increases to around 0.8 under FLUSH-T1/T2.

8. Results are similar using a single stock (Marsh Creek).

Results for a single stock (Marsh Creek), using a different standard (0.75 probability of exceeding survival escapement levels over 24 years) show the same general patterns. CRiSP-T3 assumptions tend to favor A2, while FLUSH-T1/T2 assumptions favor A3. Passage model assumptions and extra mortality hypotheses were important in determining which actions met a 24-year survival standard. FGE, PREM, and prospective model alternatives were also important in the results for the single stock.

Sensitivity of Outcomes and Decisions to Weightings on Alternative Hypotheses

PATH will attempt to assign weights to those key uncertainties based on the weight of evidence for and against particular hypotheses. The first step in assigning these weights is to establish just how sensitive the decision is to the weightings that are placed on alternative hypotheses. For example, what is the critical weighting that must be placed on the hydro-related hypothesis for extra mortality before the 0.7 threshold is reached? This information can help to frame the assignment of weights by identifying what the critical weights are. Precise framing of this discussion will be particularly important where there is disagreement among PATH scientists and agencies over what these relative weights should be.

Our results indicate that passage model, extra mortality, and best/worst combinations of drawdown hypotheses had the greatest effects on decisions. Unfortunately, these uncertainties will also likely be the most difficult to assign weightings to, because of firmly-held beliefs about the interpretation of historical data and because extra (post-Bonneville) mortality and drawdown effects are the most difficult to measure. Therefore, we looked at the effects of different weightings on these hypotheses on the expected ability of actions to meet all three survival and recovery standards.

9. There has to be a great deal of certainty about passage model assumptions and the hydro extra mortality hypothesis before any of the actions achieve an expected ability to meet all survival and recovery standards greater than 0.7.

For A1 and A2, this criterion is only met if we are absolutely certain (i.e. weight=1.0) that CRiSP-T3 and the hydro extra mortality hypothesis are correct. For A3, the 0.70 criterion is met when FLUSH-T1/T2 is assigned a weight of at least 0.8 and the hydro hypothesis is assigned a weighting of 1.0, or when FLUSH-T1/T2 is assigned a weighting of 1.0 and the hydro hypothesis is assigned a weighting of 0.8.

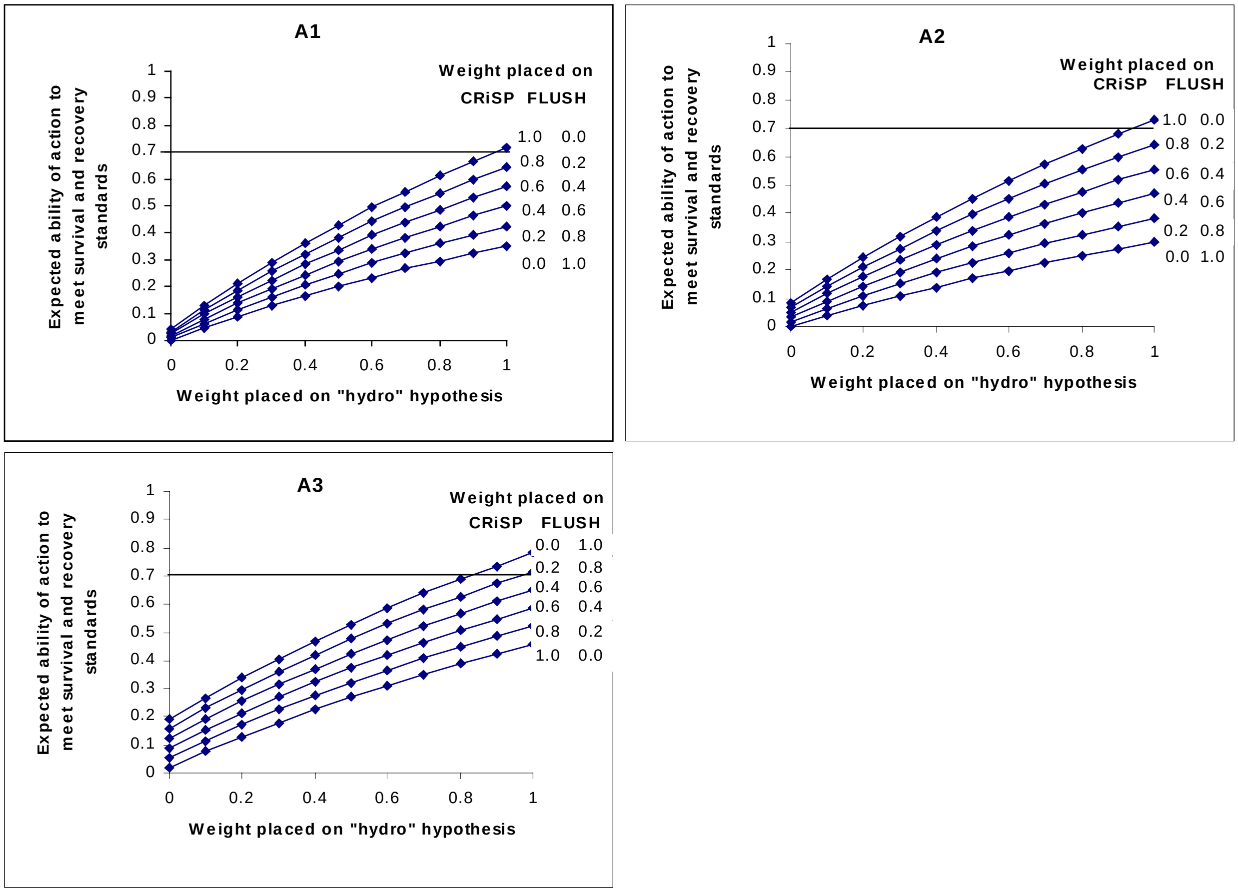


Figure E-3: Sensitivity of expected ability to meet survival and recovery standards to relative weights placed on the "hydro" extra mortality hypothesis and the passage models. Note that the remaining weight placed on the extra mortality hypotheses (i.e., 1 – weight placed on hydro hypothesis) is divided evenly between the "BKD" and the regime shift hypotheses. For example, when the weight placed on the hydro hypothesis is 0.8, the weights placed on the "BKD" and the regime shift hypotheses are both 0.1.

10. There has to be virtual certainty that the pre-removal and transition periods for drawdown to achieve an expected ability to meet survival and recovery standards greater than 0.7, and this only occurs for FLUSH-T1/T2.

A3 under FLUSH-T1/T2 only met the 0.7 criterion when it was certain that the pre-removal period was three years, and the transition period two years. If the weights on these hypotheses are high, then the assumptions about equilibrated juvenile survival rate did not matter.

11. Key uncertainties are unlikely to be resolved with existing data.

There will have to considerable agreement on three key uncertainties (in-river survival, transportation assumptions, and extra mortality) before one of the actions is clearly able to meet the survival and recovery standards. Given the lack of data that gave rise to the uncertainties, and the strongly-held beliefs which fill in dats gaps, this consensus is not likely to be achievable without a well-planned experimental design that is specifically directed towards answering questions about extra mortality and passage model assumptions.

Other Performance Measures

The NMFS jeopardy standards are only one of a number of different measures of performance produced by PATH modeling analyses. In this report, we also briefly report results for two additional measures: projected harvest rates, and Smolt-to-Adult survival rates from the time they pass the upper-most dam as smolts to the time they return to that dam as adults.

12. Projected harvest rates are highly variable.

We showed an example of the trends in mainstem harvest rates for a single stock (Imnaha), and a single action (A1) over time. We showed this for an optimistic aggregate hypothesis and a pessimistic hypothesis. In most years, harvest rates can range from below 0.1 to above 0.35 for a particular scenario. Such uncertainty is important to communicate to decision-makers and to others who will be using this information, such as the economic workgroup.

13. Median SARs of between 2 and 7% are associated with meeting the 100-year survival standard.

This is consistent with the interim SAR goal of between 2 and 6% identified by the PATH hydro workgroup (Ch. 6 in PATH FY1996 Retrospective Report). Note that these 'median SARs' are computed over a 100-year period.

In addition to quantitative performance measures, we would also like to look at how well the alternative management actions do in terms of qualitative measures of performance such as the concepts discussed in the ISG's "Return to the River" report. Such qualitative measures can allow us to incorporate less quantitative but nonetheless important issues relating to the relative health of individual salmon populations, aquatic communities, and entire ecosystems.

Again, we caution that these results are preliminary. We anticipate that review of this preliminary analysis will lead to refinements in methods and consideration of additional alternative hypotheses. These improvements will be incorporated in the final decision analysis report for spring/summer chinook. The final report will also present analyses of additional management actions to those evaluated in this report (including drawdown of John Day Dam), and will endeavor to reach consensus to the maximum extent possible on the relative weights assigned to alternative hypotheses based on the strength of supporting evidence and our professional judgements. We anticipate, however, that lack of evidence will constrain our ability to reach consensus on the relative likelihood of some alternative hypotheses.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Introduction

The Plan for Analyzing and Testing Hypotheses (PATH) is a formal and rigorous program of formulating and testing hypotheses. It is intended to identify, address and (to the maximum extent possible) resolve uncertainties in the fundamental biological issues surrounding recovery of endangered spring/summer chinook, fall chinook, and steelhead stocks in the Columbia River Basin. This iterative process grew out of previous efforts by various power regulatory agencies and state, federal, and tribal fisheries agencies to compare and improve the models used to evaluate management options intended to enhance recovery of these stocks.

The objectives of PATH are to:

1. Determine the overall level of support for key alternative hypotheses, and propose other hypotheses and/or model improvements that are more consistent with existing data;
2. Assess the ability to distinguish among competing hypotheses from future information, and advise institutions on research, monitoring and adaptive management experiments that would maximize learning; and
3. Advise regulatory agencies on management actions to restore endangered salmon stocks to self-sustaining levels of abundance.

PATH has done considerable work on Objective 1. This report focuses on the third objective. We intend to return to Objective 2 in FY98.

PATH analyses thus far have focussed on Snake River spring/summer chinook salmon, and have progressed in two stages: “Retrospective” analyses and “Prospective” analyses. The retrospective analyses attempt to identify the major spatial and temporal patterns in abundance, productivity, and survival of these stocks over the last 40 years and to determine the relative contribution of Habitat, Harvest, Hatchery, Hydro, and Climatic influences to these patterns. Our first major set of analyses was published in the PATH FY96 Retrospective Analysis report, and has received generally favorable reviews by an independent Scientific Review Panel (SRP). Much of this work will be published in peer-reviewed scientific journals in the near future. PATH also summarized the FY96 analyses in “Conclusions of FY96 Retrospective Analyses”, a document that represented the consensus of PATH scientists as of December, 1996. The Conclusions Document assessed the strength of evidence for each conclusion, considered alternative interpretations of historical information, and additional information needs required to strengthen those conclusions. This document was also favorably reviewed by the PATH SRP. Several additional retrospective analyses were completed in FY97 and a report documenting these analyses is planned in the near future. All of the retrospective analyses completed to date are considered in this report.

The PATH retrospective analyses have helped to bring a substantial set of empirical information to bear on alternative hypotheses to explain recent declines (e.g., stock-recruitment information, in-river survival studies, transportation experiments, smolt-to-adult return rates [SARs]) and have led to considerable improvements in both our understanding and modeling approaches. In addition, there has been considerable convergence on the historical data sets to use in calibrating and testing models, and on many of the assumptions to be made when projecting future population changes. Because the future is uncertain, and we are uncertain about several functional relationships in the system, the range of possible futures is best captured through a set of alternative hypotheses about different components of the system.

A given set of alternative hypotheses about all components of the system (stock productivity, downstream migration, marine survival, transportation, future climate, etc.) is referred to as an "aggregate hypothesis". Historical information can be used to assess the likelihood of alternative hypotheses about one or more system components. For example, the 1996 retrospective analyses concluded that escapement, productivity and survival have been poorer for upriver stocks than for downriver stocks, since 1975. Different aggregate hypotheses address the question of why the performance of upriver stocks is poorer than downriver stocks, and attribute different degrees of influence to the various factors that can potentially explain this pattern. These factors include direct passage mortality within the hydrosystem, delayed passage mortality (after Bonneville Dam) that is related to the hydropower system, and extra mortality that is independent of the hydropower system (due to changing climate, or the effects of hatcheries). These alternatives add uncertainty to decisions about the hydrosystem. A major objective of this report is to understand the effects of these uncertainties on the future condition of populations under different management actions. Aggregate hypotheses are discussed in more detail in Chapter 4.

PATH "prospective" analyses attempt to evaluate the ability of alternative management actions to restore depressed populations of spring/summer chinook stocks. These prospective analyses are based on results of the retrospective analyses, and use this information in two different but complementary ways to assess alternative management actions. The first approach is a weight-of-evidence approach, which synthesizes existing information around specific questions in the decision-making process. This approach was applied in Chapter 6 of the FY96 Retrospective Report, which developed a flowchart of key questions to consider when making decisions about the hydropower system and compiled available evidence to answer those questions.

The second approach to prospective analyses is a formal decision analysis which systematically looks at the outcomes of management actions under several alternative hypotheses about biological mechanisms that link actions to outcomes. Based on their outcomes, actions are then ranked according to specified management objectives. This approach was recommended by the SRP and by independent scientists within PATH as a tool for explicitly considering uncertainties in the decision-making process, in recognition that decisions cannot wait for all uncertainties to be resolved. The biological rationale for alternative hypotheses uses much the same information as the "weight-of-evidence" approach completed in FY96. A variety of management objectives can be used to evaluate alternative actions. This report focuses mostly on the survival and recovery standards used by the National Marine Fisheries Service (NMFS), which are defined precisely in Section 3 and Appendix D. In general terms, the survival standard is the fraction of time during many simulations that the spawning abundance of a stock is above a certain specified low threshold. This fraction is calculated for both 24 and 100 years (about 6 and 25 salmon generations, respectively). The recovery standard is the probability that the spawning abundance exceeds a specified recovery level during the last 8 years of a 48-year period.

Over the past two years, PATH has developed a decision analysis framework and completed a preliminary decision analysis for Snake River spring/summer chinook stocks. The purpose of this preliminary decision analysis report on spring/summer chinook is to:

1. test the methods of decision analysis we have formulated over the last two years both to provide technical guidance to future PATH analyses to acquaint regulatory agencies with these methods, and to obtain feedback on the utility of our approaches;
2. provide decision makers with our preliminary insights into the range of potential responses of Snake River spring/summer chinook to alternative management decisions (while cautioning that these insights may change significantly in future reports); and

3. characterize the magnitude of various uncertainties and demonstrate their relative importance in affecting the outcomes of alternative management decisions, both to aid decision makers in understanding the complexity of decision making, and to focus further efforts of PATH participants and other scientists/managers on those critical uncertainties.

We stress that this is a work in progress. We anticipate that review of this preliminary analysis will lead to refinements in methods and consideration of additional alternative hypotheses. These improvements will be incorporated in the final decision analysis report for spring/summer chinook. The final report will also present analyses of additional management actions to those evaluated in this report (including drawdown of John Day Dam). To the extent possible, we will endeavor to reach consensus on the relative weights assigned to alternative hypotheses based on the strength of supporting evidence and our professional judgements. We anticipate, however, that lack of evidence will constrain our ability to reach consensus on the relative likelihood of some alternative hypotheses.

For the next four months, PATH intends to focus on completing analyses for fall chinook before returning to spring/summer chinook. Therefore, we distributing this preliminary report now rather than wait until the above refinements are made to show what we have been doing and where we are headed. The final report for spring/summer chinook (which will incorporate the above revisions) will be completed by the fall of 1998.

It is important to recognize the different levels of decision making which exist in the Columbia hydropower system. Many agencies have drafted some very broad goals to help direct decision making, for example, the draft Multi-Year Implementation Plan being developed by the Columbia Basin Fish and Wildlife Authority identifies the following goal

> "Restore sustainable, naturally producing fish and wildlife populations to support tribal and non-tribal harvest and cultural and economic practices. This goal will be achieved by restoring the biological integrity and the genetic diversity of the Columbia River ecosystem and through other measures that are compatible with naturally producing fish and wildlife populations."

While the PATH group considers these goals to be of importance, this decision analysis is focused on a narrower question: **To what extent can alternative hydrosystem actions prevent extinction and lead to recovery of stocks either listed or proposed for listing, including wild spring/summer chinook, fall chinook and steelhead stocks in the Snake River and mid-Columbia region?**

In addressing this question, the PATH group adopted several principles:

1. recognize that there are alternative hypotheses regarding the causes of historical population changes, the responses of fish populations to future management actions, and the range of climatic conditions fish may encounter in the future;
2. agree to the greatest extent possible on the set of empirical studies to be used for evaluating alternative hypotheses;
3. develop an analytical framework which can easily and clearly implement alternative hypotheses about different components of the system, as well as aggregate hypotheses that combine component hypotheses;

4. demonstrate the implications and relative importance of alternative hypotheses for future decisions; and
5. document the biological rationale for alternative hypotheses, based to the greatest degree possible on empirical evidence, utilizing previous PATH work and other studies.

Through these principles, the PATH process will ensure that the region has the benefit of the best available scientific methods and information in the analyses supporting efforts to recover and rebuild endangered fish stocks. The focus of PATH analyses will be on spring / summer chinook, fall chinook, and steelhead.

Conclusions on sockeye will be based on very general inferences from the spring / summer chinook analyses, but will not consider the sockeye captive brood stock program and supplementation issues, due to the limitations of both time and information.

1.2 Structure of this Report

The main part of this report is intended to be read by decision makers and their technical advisors. The appendices provide more detailed information on the methods used and the rationale for alternative hypotheses. The structure of Appendices A and B parallels the structure of Sections 4 and 5 of this report. In particular, each of the system components described in Section 4, and the alternative hypotheses associated with these components, are elaborated upon in Appendix A. Appendix B provides more detailed results of our analyses, while Section 5 provides an overview of the main results. For a general overview of the report, we recommend reading Sections 1-3, 4.1, the introductions to Sections 4.2 and 4.3 and Section 5.

Thus, a considerable amount of uncertainty will remain, and we will rigorously assess the implications of this uncertainty.

2 Decision Options

The set of actions currently under consideration for managing the hydrosystem is shown in Table 2-1. This set of actions has been developed by the Implementation Team (I.T.), and draws from previous experience of analyzing a much larger set of options (refs: Biological Opinion; System Operating Review; System Configuration Study). Appendix C describes the hydrosystem operating requirements associated with each option. This preliminary decision analysis only considers three of the options in Table 2-1: A1 (current operations), A2 (maximize transportation without surface collectors) and A3 (drawdown to natural river level of the four Lower Snake dams). We chose this restricted set of options so as to allow us to proceed with a reasonably thorough test of our biological decision analysis and modeling tools, without having to wait for further work by the hydrologic modelers who simulate the flows expected under different scenarios. We believe that the next options to be evaluated should be B1 (natural river drawdown of both the four Lower Snake dams and John Day Dam), maximizing transportation with surface bypass collectors (A2'), and the in-river option (A6), so as to bracket the potential range of responses of fish populations. Later in this report (Appendix A) we refer to historical conditions from 1970 to the present as Scenario A0.

While PATH is only looking at hydro system decisions explicitly, the effects of habitat and harvest management actions are being considered in sensitivity analyses. Sections 4.3.5 and 4.3.7 contain descriptions of how habitat and harvest are incorporated in the analysis, but generally for habitat issues we considered a range of productivities for each stock, while for harvest we explored the effects of more conservative and more liberal harvest regulations. We are developing approaches to including uncertainties with respect to management of hatcheries, to be added to our final report. The approaches used for all four H's (hydro, habitat, harvest, hatcheries) will be re-examined following peer review of this report.

Table 2-1: Hydro system management actions currently under consideration. The in-river improvement option has not yet been quantitatively defined.

Scenario	**Flow Augmentation**		**Drawdown of 4 Snake River dams**	**Drawdown of John Day Dam**
	Columbia	**Snake**		
A1 (Current Operations)	X	X	-	-
A2	Maximize transportation (without surface collectors)			
A2'	Maximize transportation (with surface collectors)			
A3	X	X	Natural River	-
A5	X	-	Natural River	-
A6 (In-river)	No transportation, flow augmentation, surface bypass collection			
B1	X	X	Natural River	Natural River
B2	-	-	Natural River	Natural River
C1	X	X	Natural River	Spillway Crest
C2	-	-	Natural River	Spillway Crest

Some members of the PATH Scientific Review Panel have recommended that, in light of some of the major uncertainties that are difficult to resolve with current information, we focus attention on experimental management options which vary management actions over time and space in a deliberate attempt to test key hypotheses pertaining to the response of fish populations. Though experimentation may pose risks to these stocks, there is risk inherent in any actions, including continuing present operations, as these populations are at dangerously low levels. We intend to explore what kinds of options for experimental management actions may be feasible for the hydrosystem, recognizing that external factors (such as ocean conditions) could confound the results of an experimental change to the hydropower system. We also recognize that there are many people working to develop and analyze alternative actions, and the creation of additional actions could affect the schedule for completion of biological, economic and social impact analyses. This report does not evaluate any experimental management strategies.

3 How to Assess the Outcomes of the Options

3.1 Overview

We have focused our analysis on future decisions, but used the past to develop our understanding of how the hydrosystem and nature interact to affect fish populations. Here we provide a general overview of our approach (Figure 3-1). The previous PATH retrospective analyses have elucidated a great deal (see PATH 1996 Conclusions Document), and have also pointed out uncertainties in past conditions due to incomplete data and potentially confounding influences (Box 1 in Figure 3-1). These uncertainties generate a range of alternative assumptions about historical conditions, such as the mortality of fish at specific dams in past years, or the success of past transportation experiments (Box 2). These alternative assumptions about the past, together with historical flow information (Box 3), are used in retrospective modeling analyses that generate quantitative estimates of parameters needed to run models into the future. This requires running both passage models, which estimate survival from Lower Granite Reservoir to Bonneville Dam (Box 4) and life cycle models (Box 6). Spawner-recruit data and environmental data (e.g. climate indicators) are also used in the life cycle models (Box5). The retrospective modeling analysis quantifies our understanding of the variability in survival rates, and the factors which affect them. Results from the retrospective analysis are passed to the prospective analysis (Box 7). The prospective modeling analysis (Boxes 11 and 13) quantifies the range of possible futures, expressed as specific performance measures (Table 3-1). This set of possible futures depends on:

- the understanding and estimated parameter values gleaned from the retrospective analysis (Box 7);
- the specific future action under consideration (Box 8; scenarios A1, A2, or A3);
- the expected flows associated with this action (Box 10); and
- assumptions about future conditions, including passage survival assumptions (Box 9) such as fish guidance efficiency through bypasses around dams, and non-passage assumptions (Box 12) such as harvest schedules, habitat improvements and future climate.

For the prospective analysis, the alternative hydrosystem management actions described in Section 2 are evaluated by simulating their consequences using a linked set of models in a four-step process to generate performance measures (Figure 3-2):

1. A hydro-regulation model translates each management option into the mean monthly flows which would be observed in the Snake and Columbia Rivers at various locations, (the U.S. Army Corps' HYSER model has been used for the scenarios included in this report). The hydro-regulation model is run for the water years 1929-1988 to generate a representative set of flows, and this information is used as input to the passage models.

2. A passage model translates the projected set of flows and dam configurations and operations for a given year into the estimated passage survival of both transported and non-transported smolts through the migration corridor from the head of Lower Granite Reservoir to the tail-race of Bonneville Dam. The passage models simulate passage survival rates under each management action for the water years 1977-1992, to permit comparison with survival rates in retrospective simulations, and compute the relative improvement in survival. The longer term water record (i.e., 1929-1988) is considered in step 4. We have used two different passage models, CRiSP and Spring FLUSH, which use different approaches to predicting passage survival rates. The primary differences between these two models are described in Section 4.2.1 and Appendix A, Section A.2.1.

3. One of the key pieces of information passed from the retrospective modeling analysis to the prospective analysis are estimates of the ratio of post-Bonneville survival rates of transported and non-transported fish. These ratios are generated by combining estimates of historical passage survival rates with the results of transportation experiments.

4. A life-cycle model generates a range of possible spawner abundances for each stock and year, under each management option. It does this by combining information produced by the passage models (i.e., the projected passage survivals, fraction of fish transported, and post-Bonneville survival assumptions) together with estimates of the other (non-passage) influences on survival (i.e., stock productivity, adult survival during upstream migration and harvest, post-Bonneville mortality, climate conditions, habitat changes, and harvest). The life-cycle model performs a thousand simulations for a given set of passage model inputs to ensure that the full range of possible ways the system works, and thus the full range of possible futures, is adequately simulated, and that the uncertainty in performance measures (Table 3-2) is properly estimated. These simulations randomly select passage model outputs from each of the years 1977-1992 according to how frequently the flow in each year occurred in the long term historical record (1929-1992). For example, an extremely low flow year like 1977 (the lowest flow in the entire 1929-1992 period) is selected much less frequently than a more typical flow year like 1979 or 1985. The life cycle model also considers alternative assumptions with respect to whether upstream and downstream stocks have some common responses to climate fluctuations (DELTA approach) or respond independently (ALPHA approach).

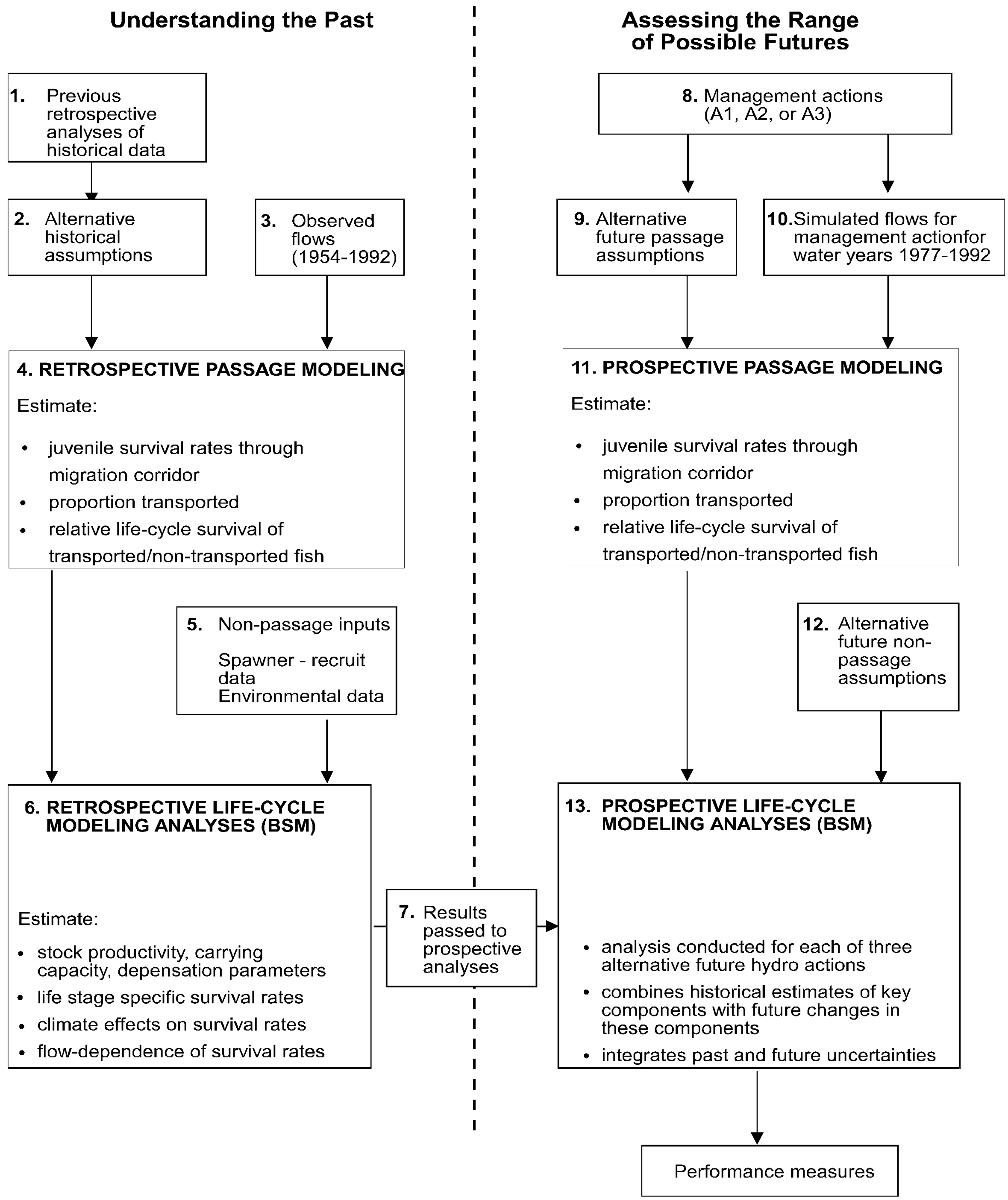


Figure 3-1: General overview of analysis. This diagram oversimplifies the actual analytical approach. More details are provided in Appendix A, section A.1.

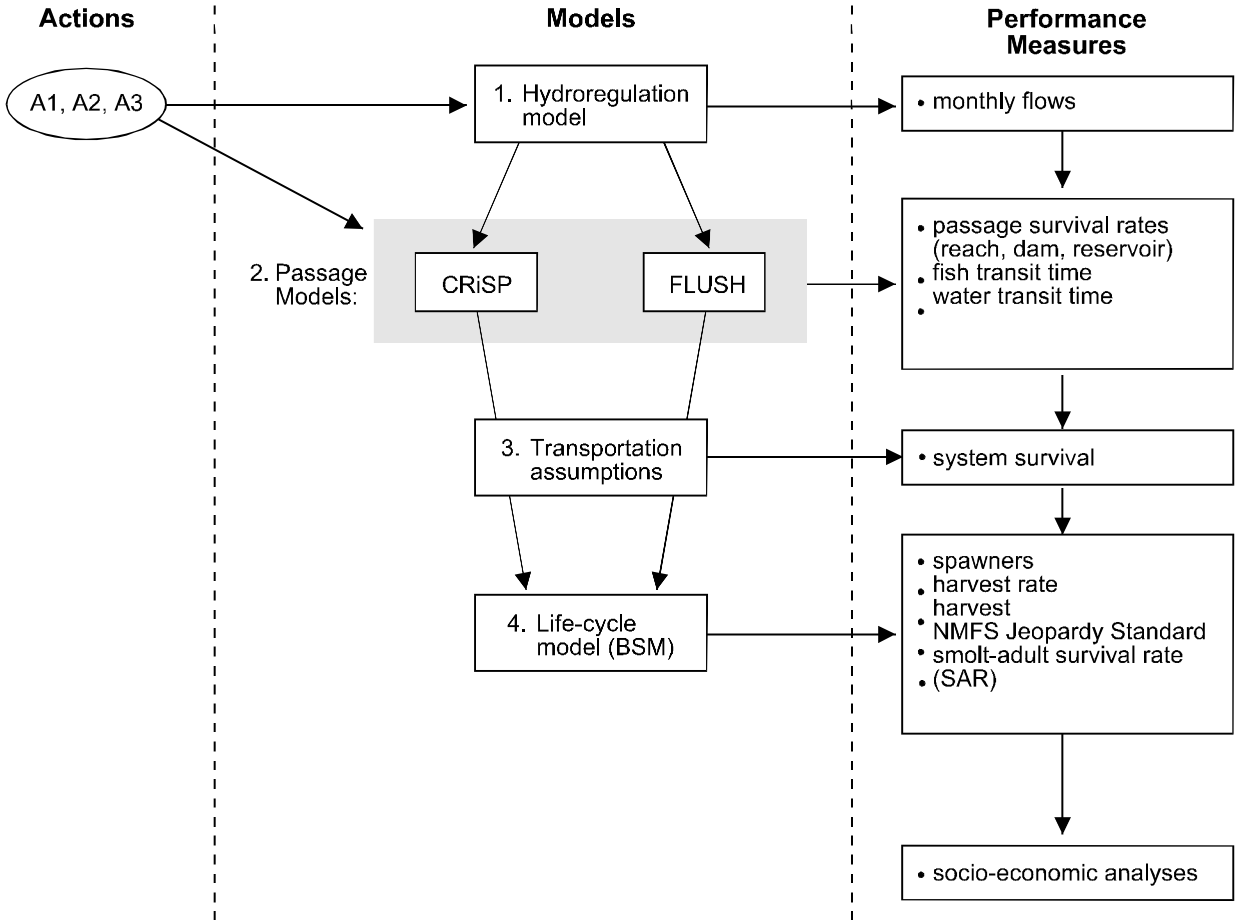


Figure 3-2: Links between the three model types and performance measures. Performance measures are explained in more detail in Section 3.2.

Because there is randomness and uncertainty in the hydrosystem and in salmon populations, as well as in future climate, we need to assess different hydrosystem actions in terms of the chances of various outcomes. To do this, we make many simulation runs, with each run based on a randomly selected value for factors such as future flows and productivity parameters. We then calculate the probability that a certain outcome occurs as the fraction of simulations that produce that outcome. For example, we might want to know the probability that the abundance of a salmon stock was going to become larger than a certain level within 50 years. If we ran 1,000 50 year simulations and the stock exceeded that level in 600 of the simulations, the estimated probability would be 0.6.

3.2 Performance Measures

The outcomes of alternative hydro management actions are evaluated in terms of various performance measures. These measures are used to rank alternative actions according to how well they meet specified management goals. A variety of performance measures have been developed to assess the biological implications of different management actions (Table 3-1). This list of performance measures was generated through discussions with the Implementation Team and analysts working on the socioeconomic analyses associated with the U.S. Army Corps Lower Snake River Feasibility Study. Because our primary goals are to determine the hydrosystem actions that should be taken to prevent extinction and lead to recovery of stocks, we focus here on the NMFS jeopardy standards that account for each of these. These standards are defined below.

Table 3-1: List of performance measures in PATH decision analysis of spring/summer chinook

1. QUANTITATIVE PERFORMANCE MEASURES	
A.	**Direct output from life cycle simulation model (BSM)**
	X Median, 10th and 90th percentiles of number of spawners in each simulated year for each stock [24-, 48- or 100-year simulations that sample entire range of parameter values]
	X Harvest rate (calculated for aggregate Snake River run; assumed to apply to each individual stock)); median, 10th and 90th percentiles of in-river and terminal harvest, by stock
	X SARs (smolt to adult return rates)[1]
	X NMFS Jeopardy Standards
	X other measures of probability of survival
B.	**Diagnostic or Intermediate (produced for verification of results and understanding influences on primary and secondary measures)**
	X estimated dam passage survival rate, estimated reservoir survival rate, reach survival rate
	X Fish Transit Time (FTT), Water Transit Time (WTT)
	X system survival[2] (survival rate through the hydro system, considering delayed mortality)
2. QUALITATIVE PERFORMANCE MEASURES	
	X To what degree is action consistent with normative river concepts (Return to the River)
	X Genetic/life history **-** does action concentrate total production in a few stocks, or does it improve survival over all stocks (compare changes in distn. of spawners over index stocks)

NMFS Jeopardy Standards

The way in which a specific hydrosystem action affects the chance of an individual spawning stock going extinct is difficult to estimate because it involves population behavior at low abundance, something we have little experience with. The performance measure we use to describe the possibility of extinction here is called a "Survival" standard. This was developed by the Biological Requirements Working Group (BRWG 1994), and has largely been accepted by NMFS for use in Snake River chinook salmon jeopardy determinations (the NMFS's approach is described in Appendix D). The Survival standard is the fraction of time during many simulations that the spawning abundance of a stock is above a certain specified low threshold. The threshold level used is either 150 spawners or 300 spawners depending on the characteristics of the stock and the stream. These levels were chosen because below these levels, spawner-recruit relationships are poorly known and unpredictable changes in population behavior are likely to occur. This standard is calculated for simulations run over 24 years and simulations over 100 years.

The effect of a certain hydrosystem action on the chance of a spawning stock recovering is described by the

[1] In this document, smolt to adult return rates (SARs) are defined as the percent of smolts counted at Lower Granite Dam which survive to return to Lower Granite Dam. However, SARs can be measured from any smolt counting location (e.g., the smolt trap near the mouth of the Warm Springs River). For clarification on how SARs are computed from BSM, see Appendix A.

[2] System survival is calculated as the number of "in-river equivalent smolts" below BON divided by the population at the head of the first reservoir. The numbers of transported smolts at each collector project that survive to BON are converted into in-river equivalents by considering the relative survival of transported and non-transported fish post-BONN. See Figure 4.2-1 and Appendix A, section A.3.2.

"Recovery" standard chosen by the BRWG (see details in Appendix D). The recovery standard is the fraction of simulation runs for which the average spawner abundance over the last 8 years of a 48-year simulation is greater than a specified level. The specified level of abundance (the recovery level) is different for each stream, and is 60% of the pre-1971 brood-year average spawner counts in each stream. We used the geometric mean abundance of spawners over the last eight years as an index of escapement to reflect the skewed distribution of abundances normally observed over time.[1]

<insert Minam River figure here>

Figure 3-3: Recent trends in Minam River spawning abundance to 1991, relative to survival and recovery levels under NMFS jeopardy standards. Also shown are the 24, 48, and 100-year periods for future projections.

These descriptions are for single stocks, but the overall performance of the system under different options needs to be described in terms of how each option affects a representative sample of all listed stocks in an Evolutionarily Significant Unit (ESU). To apply these performance standards to a number of stocks, NMFS has defined an overall Jeopardy Standard which considers, among other things, these model-derived probabilities as measures of the ability of an action to prevent extinction of an endangered stock. To meet this standard, an action must result in a "high percentage" of available populations having a "high likelihood" of being above the survival threshold level and a "moderate likelihood" of being above the recovery level. "High" and "moderate" likelihoods have been informally defined as being 0.7 for survival standards, and 0.5 for recovery standards. NMFS has defined "high percentage" of stocks as 80% of the available populations. For the cases in which we are focussed on the seven Snake River index stocks, this means that for an action to be considered to have met the overall jeopardy standard, the action must result in six stocks having a probability of 0.7 or greater of being above the survival threshold and a probability of 0.5 or greater of being above the recovery threshold.

[1] We are comparing the geometric mean of simulated future escapements with the arithmetic mean of historical abundances (recovery standard). This difference in summary statistics is recognized, but we use this method because the recovery levels are generally accepted targets, and the geometric mean is an accepted summary statistic for skewed distributions.

3.3 Stocks Considered

This analysis focuses on performance measures for seven index Snake River spring-summer chinook stocks, with comparisons to the same projections for six Lower-Mid Columbia River stocks, which pass a smaller number of dams (Table 3-2). Within these two sub-regions, these thirteen stocks are the only ones for which sufficient historical data exist to develop spawner-recruit relationships, required for generating projections of future stock sizes. The next version of this report will include three Upper Columbia stocks (the Methow, Entiat and Wenatchee). Further work is required to generalize results from these stocks to all wild chinook populations of the Snake River, Lower Columbia and Upper Columbia basins.

Table 3-2: Index stocks used in this report.

River / Region	Index Stock	Brood years with Spawner-Recruit Data
Snake R.	Minam (Snake R.)	1954-1990
	Imnaha (Snake R.)	1949-1950, 1952-1990
	Bear Valley/Elk (Snake R.)	1957-1990
	Poverty Flat (Snake R.)	1957-1990
	Johnson (Snake R.)	1957-1990
	Sulphur (Snake R.)	1957-1990
	Marsh (Snake R.)	1957-1990
Low-Mid Columbia	John Day Mainstem	1959-1990
	John Day Mid Fork	1959-1990
	John Day North Fork	1959-1990
	Warm Springs	1969-1990
	Klickitat	1966-1990
	Wind River	1970-1990

3.4 Limitations of Current Performance Measures

Because the NMFS performance standards which we have used here do not directly address extinction, some qualifications are necessary. First, the jeopardy survival standard is not a probability of a population or a cohort surviving over a certain time period, but rather the probability of a spawning escapement being above a certain spawning abundance. This assumes the population does not go extinct in the time period. Interpreting the survival standard as a probability of survival can lead to apparent inconsistencies. For example, the survival standard over 100 years can be greater than the standard over 24 years under the same conditions, something which would not be possible if it were a true survival probability. A second qualification is that the recovery standard does not explicitly recognize the risks of extinction during the first 40 years (where extinction means that the population falls below a 'quasi-extinction' level that inevitably leads to extinction). This would cause the estimates of probability of recovery given by this performance standard to be biased high. This bias will probably be inconsequential if conditions are constant throughout the 48 years, but in situations where poor initial conditions are followed by good conditions during the last years of the period, the bias may be substantial. This sort of fluctuation is considered under some climate hypotheses (see Section 4.5.3).

While the performance measures we have used provide valuable information regarding the probability of extinction, it has certain drawbacks and we are working on ways of more realistically describing the risks to stocks. For example, when spawner abundances drop below 150 or 300, there is an increasing chance that the spawner-recruit relationship will change in a way that increases the probability of extinction, but such changes are not included in the current life cycle model. Another shortcoming of the current approach is that it focusses on the numbers of spawners in single years. This does not accurately reflect the risk to stocks because a population with low spawning abundance in one year could have large cohorts in the ocean ready to spawn in subsequent years. Conversely, it would be particularly serious if a stock remained below 150 spawners for several consecutive years.[1] We intend to include other performance measures in our next report.

There are other performance measures of potential value for our final report. For example, when a particular action falls short of reaching one or more jeopardy standards, it may be worthwhile to assess how much of an increase in life cycle survival is required to meet the standard. This would help to quantify the magnitude of shortfall in the proposed action.

[1] This is not a hypothetical situation. Out of the last five brood years (1991-1995), numbers of spawners for Snake River index stocks were frequently below their respective threshold levels: Bear Valley/Elk, Marsh, Sulphur, Poverty Flat, Johnson, Imnaha, and Minam populations were less than the threshold 4, 4, 3, 2, 2, 2, and 3 times respectively.

4 Uncertainties in the Response of Populations to Management Actions

> Conscientiously followed, the method of the working hypothesis is an incalculable advance upon the method of the ruling theory; but it has some serious defects. One of these takes concrete form, ... in the ease with which the hypothesis becomes a controlling idea. To avoid this grave danger, the method of multiple working hypotheses is urged. It differs from the simple working hypothesis in that it distributes the effort and divides the affections. … The investigator thus becomes the parent of a family of hypotheses; and by his parental relations to all is morally forbidden to fasten his affections unduly upon any one. … the right use of the method requires the impartial adoption of all alike into the working family. The investigator thus at the onset puts himself in cordial sympathy and in parental relations (of adoption, if not of authorship) with every hypothesis that is at all applicable to the case under investigation. Having thus neutralized so far as may be the partialities of his emotional nature, he proceeds with a certain natural and enforced erectness of mental attitude to the inquiry, knowing well that some of his intellectual children (by birth or adoption) must needs perish before maturity, but yet with the hope that several of them may survive the ordeal of crucial research, since it often proves in the end that several agencies were conjoined in the production of the phenomena. Honors must often be divided between hypotheses. …
>
> (excerpted from: T.C. Chamberlain. 1890. "The Method of Multiple Working Hypotheses". Science 15:92. Reprinted in R. Hilborn and M. Mangel. 1997. The Ecological Detective. Confronting Models with Data. Princeton University Press. 315 pp.)

4.1 Overview

The hydrosystem management actions under consideration can affect fish populations through changes in juvenile survival past dams and through reservoirs, changes in estuarine and ocean survival, or changes in the survival of adults returning upstream. During the last two years, the PATH group has made considerable progress in clarifying the historical changes that have occurred to spring-summer chinook and standardizing the historical data sets used by models (FY96 PATH Retrospective Report and Conclusions Document). Nevertheless, there do remain many uncertainties which can potentially affect the responses of fish populations to management actions. We have laid out a range of alternative hypotheses for each of these uncertainties.

In this chapter, we provide an overview of the different uncertainties and hypotheses used in the preliminary decision analysis. Further details, the biological rationale, and the mathematical representation of alternative hypotheses are described in Appendix A. Chapter 5 presents the results of our modeling analyses to assess both the range of responses of fish populations to management actions, and the relative importance of these uncertainties in determining those responses. This draft does not attempt to weigh simulation outcomes on the basis of the strength of evidence for alternative hypotheses. In the results presented in Chapter 5, each alternative hypothesis is given equal weight. We do, however, assess the sensitivity of decisions to different weights for key hypotheses. The final report will contain more complete descriptions of alternative hypotheses, provide structured evidence for and against each alternative, and, wherever feasible, assign different weights to these alternatives.

4.1.1 Alternative Futures for Spring Summer Chinook

In exploring the range of possible futures for these stocks under different management actions, we examined twelve different uncertainties, and formulated alternative hypotheses for each one (Figure 4.1-1, Table 4.1-1). Figure 4.1-1 shows the management actions on the left side of the diagram, and then a series of branches to incorporate alternative hypotheses about different components of the system. Each small circle represents a "node" in the decision tree, where a choice of alternative hypotheses is possible. We have not included all of the branches because it would make the figure too cluttered (altogether there are over 5,000 possible paths through the tree). Table 4.1-1 represents the same information as Figure 4.1-1, but in tabular form. It includes a brief description of the differences among alternative hypotheses. These alternatives are outlined in greater detail both within this chapter and in Appendix A.

The uncertainties in Figure 4.1-1 and Table 4.1-1 are of two types: uncertainty regarding the future environment, and uncertainty regarding how the system works (i.e., in the functional relationships we use to predict the future). Though the future environment may be beyond human control (e.g., future climate), it nevertheless is of potential significance in determining future population sizes. We consider our uncertainty in how the system works primarily in terms of the survival changes caused by management actions (e.g., changes in reservoir survival in response to changes in river velocity; changes in juvenile survival after natural river drawdown). Alternative hypotheses to describe how the system works often hinge on the interpretation of historical information, because the functional relationships in models are based on both general principles and historical data. Results from historical transportation experiments, for example, affect the transportation-survival relationships included in the models, which in turn are used to predict the response to a management action. Similarly, uncertainties regarding dam survival in particular past years can, in some models, affect functional relationships that relate river velocity to fish survival through reservoirs. Differences in interpretation of historical data therefore indirectly affect predictions of future states. A final note of caution is necessary with respect to historical data. While reasonable consistency with historical information is necessary to provide some minimum level of confidence in a model, the degree to which a model matches historical information does not determine how well it will perform in predicting the response to new management actions. This is because: 1) future conditions may not be the same as past ones, particularly if there are major changes made to the dams and reservoirs; and 2) the fact that a model fits past data well does not mean that the mechanisms incorporated into that model reflect those which exist in nature, so that the model may not correctly simulate the response of the system to new conditions.

As you move through the different uncertainties in Figure 4.1-1 and Table 4.1-1, the number of combinations of assumptions increases rapidly, just as the branches of a tree increase as you move up the trunk to the crown. A particular combination of hypotheses about how the system works and the future environment is a unique "prospective aggregate hypothesis", like a unique path chosen by an ant climbing from the trunk to the last branch at the top of a tree. Each prospective aggregate hypothesis potentially yields a unique future biological response. One of the objectives of the preliminary analysis reported here is to 'prune the tree', by determining which uncertainties have limited effects on performance measures and the resulting decision, so that we can focus on the most critical alternative hypotheses. We may also develop new variations or combinations of hypotheses that better reflect recent evidence. Model runs were completed for all combinations of the alternative hypotheses in Figure 4.1-1 and Table 4.1-1.

n:\davem\ew788\decana;\modelrun\document\A2treef.cdr

Figure 4.1-1: Decision tree used to incorporate alternative hypotheses in prospective analysis for spring-summer chinook. Further details on the alternative hypotheses are provided in Table 4.1-1. line is an example of a prospective aggregate hypothesis (one of 5,148 possible combinations), which is generally consistent with the retrospective aggregate hypothesis H1, described in Table 4.1-2.

Table 4.1-1: Set of uncertainties and alternative hypotheses considered in this analysis.

Uncertainty (Report section with further details)	Hypothesis Label	Description
Uncertainties / hypotheses related to downstream passage to Bonneville Dam		
In-river survival assumptions – Passage Models (4.2.1)	PMOD1	CRiSP direct survival estimates.
	PMOD2	FLUSH direct survival estimates.
Fish Guidance Efficiency (FGE) (4.2.2)	FGE1	FGE w/ESBS > FGE w/STS. (ESBS = extended length submersible bar screens). (STS = standard length submersible travel screens).
	FGE2	FGE w/ESBS = FGE w/STS.
Historical ∪ Turbine + Bypass Survival (4.2.3)	TURB1	Turbine survival = 0.9. Bypass survival = 0.97 - 0.99, depending on the project.
	TURB4 TURB 5 TURB 6	Various mechanisms for turbine/bypass survival during some historical years. Survival is lowest under TURB4, and highest under TURB5.
Predator removal efficiency (4.2.5)	PREM1	0% reduction in reservoir mortality resulting from predator removal program.
	PREM3	25% reduction in reservoir mortality.
Duration of pre-removal period under drawdown (4.2.6)	PRER1	3 years
	PRER2	8 years
Equilibrated Snake River juvenile survival rate under drawdown (4.2.6)	EJUV1	0.85
	EJUV2	0.96
Transition Period: Juvenile survival (4.2.6)	TJUVa	Survivals reach equilibrated values 2 years after dam removal.
	TJUVb	Survivals reach equilibrated values 10 years after dam removal.
Other uncertainties / alternative hypotheses		
Transportation models (section 4.3.1)	TRANS1 or T1 (FLUSH only)	Relationship established between TCR and FLUSH in-river survival, based on data from **all** transport studies conducted at LGR and LGO dams between 1971-1989. This relationship, and FLUSH in-river survival, used to estimate relative post-BONN survival of transported fish (D) in both retrospective and prospective analyses (see Fig. 4.2-1).
	TRANS2 or T2 (FLUSH only)	TCRs derived from TRANS1 adjusted by 0.83 to reflect poorer survival of transported fish from last dam to spawning grounds.

Uncertainty (Report section with further details)	Hypothesis Label	Description
	TRANS3 or T3 (CRiSP only)	For pre-1980 retrospective analyses, relative post-BONN survival set at average D-value estimated from seven T:C studies in 1970's and associated CRiSP in-river survival rate estimates. Post-1980 retrospective analyses use average D-value estimated from four T:C studies in 1980's, and CRiSP in-river survivals. For prospective analyses, D-value randomly selected from four 1980 values.
Distribution of Extra Mortality (4.3.2)	ALPHA	Extra mortality[1] is specific to each sub-region, and affected by climate variables.
	DELTA	Extra mortality is independent of the common year effects which affect several subregions.
Extra mortality / Future climate (4.3.3, 4.3.4)	EMCLIM1	Extra mortality is here to stay; future climate is sampled from historical distribution with autoregressive properties.
	EMCLIM2	Extra mortality is here to stay; future climate follows cyclical pattern.
	EMCLIM3	Extra mortality is proportional to hydropower-related mortality, future climate is sampled from historical distribution with autoregressive properties.
	EMCLIM4	Extra mortality is proportional to hydropower-related mortality, future climate follows cyclical pattern, with both long (60-year) and shorter (18-year) cycles.
	EMCLIM5	Both extra mortality and future climate follow cyclical pattern.
Habitat Effects (4.3.5)	HAB0	Same management as current.
	HABB	Implementation of all possible habitat restoration or protection.

4.1.2 Differing Explanations of Historic Declines in Spring Summer Chinook and Their Link to Future Projections

We base our future projections on our understanding of the past, much of which is summarized in the PATH FY96 Retrospective Analyses, the PATH FY96 Conclusions Document, and FY97 retrospective analyses (to be published). However, as past information is incomplete, there are differing interpretations of the relative importance of different factors in causing recent declines of Snake River spring-summer chinook. As described above, a set of alternative hypotheses about all components of the system (stock productivity, downstream migration, marine survival, etc.) is referred to as an aggregate hypothesis. We call a set of hypotheses about the future a "prospective aggregate hypothesis", and a set of hypotheses about the past a "retrospective aggregate hypothesis". The set of hypotheses about the past is smaller than the set of hypotheses about the future, because though there are many possible alternative futures, there is only one past.

[1] Extra mortality is any mortality that is not accounted for by either: 1) spawner-recruit relationships; 2) estimates of direct mortality within the migration corridor; or 3) for the Delta model only, common year effects affecting both Snake River and lower Columbia River stocks.

Two example retrospective aggregate hypotheses are presented in Table 4.1-2. There are many other possible hypotheses, but these two examples illustrate some key differences in our interpretation of the past which have important consequences for influencing the range of future projections. The H1 aggregate hypothesis proposes that recent declines are primarily due to the hydrosystem, while the H2 aggregate sypothesis attibutes declines primarily to non-hydro factors. The logical consequences of these two *retrospective* aggregate hypotheses are represented by subsets of the *prospective* aggregate hypotheses considered in our future projections. For example, the branches in Figure 4.1-1 involving FLUSH-T1/T2, the DELTA model, and Hydro-related extra mortality are more consistent with the H1 retrospective aggregate hypothesis in Table 4.1-1. The branches involving CRiSP-T3, the ALPHA model, and regime shift or BKD extra mortality are more consistent with the H2 retrospective aggregate hypothesis in Table 4.1-1. Note that for virtually all system components (the rows of the table) there are common areas of agreement (statements which span the H1 and H2 columns in Table 4.1-2). There are, however, some significant differences among these two retrospective aggregate hypotheses; the last column outlines why such differences exist, and what might be necessary to resolve them. The table references various figures which follow, as well as chapters of the PATH FY96 Retrospective Analyses, that help to illustrate the differences among alternative hypotheses for various components. *Readers primarily interested in the prospective analysis may wish to skim Table 4.1-2 and move directly to Section 4.3.*

Table 4.1-2: Examples of two retrospective aggregate hypotheses (H1 and H2). These aggregate hypotheses are two of the many alternative explanations of the 1970-1990 declines in Snake River spring-summer chinook stocks. Both hypotheses recognize that the relative importance of different factors has varied over time. Statements which span the columns for H1 and H2 are common to both hypotheses.

System Component [Abbreviations in Table 4.1-1, Figure 4.1-1; Chapters from FY96 Report; relevant sections in this report]	**Aggregate Hypothesis H1. Mostly Hydro** **BY1970-74: Hydro** **BY75-90: Hydro primary; climate secondary**	**Aggregate Hypothesis H2. Mostly Non-Hydro** **BY1970-74: Hydro** **BY74-90: climate primary; hydro, hatcheries, and habitat secondary**	**Key sources of differences between H1 and H2, and possible ways to resolve them.**
a. Direct Survival of In-River Fish	- poor in-river survival for brood years 1970-74 in part due to inadequate passage facilities at dams		
PMOD, TURB [Ch. 6] [[4.2.1, 4.2.3]	- fish survival rate vs. FTT[1] non-linear - cumulative effect of longer FTT leads to lower reservoir survival per day in JDD -> BONN reach than in other sections - historic dam mortality lower than in H2; reservoir mortality higher	- fish survival rate vs. FTT essentially linear - similar per project survival in JDD -> BONN projects as in other ones - historic dam mortality higher than in H1; reservoir mortality lower	- reach survival data mostly for LGR to MCN; general H1/H2 agreement there. - differences between H1 & H2 caused by lack of PIT-tag studies for lower reaches, and different interpretations of 1973 and 1979 survivals.
b. Transportation TRANS or T	- success of transportation varies both among years and within years		
[Ch. 6] [4.3.1]	- benefits are variable - transport studies from all years (1971-89) are representative of future - transported fish do much worse in the ocean than in-river fish - future D (see footnote 2) values likely to average <0.5 - transport survival not improving over time (SAR data).	- large benefit overall - only studies from 1980's and beyond are representative of future - transported fish do about as well as in-river fish in the ocean - future D values likely to be 0.8 or higher - transport survival improving over time (T:C data, descaling rates)	- no direct way to measure D - all estimates of D are indirect, based on T:C ratios[3] and estimates of in-river survival. - differences in estimated in-river survival (a) magnify differences in transportation benefits (b) - well designed tagging studies to assess the effectiveness of transport - well designed tagging studies to assess the effectiveness of transportation
c. Post-BONN mortality EM; ALPHA/DELTA [Ch. 5] [4.3.2, 4.3.3]	- mostly caused by hydrosystem during brood years (BY) 1970-1974		- not enough CWT recoveries to know if significant upstream-downstream differences exist in marine distribution; differing interpretations of this meager data; no empirical estimates of upstream-downstream differences in marine survival - climate data confirm temporal changes in conditions, but no way to estimate upstream stocks' relative resilience to climate shifts.

[1] FTT=Fish Transit Time, the time it takes smolts to travel from the head of Lower Granite pool (LGR) to the Bonneville tailrace (BONN).

[2] D=(post-BONN survival of transported fish)/(post-BONN survival of in-river fish). If D=1, post-BONN survivals of the two groups are equal. Estimates of D are affected by which T/C studies are used, and estimates of in-river survival. See Figure 4.2-1.

[3] T:C is defined in Figure 4.2-1. The T:C ratio is also referred to as TCR.

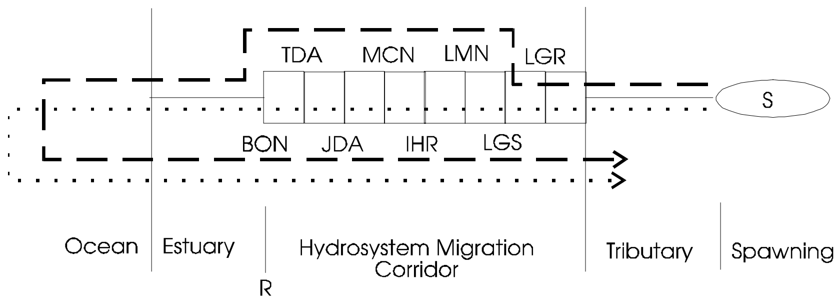
System Component [Abbreviations in Table 4.1-1, Figure 4.1-1; Chapters from FY96 Report; relevant sections in this report]	**Aggregate Hypothesis H1. Mostly Hydro** **BY1970-74: Hydro** **BY75-90: Hydro primary; climate secondary**	**Aggregate Hypothesis H2. Mostly Non-Hydro** **BY1970-74: Hydro** **BY74-90: climate primary; hydro, hatcheries, and habitat secondary**	**Key sources of differences between H1 and H2, and possible ways to resolve them.**
c. Post-BONN mortality (cont.) EM; ALPHA/DELTA [Ch. 5] [4.3.2, 4.3.3]	- mostly driven by hydrosystem in BY75-90, except in BY80-83 [Ch. 5] - overall effect of hydrosystem (m) indicated by differences in (R/S) between Snake R. and lower Columbia R. stocks, after accounting for common climate effects (δ) affecting both groups [Ch. 3 and 5] - post-BONN mortality assumed due to hydrosystem, and estimated from difference between overall effects (m) and direct effects (system survival[1]). - much spatial/temporal overlap of upstream/downstream stocks in estuary and near ocean environments.	- mostly climate-driven for BY75-90 (climate good for 1945-1975, poor for 1975-2005, then good for 2005-2035) [Ch. 12] - upstream-downstream differences in (R/S) caused by factors other than hydrosystem (i.e., different genetic composition, ocean distributions, marine survival, resiliency to climate shifts) - post-BONN mortality (for Snake R. stocks) assumed due to shift in climate regime, estimated from 'step-drop' in (R/S) of Snake R. stocks, and system survival. - upstream stocks distinctly different from downstream stocks.	- resiliency of upstream Snake River stocks may have been lowered by hydrosystem, confounding effects of climate and hydrosystem.
d. Climate	- climate had more negative effects on salmon after the mid-1970's		
CLIM [Ch. 12] [4.3.4]	- climate generally positive for BY 1950-1970; generally negative for BY 1974-1990 (except 1983, 84, 85, 88)	- climate strongly positive for BY1945-1974; strongly negative for BY 1975-2005. - minor 18.5-year cycles on top of major shifts can improve or worsen conditions	- climate regime-shift coincided with major transportation and hatchery programs; - difficult to demonstrate relative contribution of each factor to depressing (R/S) and Smolt-Adult return rates without an experimental change in hydrosystem, transportation, or hatcheries, out of phase with climate changes

[1] System survival = the weighted average of 1) the passage survival of non-transported smolts from the head of LGR pool to BONN tailrace; and 2) the survival of transported fish from the point of collection back to this point. The two survivals are weighted according to the fraction of fish that end up being transported.

System Component [Abbreviations in Table 4.1-1, Figure 4.1-1; Chapters from FY96 Report; relevant sections in this report]	**Aggregate Hypothesis H1. Mostly Hydro** **BY1970-74: Hydro** **BY75-90: Hydro primary; climate secondary**	**Aggregate Hypothesis H2. Mostly Non-Hydro** **BY1970-74: Hydro** **BY74-90: climate primary; hydro, hatcheries, and habitat secondary**	**Key sources of differences between H1 and H2, and possible ways to resolve them.**
e. Hatcheries [Ch. 11] [4.3.6]	- no significant effects of separate hatchery additions in individual streams		
	- effects of aggregate hatchery additions difficult to determine due to overlap in space and time with hydrosystem effects; presumed low [Ch. 11] - important to examine coincidence in timing of ocean entry of hatchery and wild fish for both upriver and downriver stocks to assess if impact is feasible	- Hatchery fish: 1) spread disease to Snake River wild fish, lowering their resilience to climate changes; and 2) compete with Snake River wild fish for food in early ocean period, when productivity lower due to shift in ocean conditions - these effects are greater on Snake River fish than lower Columbia River fish - Snake R. spring/summer chinook declined as hatchery chinook and steelhead releases increased	- changes in disease rates in Snake R. fish insufficient to explain declines without assuming greater sensitivity to disease than lower Columbia stocks - differing interpretations of comparibility of upstream and downstream stocks - impossible to disentangle historical effects of: 1) hydrosystem on early ocean productivity (i.e., blocked nutrients, changed seasonal flow patterns); 2) natural changes in ocean productivity; 3) competition from hatchery fish; 4) impact of hydrosystem on disease transmission (bypasses, barges); and 5) reduced fitness of smolts due to barge transportation - only way to test these effects would be to alter hydrosystem, shut off transportation or hatcheries in successive years, while monitoring 'control' stocks for climate changes (see (d))
f. Spawning and Rearing Habitat HAB [Ch. 4, 9] [4.3.5]	- important factor pre-1975; maintaining habitat is critical to survival of stocks		- scope for increasing productivity and survival rates through habitat improvement - adaptive management experiments to improve degraded habitat
	- not significant in determining rate of 1975-1990 decline - smolts/spawner in 1990's not significantly different from 1960's for Snake R. aggregate stock [Ch. 9] - no significant correlations between changes in land use and trends in (R/S) [Ch. 4]	- reduced resiliency of some stocks in poorer habitat (linked to (h) depensatory mortality) - degradation in some rearing, migratory corridor, estuary, and ocean habitats may have impact on stocks	
g. Harvest [Ch. 13] [4.3.7]	- minor effects on BY 1970-1975		- spring/summer harvest recognized as less significant factor than hydro, climate, hatcheries, and habitat
	- not significant after 1975	- current low harvest rates continue to have some effect	
h. Depensatory mortality and stock productivity [4.3.2]	- generally higher inherent productivities - depensation may have occurred, but not yet observed; could be increasingly important in future	- generally lower stock productivities - low spawning numbers may cause reduced size and fitness of smolts due to lack of carcasses and marine nutrients	- not enough observations at low stock sizes to show reduced (R/S) or reduced smolt size at low numbers of spawners

4.2 Uncertainties/Alternative Hypotheses Related to Downstream Passage

Models simplify reality. The parts of the salmon's life cycle represented in the passage models are shown in Figure 4.2-1. The caption to Figure 4.2-1 also defines some of the terms used to distinguish among alternative hypotheses. Abbreviations for Columbia River dams are defined in Table 4.2-1. As described in Section 4.1, key outputs from the passage models include direct survival of in-river and transported fish, the partitioning of in-river survival between dam and reservoir survival, expected transport:control ratios, and the proportion of fish transported. These outputs feed into a life-cycle model, which is described in Section 4.3.



LEGEND	
. . . .	path of a fish travelling only in-river (not transported)
	Path of a fish collected at Little Goose Dam (LGS), barged to Bonneville (BON) tailrace, and released

Figure 4.2-1: Schematic showing components of life cycle modeled and definition of terms. 'Direct Survival' of in-river juvenile fish (V_n) is survival from head of Lower Granite (LGR) pool to tailrace of Bonneville Dam., including reservoir and dam survival at each project. 'Direct Survival' of transported juvenile fish (V_t) is in-river survival from head of LGR to point of collection (LGS in the example), multiplied by bypass survival at collection project, multiplied by barge survival to BON tailrace. The passage models predict V_n and V_t. Transport:Control ratio (T/C or Φ) is the ratio of survival of transported fish survival to in-river fish survival from juveniles at collection point to adults at the same point (i.e., in example shown, from juveniles at LGS through ocean and back to adults at LGS). Recruits per spawner (R/S) is the number of adult fish returning to BON (R in Figure 4.2-1) , divided by the number of spawners in the parent generation (S). D is the ratio of survival of transported fish measured from BON tailrace through the ocean and back to the point of collection (i.e. $\lambda_{t;}$ excluding downstream migration corridor) to survival of in-river fish, measured over the same interval (λ_n). Putting all these together gives the equation:

$$(T/C) \text{ or } \phi = [V_t/V_n] * [\lambda_t/\lambda_n] = [V_t/V_n] * D \qquad \text{[Eq. 4.2-1]}$$

Table 4.2-1: Abbreviations used for Columbia River System Dams

BON	Bonnevile
TDD	The Dalles
JDA	John Day
MCN	McNary
IHA/IHR	Ice Harbor
LMO/LMN	Lower Monumental
LGO/LGS	Little Goose
LGR	Lower Granite

4.2.1 Passage Models

PATH has used two passage models in our analyses of spring/summer chinook: CRiSP and Spring FLUSH (throughout this document, we use "FLUSH" to refer to Spring FLUSH. There is a version of FLUSH for fall chinook called Fall FLUSH). The reason for using two models is that they represent different approaches to modeling reservoir mortality, dam passage mortality and transportation mortality. These are the three main components for which different hypotheses exist within these two models, and different ways of representing these hypotheses mathematically. CRiSP simulates changes to fish populations using a more detailed, mechanistic approach, while spring FLUSH relies on a more aggregated, empirical approach. Peer reviewers have found strengths and weaknesses in both approaches (SRP, 1994, 1996).

We provide here a brief overview of the major differences in structure and hypotheses among these two models; more details are provided in Appendix A.2.1. CRiSP simulates conditions for each day of the juvenile migration season: the monthly flows generated by the hydroregulation models are interpolated to daily flows and velocities; each day's simulated group of fish moves down the river at a speed consistent with that day's velocities and other factors; the dam mortality they experience is determined by the daily proportions of fish passing through turbines, spillways or bypasses; and the reservoir mortality of fish is determined by their daily encounter rates with predators and high gas levels. FLUSH uses a more aggregated approach: the monthly flows generated by the hydroregulation models are converted into the average water velocities and fish transit times experienced by smolts over the migration season; mortalities at turbines, spillways and bypasses are calculated based on the average proportions of fish passing through these alternative routes; and reservoir mortality is based on an empirical relationship between fish transit time and reservoir survival (i.e., the longer the transit time, the poorer the survival). This empirical relationship in FLUSH is based on fits to historical data on reach survivals and assumptions about the level of turbine and bypass mortality before 1980 ('TURB' alternative hypotheses in Table 4.1-1 and Figure 4.1-1). Hypotheses with greater historical turbine and bypass mortality (e.g. TURB4) yield a weaker relationship between reservoir survival and fish travel time. The key difference in the behavior of the two models is that the survival of fish through reservoirs is more responsive to changes in fish transit times within FLUSH than within CRiSP (Figure 4.2-2). FLUSH shows lower total in-river survival than CRiSP during the 1970-1995 period (Figure 4.2-3).

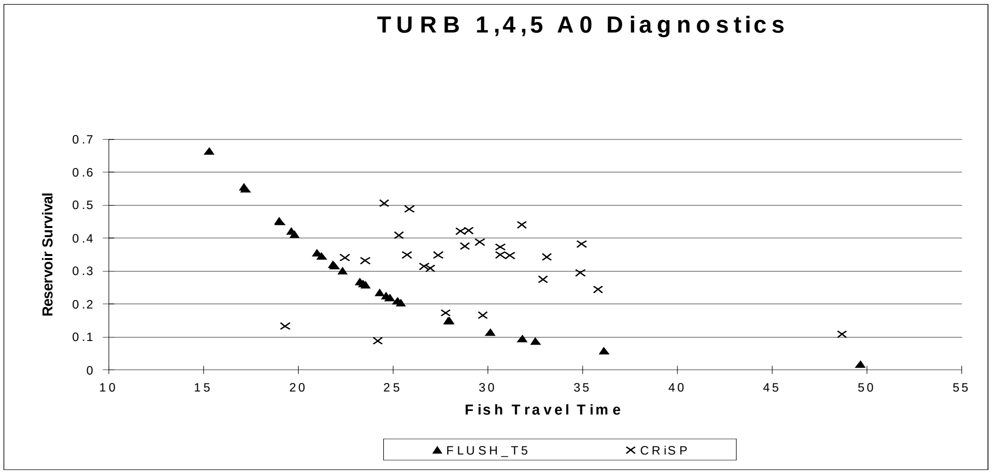


Figure 4.2-2: Model reservoir survival vs. fish travel time relationships in FLUSH and CRiSP. FLUSH example shown is for the TURB5 hypothesis.

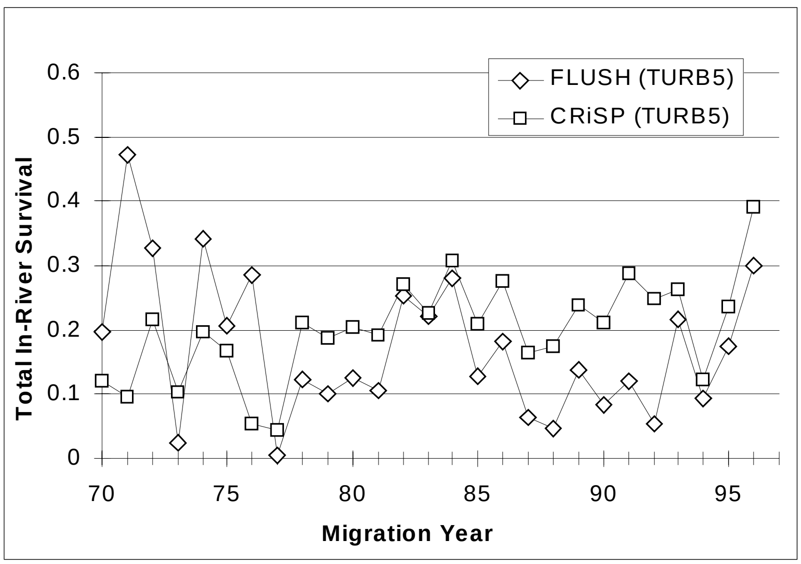


Figure 4.2-3: Comparison of CRiSP and FLUSH estimates of historical in-river survival rates from uppermost reservoir to below Bonneville Dam, not including transported fish.

4.2.2 Fish Guidance Efficiencies and Surface Collectors

Fish guidance efficiency (FGE) is typically defined as the percentage of the (juvenile) fish committed to pass through the turbine intake that, intercepted by special screens, are guided upward into the gatewell and then into a turbine bypass channel (Figure 4.2-4). However, in the past fish could only avoid turbines by swimming through a gatewell salvage system or an ice and trash sluiceway at most of the lower Snake and Columbia River dams. Fish guidance efficiency was estimated for the structural configuration in place at each project during each year of service.

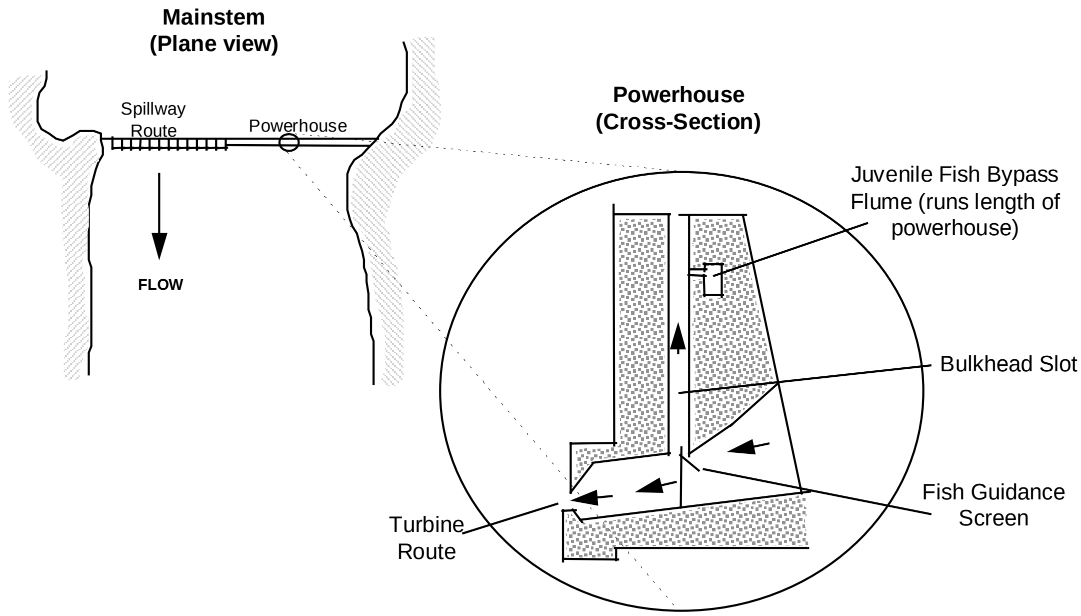


Figure 4.2-4: Spillway, juvenile bypass flume, and turbine routes of passage at a hypothetical mainstem dam.

PIT-Tag versus Fyke Net Estimates of Guidance Efficiencies

Data on fish guidance efficiencies have been obtained with fyke-net tests in the past and with PIT-tag studies in more recent years. Fyke-net estimates of FGE are thought to be biased upward and to characterize the behavior of only a limited portion of the yearling chinook run (see Appendix A, Section 4.2). Therefore, where both types of estimates were available, estimates derived from PIT-tag studies were given precedence (Table A.2.1.1.4-1)

Sensitivity Analysis for the Effect of Extended-Length Screens on FGE

Standard-length submersible traveling screens (STS) have been replaced by lowered STS or extended-length submersible bar screens (ESBS) at Lower Granite, Little Goose, and McNary Dams. Screens were extended to deflect more of the intake flow upward, with the goal of improving fish guidance into the gatewell and the bypass system. Two alternative hypotheses were derived for the effect of extended-length submersible screens on FGE, to assess the sensitivity of outcomes to FGE assumptions.

(FGE1): Assume that extended-length screens have significantly improved FGE. The results of side-by-side (fyke-net, downstream slot) tests of guidance efficiencies for yearling chinook indicate that extended-length screens guide half of the fish that are not guided by the standard-length screens.

(FGE2): Assume that extended-length submersible screens have had no effect on fish guidance efficiency. Recent PIT-tag studies with wild yearling chinook by the Idaho Department of Fish and Game indicate considerable overlap between detection rates at Lower Granite Dam during 1996 (when extended-length screens were in place) versus 1993 (standard-length submersible traveling screens in place), at similar spill levels.

4.2.3 Turbine/Bypass Survival

Standard Estimates

"Standard" estimates represent our best understanding of turbine and bypass survival under current and recent conditions. In determining the standard estimates, greatest weight was given to the most recent data, particularly those derived from PIT-tag studies.

Standard Estimate of Turbine Survival (TURB1): Turbine survival is defined as the proportion of fish surviving direct turbine passage injuries as well as any indirect mortality experienced in the tailrace by fish that passed through turbines, above the tailrace mortality experienced by fish passing through other routes (i.e., spillway or bypass). Based on a review of field studies at mid-Columbia projects and Lower Granite Dam, a value of 0.90 was adopted as the standard estimate of turbine survival. That is, of the fish that pass through turbines, 90% survive.

Standard Estimate of Bypass Survival (TURB1): Bypass survival is defined as survival past turbine intake screens, gatewells, orifices, bypass flumes, and, in some cases, dewatering screens, wet separators, sampling facilities (including holding tanks), and bypass outfall conduits. These estimates also apply to juvenile bypass through sluiceways at The Dalles, Ice Harbor, and the Bonneville Powerhouse One during certain years.

Based on a review of field studies at lower Snake and Columbia River dams, a range of 0.97-0.99 was adopted as the standard estimate of bypass survival, depending upon the bypass method at a given project in a given year.

Non-standard Estimates

Historical estimates of bypass and turbine mortality vary from current estimates for some projects during some years. There was general agreement that, between 1980 and the present, the standard estimate of turbine and bypass applied. However, there is less certainty about survival estimates prior to 1980, so several alternative hypotheses were described.

TURB4: Survival due to passage through these routes is significantly lower than would be predicted based on bypass structure alone. Turbine and bypass survivals are described by an exponential regression function relating passage mortality to the rate of descaling at 0.25, 2, or 6 days after passage.

TURB6: Some additional debris-related mortality occurred during early years but survival was higher than estimated by TURB4. Bypass survival is a function of mortality due to descaling assuming that the rate of mortality is equal to the rate of descaling. The survival of fish passing through the turbine route is the same as that described in TURB1 (i.e., 0.90 ± 0.03). However, there is no time element in the function relating mortality to descaling. All mortality is assumed to occur instantaneously.

TURB5: As described above for TURB6 but assuming that the rate of turbine mortality is equal to one-half the rate of descaling because studies indicated that a large percentage of the descaling was caused by the screens used in those years.

All three non-standard estimates of turbine and bypass survival were incorporated into both the CRiSP and FLUSH models because studies indicated that a large percentage of the descaling was caused by the screens used in those years.

4.2.4 Spill Survival and Spill Efficiency

Spill Survival

A value of 0.98 was adopted as the standard estimate of spillway survival. This estimate is conservative and represents the findings of most of the spill survival studies conducted to date. Uncertainty in this value was explored in initial passage model runs, but it was apparent from these results that different assumptions about spill survival had very little effect on the model output. Therefore, the analyses presented in this report are based solely on the standard estimate of 0.98. We are conducting further sensitivity analyses on the effects of spill efficiency assumptions.

Spill Efficiency

Spill efficiency is defined as a ratio of the proportion of the smolt population passed via the spillway (spill effectiveness) to the proportion (percent) of total flow discharged as spill. The ratio of 1:1 has been adopted as the standard assumed value, at all dams except The Dalles, based upon Steig (1994). The Dalles has an alternative configuration to other projects, with the spillway oriented perpendicular to the natural course of the river and the powerhouse oriented nearly parallel to the river, which should produce higher spill efficiency (e.g., ISG 1996). We applied an equation that set spill efficiency at 2:1 at spill proportions less than 30% of total river flow, decreasing at higher spill proportions until 1:1 is reached at 100% spill (Willis 1982, Giorgi and Stevenson 1995, and Holmberg et al. 1997).

4.2.5 Predator Smolt Removal Efficiency

Predator removal efficiency is expressed as a percent reduction in reservoir mortality. For the preliminary decision analysis, PATH explored two alternative hypotheses. Hypothesis 1 is that the predator removal program (i.e., removal of squawfish for bounties) has no effect on reservoir mortality. Hypothesis 2 states that removal of predator results in a 25% reduction in reservoir mortality. These two values were chosen to represent the extreme bounds of probable effects in the preliminary decision analysis, and follows the approach taken in the 1995 Biological Opinion. A more empirical approach to defining these bounds is described in Appendix A.2.5; this approach or one similar to it could be used in the next round of analyses.

4.2.6 Drawdown

The only drawdown scenario evaluated in the preliminary analysis was to a natural river level drawdown of all four Snake River dams. Although alternative hypotheses for drawdown scenarios were developed by the Drawdown Workgroup, the hypotheses that were actually implemented in the preliminary analysis are slightly different because of modeling constraints. Further revision and specification of drawdown hypotheses will occur after this preliminary analysis.

The group defined four time periods that were important to consider when predicting the effects of drawdown:

1. Pre-removal - the period between when the region decides to proceed with drawdown and when physical removal of dams begins.
2. Removal - period in which engineering work to breach or circumvent the dams is carried out.
3. Transition - period beginning just after the dams are removed and continuing until fish populations attain some equilibrated conditions.
4. Equilibrium - period of time from when fish populations equilibrate to the end of the simulation period.

For each period, we need to estimate: a) the duration of the period, and b) the adult and juvenile survival rates that are expected during this period. Combining these two elements gives a trajectory of adult and juvenile survival rates before, during, and after drawdown that can be used in the model to project spawner abundances over time.

Estimates of duration, juvenile survival, and adult survival for each of the four time periods are summarized in Table 4.2.6-1. Alternative hypotheses were considered for the following elements:

a) Duration of pre-removal period - to reflect uncertainty in the Congressional appropriations process and the possibility of litigation
b) Duration of transition period - to reflect uncertainty in the physical and biological responses to drawdown (e.g., response of predators, release of sediment)
c) Equilibrated juvenile survival rate - to reflect uncertainty in the long-term physical and ecological effects of drawdown (e.g., change in density of predators)
d) Equilibrated adult survival rate – to reflect uncertainty in the long-term effect on adult survival conversion rates (the current model runs include only a single, high conversion rate)

Table 4.2.6-1: Summary of estimates of duration, juvenile survival, and adult survival for each of the four time periods

Time period	Duration (years)	Juvenile survival (over the reach corresponding to 4 Snake River projects)	Adult survival (conversion rates)
Pre-removal	3 years or 8 years	determined by passage models	current estimates
Removal	2 years	no change from pre-removal period	no change from pre-removal
Transition	2 years or 10 years	linear increase from pre-removal survival to equilibrated survival	linear increase from pre-removal to equilibrated value
Equilibrium	determined by length of simulation period	0.85 or 0.96	0.97

Two examples of the juvenile survival trajectory used in the preliminary analysis are shown in Figure 4.2-5. The examples use the same equilibrated juvenile survival rate (equates to a survival rate of 0.85 over the reach corresponding to four Snake River projects) and the same three year pre-removal period, but differ in the length of the transition period between dam removal (which is completed in 2004 in this scenario) and equilibrated levels. In these examples, a regional decision is made in 1999 and removal of dams takes place between 2002 and 2004.

For our next report, we may include sensitivity analyses to explore the potential impacts if drawdown were to concentrate predators during the initial part of the transition period.

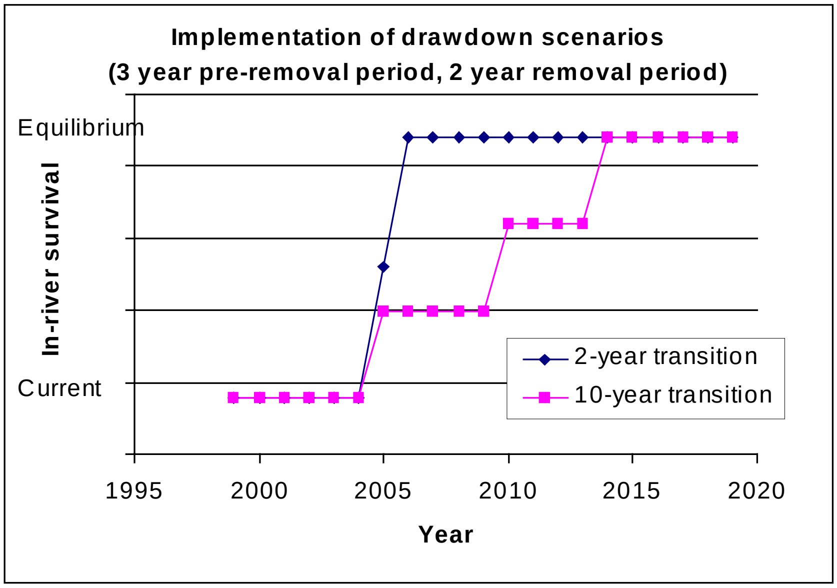


Figure 4.2-5: Example trajectories of juvenile survival following drawdown

4.3 Other Uncertainties/Alternative Hypotheses

There are six system components outside of the downstream passage portion of the life-cycle which can have alternative hypotheses:

- transportation (the processes occurring after the end of the migration corridor);
- stock productivity;
- extra mortality occurring beyond Bonneville (for both in-river and transported fish);
- future of climate conditions;
- the effects of changes in habitat management; and
- the effects of changes in hatchery operations.

Figure 4.3-1 shows the general model structures that are used to represent alternative hypotheses for both passage and non-passage components of the life-cycle. The recruitment of the fish population (R) depends on four components:

1. the basic productivity of the stock, represented by a stock recruitment relationship;
2. the system survival, which includes both the direct survival of in-river fish and the survival of transported fish, converted into in-river equivalents (see Appendix A, Section A.3.2 for more

details);

3. the post-Bonneville survival of non-transported fish; and
4. future climate changes.

Figure 4.3-1 shows how alternative hypotheses can be included for each of these four components. The basic recruitment of the stock depends on the number of spawners present (S), the parameters that define the productivity and carrying capacity of the stock (a and b), and a parameter which affects how productive the stock is at very low numbers of spawners (p). The system survival is determined by three factors: 1) the direct mortality (M, estimated by either CRiSP or FLUSH); 2) D, the ratio of post-Bonneville survivals of transported and non-transported fish; and 3) P, the fraction of smolts at Bonneville which were transported. These terms are described in Figure 4.2-1. The post-Bonneville survival of non-transported fish is affected by whether extra mortality is considered to be unique to each region (ALPHA model), or considered to have common temporal patterns in both Snake River and lower Columbia River stocks (DELTA model). The extra mortality is represented either as a function of changes in hydro system survival, changing ocean regimes, or as something completely independent of hydrosystem actions or climate. The fourth factor, shorter term climate fluctuations, can be represented either randomly or in cycles. Each of these alternative hypotheses for components of the salmon's life history are explained in more detail below.

Future projections of salmon populations are based partly on historical calibrations to spawner recruit data (Box 6 in Figure 3-1). In fitting a basic equation shown in Figure 4.3-1 to historical data, whenever one factor shows a smaller effect (e.g., a higher system survival), one or more of the other factors must "take up the slack" by having lower survival, so that the historical declines of salmon are accurately simulated. Therefore, an alternative hypothesis which postulated higher system survival, would need to also propose either lower stock productivity, lower post-Bonneville survival, and / or lower survival through short-term climate fluctuations in order to generate the observed year-to-year changes in recruitment.

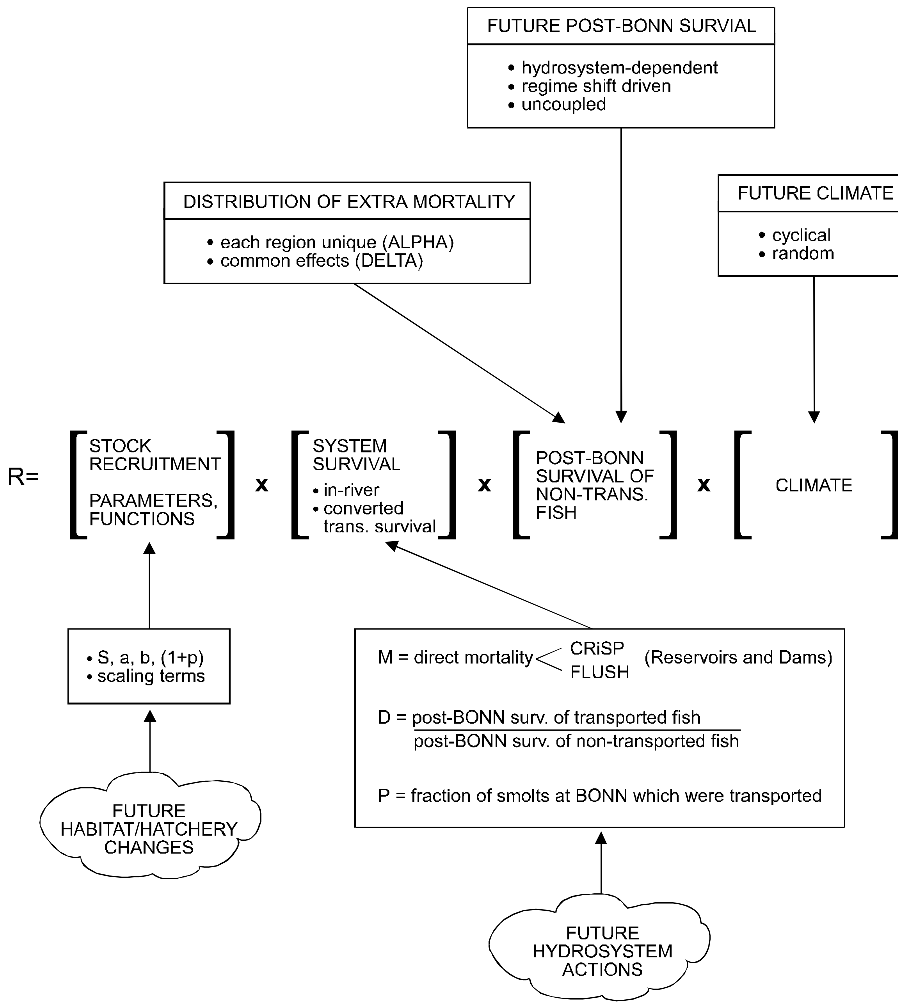


Figure 4.3-1: General structure of the life-cycle model used to integrate alternative hypotheses for all components. Further management actions shown in clouds at bottom. Changes to the schedule of in-river harvests (the only harvest for spring / summer chinoo) do not affect survival of recruits to mouth of Columbia River (R), but do affect number of returning spawners (S) Changes to hatchery operations are not currently modeled.

4.3.1 Transportation Assumptions

Transportation Rules

The proportion of fish transported in the prospective models is determined by the fish guidance efficiencies used in the passage models and the rules for spill and collection under various flows. For scenario **A1**, these rules are based on the seasonal average (April 10-June 20) flows; one of three cases is applied depending on the amount of flow relative to thresholds of 85 kcfs and 100 kcfs (see Section A.3.1 for details). In scenario **A2,** all smolts collected at LGR\LGO\LMO, and MCN Dams are transported (see Table 4.2-1 for abbreviations). There is no voluntary spill at collector projects. In scenario **A3,** no smolts are transported.

Relative Survival of Transported Fish After Bonneville Dam

Uncertainties related to transportation are focussed not on direct survival in the barge (V_t), which is acknowledged to be relatively high (around 98%), but on the relative survival of transported and non-transported juvenile fish between the time they are released below Bonneville Dam and the time they return to Bonneville Dam as adults (Figure 4.2-1).

The alternative "TRANS" hypotheses(i.e. T1, T2, and T3) considered here are different ways to use the available T:C studies to estimate the relative survival of transported fish after Bonneville Dam ("D"; defined in Figure 4.2-1). These alternative hypotheses have been implemented only partially in our current results (i.e., T1 and T2, are used only in combination with FLUSH estimates of in-river survival; T3 is only combined with CRiSP estimates of in-river survival). We intend to conduct sensitivity analyses which explore other combinations (i.e., T1 and T2 with CRiSP; T3 with FLUSH) to understand the importance of transportation hypotheses relative to passage model selection in affecting performance measures.

The effect of these alternative hypotheses is shown in Figure 4.3-2. Because of the importance of this factor, we provide more details in this section of the report below.

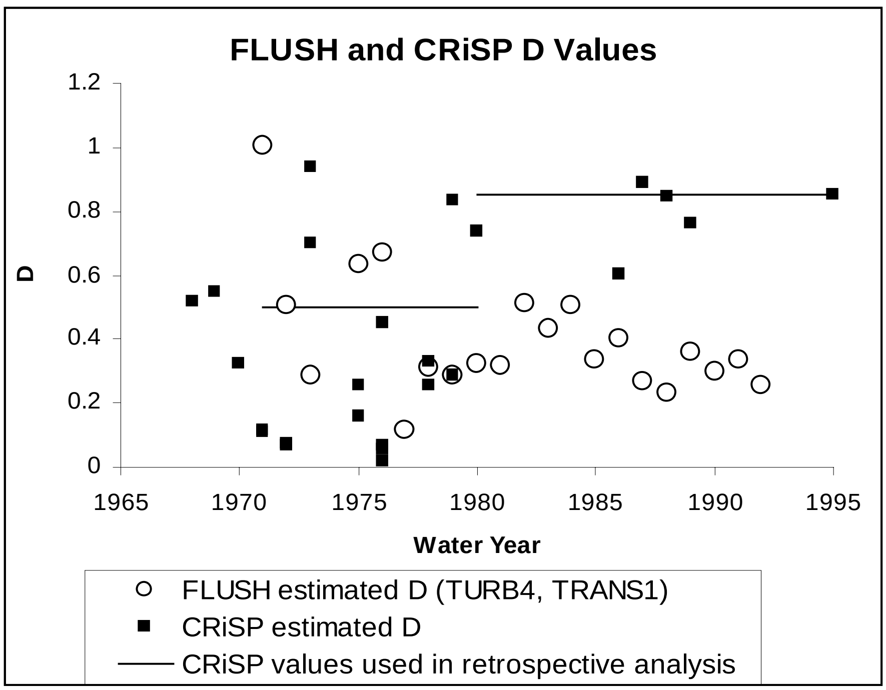


Figure 4.3-2: Comparison of D values estimated from T:C studies and passage models. D is the ratio: (post-BONN survival of transported fish) / (post-BONN survival of non-transported fish). See Figure 4.2-1.

T1and T2 (implemented only with FLUSH)

The ratio of post-Bonneville survivals of transported (8_t) and non transported (8_n) fish is represented by

$$D = (8_t / 8_n) \qquad \text{[Eq. 4.3.1-1]}$$

In T1, retrospective values for *D* are computed for brood year *y* by:

$$D_y = (T/C)_{est} * V_{n,y} / V_{t,y} \qquad \text{[Eq. 4.3.1-2]}$$

where V_n is the estimated survival from the head of Lower Granite to below Bonneville, for non-transported fish. $(T/C)_{est}$ is the T:C ratio estimated from a function in which $(T/C)_{est}$ is inversely related to V_n, and always >1 (see Table 4.3-1, Appendix A.3.1). This function is based on data from all available transport studies conducted at Lower Granite (LGR) and Little Goose (LGO) dams between 1971-1989. A second transport model (T2) adjusts the T/C estimates of transport model 1 by a factor of 0.83 to account for possible differences in T:C values measured at mainstem dams (where most adult recoveries take place) and natal areas (spawning grounds and hatcheries). Under both T1 and T2, prospective values for (T/C) are computed from the (T/C) vs. in-river survival (V_n) function, and the V_n for the simulated future year.

Table 4.3-1: D values estimated by FLUSH.

Water Year	FLUSH D
1971	1.00
1972	0.504
1973	0.285
1975	0.637
1976	0.671
1977	0.116
1978	0.309
1979	0.288
1980	0.321
1981	0.316
1982	0.511
1983	0.434
1984	0.503
1985	0.336
1986	0.402
1987	0.266
1988	0.233
1989	0.357
1990	0.297
1991	0.336
1992	0.254
pre-1980 average	0.476
1980-1992 average	0.351

T3 (implemented only with CRiSP)

In T3, retrospective *D* values were computed from the measured (T/C) for each year with a transportation study (Table 4.3-2):

$$D_y = (T/C)_y * V_{n,y} / V_{t,y} \qquad \text{[Eq. 4.3.1-3]}$$

A D_y value of 0.5 was used for years before 1980, and a value of 0.85 was used for years after 1980. Note that Equation [4.3.1-3] is like Equation [4.3.1-2] except that measured (T/C) values are used instead of estimated ones. For prospective runs, T3 uses:

$$(T/C)_{future} = D_{random} * V_t / V_n \qquad \text{[Eq. 4.3.1-4]}$$

where D_{random} is a randomly selected D_y value from the set of retrospective estimates after 1980.

The CRiSP team is now considering fitting a regression line through D_y to reflect gradual improvements in D_y since 1975, and relating D to estimates of descaling.

Table 4.3-2: CriSP-T3 estimates of D.

Water Year	CRiSP D
1968	0.518
1969	0.548
1970	0.320
1971	0.109
1971	0.118
1972	0.070
1972	0.073
1973	0.702
1973	0.938
1975	0.157
1975	0.254
1976	0.069
1976	0.452
1976	0.018
1976	0.052
1978	0.257
1978	0.326
1979	0.287
1979	0.836 (a)
1980	0.737 (a)
1986	0.603
1987	0.892 (a)
1988	0.847 (a)
1989	0.760
1995	0.852
pre-1980 average	0.321
1980-1995 average	0.782

(a) Estimates based on transport studies at McNary Dam.

It can be seen from the above equations that (T/C) is inversely related to V_n in all the three TRANS hypotheses. In other words, years with lower in-river survival have a higher T:C ratio (and vice versa). Differences in future (T/C) values are related to:

1. Differences between CRiSP and FLUSH in retrospective values for V_n, which affect the estimated historical D values;
2. Differences in the set of years of T/C information used to estimate future T/C's and corresponding *D*'s (i.e., either 1971-1989 (T1 and T2), or just the post-1980 years with transportation studies (T3); and
3. Differences in the smoothing procedure used (i.e., smoothing (T/C) estimates (T1 and T2) or smoothing (averaging or regressing) the *D* estimates (T3).

4.3.2 Stock Productivity

Productivity of spring/summer chinook populations (i.e., the number of recruits per spawner at all spawner densities except very low ones) is quantified based on a generalized Ricker spawner-recruitment model. In some salmon populations, the number of recruits per spawner declines as spawner abundance declines, which is called **depensatory**. The potential for this behavior is of concern because it allows populations to go extinct more rapidly than otherwise expected. The unknown relationship between spawners and resultant recruits at low spawner abundance is a significant source of uncertainty in productivity. This relationship is poorly understood because we have little experience with populations at these low abundances. Though there is no clear evidence for depensatory behavior in these stocks, it is difficult to detect such patterns in data with ageing errors. We account for the possibility of depensatory behavior by including it as a source of uncertainty in the relationship between spawners and recruits, but only insofar as it was evident in the data up to brood year 1990. These data do not show significant evidence of depensation, but more recent brood years (with lower escapements) have not yet been incorporated into the analysis. Further sensitivity analyses to depensation assumptions will be presented in our final report.

The Ricker spawner-recruitment model estimates sub-basin specific rates of intrinsic productivity and population carrying capacity. The prospective analyses include a stochastic relationship between spawners and resultant recruits and admission of uncertainty about fundamental parameters governing modeled productivity. Alternative hypotheses (Section 4.3.5) consider potential changes in each stock's productivity.

There are two alternative representations of chinook population dynamics (the Alpha model and the Delta model). The Alpha and Delta models are described fully in Appendix A.3.2. Briefly, the Delta model is an extension of the model used in Chapter 5 of the PATH FY96 Retrospective Analysis (Deriso et al. 1996). Deriso et al. used spawner-recruit data from Snake River and lower Columbia River stocks to infer both common-year effects due to climate affecting all stocks, and a combined 'direct plus extra' mortality. The prospective Delta model partitions out the direct and extra mortality components by using a passage model for the direct component, but keeps the common year effects as a separate term.

That is, climate effects are separated from both direct and extra mortality. The Alpha model also uses a passage model for the direct component, but does not estimate common-year effects based on similarities between Snake River and lower Columbia River stocks. Rather, the Alpha model treats each stock group independently, with an extra mortality specific to each group that includes both climate effects and any delayed effects of the hydrosystem. Thus climate effects are part of the extra mortality.

The Alpha and Delta models produce comparable rates of average productivity for the 1951-1990 brood years examined (Figure 4.3-3). There were differences, however, between these models for estimates of maximum potential productivity. Maximum potential productivity values are generally larger in the Delta model. However, when adjustments are made to the productivity measure to account for model-specific estimates of post-Bonneville "extra mortality" for non-transported smolts, no single model produces consistently higher estimates of adjusted productivity.

In the final report, we will also compute probabilities of quasi-extinction (see Chapter 3) of this report. Quasi-extinction is simply an abundance of spawners which, for modeling purposes, you assume is equivalent to extinction. One records the number of simulations where this event occurs, along with 'the first crossing time' (i.e. number of years from the start of the simulation until the event occurs).

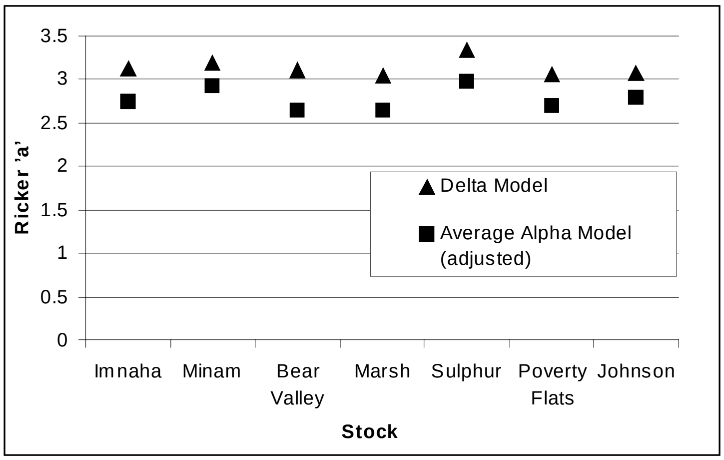


Figure 4.3-3: Comparison of stock productivities as estimated by the Delta and Alpha prospective models.

4.3.3 Extra (Post-BONN) Mortality

Extra mortality is mortality occurring outside of the migration corridor that is not captured by the inherent stock-recruitment parameters for a given stock (i.e., productivity, depensation, and carrying capacity terms). Three alternative hypotheses for extra mortality were considered.

a) Hydro-related

The completion of the Federal Columbia River Power System in the late 1960s through the mid-1970s and its subsequent operation, have increased the direct and delayed mortality of juvenile migrants, resulting in considerably sharper declines in survival rates of Snake River spring and summer chinook stocks (over the same time period), than of similar stocks which migrate past fewer dams and are not transported. This hypothesis follows from Conclusion 3a.2 of the PATH FY96 Conclusions Document:

> *We are reasonably confident that the aggregate effects of the hydrosystem have contributed to reduced survival rates of Snake River stocks (from spawners to adults returning to the mouth of the Columbia River), during the post-1974 period, as compared to the pre-1970 period. Hydrosystem effects include both direct (e.g., turbine mortality) and indirect effects (e.g.,*

delayed mortality, due to such mechanisms as changes in estuary arrival times).

b) "BKD" or Stock Viability Hypothesis

The hypothesis proposes that: 1) the viability of Snake River stocks declined as a direct or indirect result of the hydrosystem construction in the 1970s; 2) current extra mortality is not related to either the hydropower system or climate conditions; and 3) extra mortality is here to stay, even if hydrosystem direct mortality is reduced and / or the climate improves. One hypothesis to account for decreased stock viability is that hatchery programs implemented after construction of the Snake River dams increased either the incidence in the level of bacterial kidney disease (BKD) within the wild population or its severity. In both cases, the mortality increased in juvenile fish after they exited the hydropower system as compared to earlier years (or as compared to downstream stocks for the same time period). Under this hypothesis, it is unlikely that the increased rate of mortality from BKD would change back to a more favorable condition in the near future. Another stock viability hypothesis is that low stock sizes have led to increased predation rates on juveniles, and insufficient nutrients from returning adults' carcasses to support the growth of parr.

BKD is only one possible means by which stock viability may have been reduced. Occasional changes in underlying stock viability may cause some or all of the delayed mortality to remain, even if direct mortality is reduced. The consequence of falling into this category (i.e., "delayed mortality is here to stay") is that it is unknown when or if the impacts will switch back to a less benign state. For modeling purposes, we consider this the worst case, which is that these factors will stay in the present less favorable state.

c) "Regime shift" Hypothesis

Extra mortality is not related to the hydropower system, but is due instead to an interaction with a long term cyclical climate regime shift with a period of 60 years. This regime is believed to have shifted from good to poor during brood year 1975, and is expected to return to above average conditions in 2005. The signatures of a recurring pattern of interdecadal climate variability are widespread and detectable in a variety of Pacific basin climate and ecological systems. These cyclical changes affect ocean temperatures and currents which affect distributions of predators and prey; and broad scale weather patterns over land masses which then affect temperatures, rainfall, snowpacks, and subsequent flows. The changes in conditions could affect various stocks to different degrees with the effect on Snake River stocks being systematically different from lower river stocks. There is nothing that we can do to change these patterns, but they are expected over time to provide more favorable and less favorable conditions to species located in different areas.

4.3.4 Future Climate Conditions

Since climate is a factor that is beyond human control, modeling future climate conditions is necessary to explore the sensitivity of the performance of alternative management actions to good and bad climate scenarios.

Three climate hypotheses were considered in the preliminary analysis (details on the mathematical implementation of these hypotheses are included in Section A.3.4). One hypothesis assumes that climate patterns in the Northeast Pacific follow a cyclical pattern, with changes between good (i.e., increase in productivity) and bad (decrease in productivity) ocean conditions occurring on an 18.5 year cycle. This hypothesis is based on observed historical patterns in various indices of ocean productivity; based on these patterns the last change from good to bad conditions is assumed to have occurred around 1995 in one prospective model (Alpha model; see Section A.3.2) and 1991 in the other prospective model (Delta model), while the next change from bad to good conditions is expected in approximately 2005 in the Alpha model and 2001 in the Delta model. Another hypothesis does not explicitly model an underlying decadal trend. Instead, it samples from estimated climate effects during the period 1950 to 1995, a period which includes both good and bad climatic conditions as measured by various ocean productivity indices. Sampling from this period is done such that good years tend to follow good years and bad years tend to follow bad years (this method is called 'first order autocorellation'). Modeling climate in this way does tend to produce variable cyclical patterns even though no explicit assumptions are made about underlying decadal-scale trends. Finally, a larger-scale climate hypothesis involves a 60-year cycle; this climatic pattern is associated with the "Regime shift" hypothesis for extra mortality discussed in the previous section, and may be superimposed on the shorter period cycles described above. These methods are explained in more detail in Appendix A, section A.3.3 and A.3.4.

The aggregate effects of the FLUSH-T1/T2 passage and transportation models, and the Delta approach to climate effects are shown in Figure 4.3-3 for the historical period. Good climate years help to reduce overall mortality, while poor climate years (most years in Figure 4.3-3) increase it.

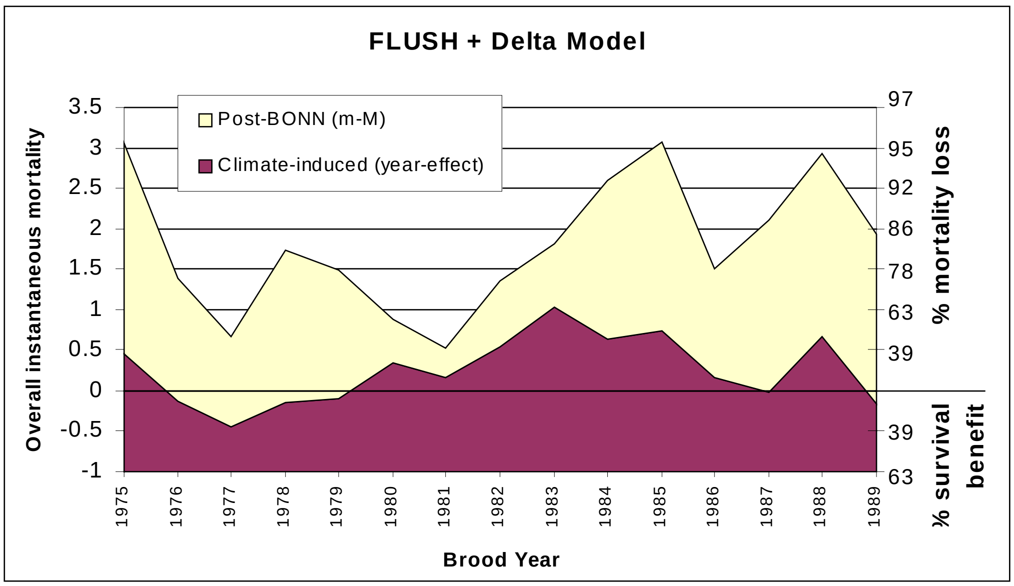


Figure 4.3-4: Comparison of historical estimates of post-Bonneville mortality and climate-related mortality as estimated by FLUSH-T1/T2 passage and Delta prospective models. Years with good climate have less dark area (climate-induced mortality).

< not ready for this draft >

Figure 4.3-5: Comparison of historical estimates of post-Bonneville mortality and climate-related mortality as estimated by CRiSP-T3 passage -transportation assumptions and Alpha prospective models.

4.3.5 Habitat

Future changes to the freshwater spawning and rearing habitats of chinook salmon may have an important influence on stock recovery. While few would disagree that freshwater habitat can be a critical limiting factor or that changes in land use can affect habitat quality and survival, it has not been possible to statistically demonstrate the effects of any given set of habitat conditions or indicators on stock productivity (Appendix A3.5). Instead we have relied upon expert judgments of plausible changes in stock productivity due to habitat management and rehabilitation. The purpose of the habitat approach is to look at how these various assumptions about future spawning and rearing habitat conditions would affect the results of the various hydrosystem management actions. This analysis does not consider the impacts of changes in fish habitat other than for the spawning and rearing part of the life cycle.

Habitat effects were defined in terms of changes to the Ricker *a* parameter, which is a measure of stock productivity at low levels of abundance and should reflect habitat quality. Plausible changes in *a* for each population within a 48-year period (the NMFS recovery time period) were defined as plus or minus 1 unit which is approximately equivalent to a three-fold change in stock productivity (recruits per spawner) or egg-to-smolt survival rates. This range is based on observed ranges of: 1) *a*-values; 2) PIT tag recovery rates; and 3) subbasin planning model smolt production assumptions. To avoid unreasonably high predicted *a* values for some stocks the maximum upward change was constrained to not exceed the highest observed value among all stocks, which was assumed to represent an upper limit for the intrinsic productivity of stocks in an area as determined by physiography and climate.

Using expert judgment we defined the probabilities of: 1) no change; 2) an increase; or 3) a decrease in Ricker *a* values for each population relative to their values estimated from 1952-1990 spawner-recruit data. We developed probability tables for two habitat management options (Table A.3.5-1): (A) status quo habitat management (in which some Ricker *a* values can increase or decrease) and (B) maximum practical effort with respect to habitat protection and restoration (i.e., greater chance of increased *a*-values, and lower chance of declines). The results presented in Section 5 contrast option B to a scenario in which Ricker *a* values did not change from their estimated values. The probability judgments were based on qualitative descriptions of habitat conditions currently affecting each population and habitat group's assessment of the potential for improvement or degradation. Productivity in a relatively pristine habitat was generally deemed unlikely to increase, so a low probability was assigned to the increased *a*-value and a high probability assigned to the no-change option. Conversely, probabilities of increased productivity (higher *a*-value) were greater in habitats degraded by past land use practices. Expert opinion was also used to judge how rapidly changes are likely to occur by assigning probabilities that changes to the higher or lower value would occur by simulation years 12 and 24 given that the change was expected by year 48. The results of the expert elicitation are described in more detail in Appendix A, Section A.3.5.

4.3.6 Hatcheries

As a response to Snake River hydrosystem development, hatchery releases of anadromous fish have increased, both from traditional hatcheries and subbasin supplementation operations (Table A.3.6-1). There are a number of reasons to believe that these releases may have a negative effect on ESA stock survival. These include transmission of disease, increased competition for food and other resources, and increasing available prey for subbasin and mainstem predators (see Chapter 11 of the FY1996 PATH Retrospective Report). These effects could manifest themselves either for stock groups as a whole (e.g., for all Snake index stocks) or at the level of individual index stocks where supplementation releases have occurred. On the other hand, supplementation proponents believe that releases have positive demographic effects by enhancing abundance of returning adults to seed future generations. Although many studies of hatchery-wild salmonid interaction have been conducted, only a handful of studies actually address the effects of hatchery releases on wild stock spawner-recruit survival. Some of these were summarized in Chapter 11 of the PATH FY96 Retrospective Report. We intend to consider hatchery hypotheses in our next report.

4.3.7 Harvest and Upstream Passage

The primary harvest of Snake River spring and summer chinook occurs within the Columbia River, very little of the harvest on these stocks occurs in ocean fisheries. Harvest levels have been significantly restricted since the late 1970s to below 15%. Currently, mainstem harvest of Columbia River spring and summer chinook stocks are managed according to schedules relating annual allowable harvest rates to estimated return levels. The current harvest schedules explicitly address harvest within the range of run strengths observed in recent years and are in place through 1998 (harvest BioOp agreement reference???). For the purposes of prospective modeling exercises, the harvest rate schedule was expanded to cover higher run sizes and tributary harvest. The rules applied at higher run levels reflect the rules included in the Columbia River Fish Management Plan (CRFMP) and historical fishing levels in the tributaries. The allocation of harvest between mainstem and tributaries resulting from application of these rules is essentially arbitrary. In practice, a different mix of the total allowable harvest rate between mainstem and tributary may occur for any indicator stock in any particular year. However, the total harvest rate impacts reflect the intent of the CRFMP. Harvest policies will be reviewed as part of the U.S. vs. Oregon negotiation process for application to 1999 and beyond.

Mainstem harvest rates on spring and summer chinook respectively are determined by the expected Snake River wild returns using harvest rate schedules. The indicator stocks comprise a portion of the total production from Snake River tributaries. For application within the modeling exercise, the run strength parameters determining the applicable harvest rate have been transformed from aggregate Snake River run size into the percent of escapement needed to achieve MSP (escapement level needed to achieve **m**aximum **s**ustained **p**roduction). Transformed harvest schedules for use in the prospective analyses are provided in the following two tables. For any given year in the prospective analysis, the appropriate mainstem harvest rate is obtained through a two step process. The run size projections to the Columbia River mouth are summed across indicator stocks and that total is expressed as a percentage of the sum of the MSP's for those same stocks. A mainstem harvest rate corresponding to that %MSP is applied to each of the individual indicator runs in the simulation. A tributary harvest rate corresponding to the average %MSP is subsequently applied to each of the indicator stocks. A primary assumption of this approach is that variability between actual run strength and the estimated run size used to set harvest rates within a particular year does not have a significant effect on harvest level.

Alternative harvest schedules were developed for the purpose of assessing the impact of variations in the harvest schedule within the prospective analysis (Tables 4.3.7-1 and 4.3.7-2). Given the starting population levels for the indicator stocks, variations in harvest rates at the low to moderate run strengths are the most likely to effect rebuilding. A more conservative harvest schedule was generated by dividing the harvest rates for run strengths below MSP by 1.5. A higher allowable harvest rate alternative was developed by multiplying the same set of rates by 1.5. This approach was applied to both the spring and the summer chinook harvest schedules.

Table 4.3.7-1: Upriver Spring chinook CRFMP harvest rate schedule to be implemented in BSM (Bayesian Simulation Model used to simulate overall life cycle changes)

	Existing Harvest Management		Conservative Harvest Management		Liberalized Harvest Management	
Run Size % of MSP /a /b	C.R. Mainstem Harvest Rate	Tributary Harvest Rate	C.R. Mainstem Harvest Rate	Tributary Harvest Rate	C.R. Mainstem Harvest Rate	Tributary Harvest Rate
< 22%	0.055	0	0.037	0	0.083	0
22%-44%	0.082	0	0.055	0	0.123	0
45%-112%	0.14	0	0.093	0	0.210	0
113%-125%	0.25	0.05	0.25	0.05	0.25	0.05
126%-175%	0.3	0.15	0.3	0.15	0.3	0.15
176%-200%	0.35	0.2	0.35	0.2	0.35	0.2
>200%	0.4	0.25	0.4	0.25	0.4	0.25

a/ run size adjusted for 77-90 average adult passage conversion and 90% prespawning survival
b/ average % of MSP for index stocks

Table 4.3.7-2: Upriver Summer chinook CRFMP harvest rate schedule to be implemented in BSM

	Existing Harvest Management		Conservative Harvest Management		Liberalized Harvest Management	
Run Size % of MSP /a /b	C.R. Mainstem Harvest Rate	Tributary Harvest Rate	C.R. Mainstem Harvest Rate	Tributary Harvest Rate	C.R. Mainstem Harvest Rate	Tributary Harvest Rate
< 25%	0.02	0	0.013	0	0.03	0
25%-49%	0.05	0	0.033	0	0.08	0
50%-99%	0.1	0	0.067	0	0.15	0
100%-129%	0.15	0	0.15	0	0.15	0
130%-149%	0.2	0.05	0.2	0.05	0.2	0.05
150%-169%	0.25	0.1	0.25	0.1	0.25	0.1
170%-200%	0.3	0.2	0.3	0.2	0.3	0.2
>200%	0.35	0.25	0.35	0.25	0.35	0.25

a/ run size adjusted for 77-90 average adult passage conversion and 90% prespawning survival

4.3.8 Other Factors Not Modeled

Any analysis must focus on a particular body or type of information and cannot account for all factors that may bear on decisions. In particular, a quantitative analysis such as this focuses on those issues that can be quantified on the basis of existing information. These missing factors need to be recognized to place the results of the biological analyses in the context of all scientific, social, and economic information that may bear upon important decisions.

The present analysis focuses on those aspects that can be quantified within a stock recruitment type of relationship. However, this does not fully account for all factors that are known to, or have been suggested to, affect the production and success of salmon in the Columbia River Basin. Information that is presently less easily quantifiable may, however, be equally important from a scientific perspective. Examples of factors that have not been included in the quantitative analysis are discussed below. We hope to deal with some of the "unmodeled" factors in our final report.

1. Individual Populations Geographically Isolated

In the modeling of future conditions, each population is treated as an independent entity and isolated from other similar populations. For example, fish do not stray from one population to another. Except to the extent they are affected by common factors, individual populations decline or increase independently. A counter point to this is that though interactions between populations are not explicitly modeled, the effects of population interactions are captured implicitly in historic stock-recruitment data and the fitted parameters of stock-recruitment functions.

In nature, interactions likely do occur between populations in regard to both behavior and productivity. Populations overlap at different life stages affecting dispersion of individuals and perhaps carrying capacity. For example, a population of salmon may spawn in habitat that is considered ideal, but disperse downstream where they intermingle with other populations. These downstream areas may be affected by habitat degredation and have limited capacity as a result. While at the present low levels of abundance such interactions may be less important, within the historical record they could have been important and could be in the future as populations rebuild. In particular, some habitats and environmental conditions could result in low capacities for some life stages. This could affect dispersal, distribution, and survival of populations.

At a level above that of individual populations, the present analysis addresses races (for example, spring chinook) in isolation from other races and species. Regional decisions, however, will be made in the context of their impacts on other races, as well as other salmon, resident fish and wildlife. It is quite possible that assessments of major actions could be different if considered across the spectrum of affected species, races, and populations, rather than for each biological group in isolation. Fall and spring chinook, for example, utilize the mainstem Columbia and Snake rivers at markedly different points in their life cycle. Evaluating management options from the perspective of either group in isolation (perhaps because the biology of one or the other is more easily quantifiable from existing information), may result in a different prioritization of actions than if the actions were evaluating in a broader context of the species (or genus) collectively in an ecosystem context. The final PATH decision analysis will jointly consider the impacts of management actions on spring / summer chinook, fall chinook, and steelhead.

2. Populations are Genetically Isolated

The analysis treats populations as genetic isolates as well. No emigration between populations is explicitly considered. Salmon populations, however, are likely structured into some higher organization reflecting differing degrees of genetic communication. Within this organization, populations may vary asynchronously as a result of local as well as regional conditions. Genetic communication (gene flow) between populations is balanced with selection for local conditions. This organization likely develops as an adaptive trait in response to environmental variation. As such, it would have an impact on survival, rebuilding and sustainability of populations that is not considered in the analysis.

In this analysis, each population declines or increases in isolation from other populations. In reality, conditions may result in increases in some populations and emigration into other populations. The effect of declines in some populations due to local catastrophes or environmental variation may be dampened by emigration from stronger populations. However, a counter argument to this criticism is that in the Snake

River, all naturally spawning spring/summer chinook populations are weak, and varying in synchrony, (i.e., there are no strong populations to serve as effective sources to mitigate the effects of local catastrophes).

3. Populations are Behaviorally Similar for Parts of their Life Cycle

Populations vary within the analysis in regard to basic productivity (Ricker *a* and *b* values) and in regard to their position relative to the number of dams encountered moving to and from the ocean. Fish move through the hydroelectric system at similar times encountering similar conditions and reacting in similar manners. Differences among stocks in basic productivity parameters may reflect differences at any point during the life cycle. Two of the key hypotheses considered in the decision analysis are: 1) whether or not there are common year effects between upriver and downriver stocks; and 2) whether differences in recruitment between upriver and downriver stocks reflect exposure to hydrosystem effects, or differences in marine survival. These alternatives are reflected in the delta and alpha models, and the alternative hypotheses about post-Bonneville survival.

Information upon which to test these alternative hypotheses is limited. The hypothesis of broadscale, regional differences in marine migration and survival of spring chinook, for example, is neither supported nor rejected by the very limited coded wire tag data available from ocean recoveries of hatchery fish. Nonetheless, based on differences observed in some races for which data exist (e.g., Columbia River fall chinook), there is reason at least to carry these alternative hypotheses through the analyses.

4. The Mainstem River is Treated Largely as a Migrational Corridor Affected only by Flow

The mainstem Columbia and Snake rivers are treated largely as reservoirs affected only by flow rate. Quantity and quality of mainstem river habitat for juvenile rearing and migration are not considered. Factors that form riverine environments include geomorphology, physical (and biological) structure, and hydrology. The interaction of these to form salmonid habitat and structure salmonid populations are not considered. The impact of simplification of the mainstem and elimination of habitat is not included except as it is embedded in the historical data. Consideration of future conditions in the mainstem rivers resulting from changed management practices is limited to its effect on fish travel time and migrational survival.

There are legitimate counter-arguments to the above criticisms. First, the key factors for successful mainstem migration are better known for spring/summer chinook than for other species/races. They are passive migrants which historically took less than two weeks to migrate the entire distance. The net effect of any uncertainty of the importance of the mainstem area for spring chinook is much more likely to result in an underestimate of the benefits of a natural river system on chinook survival than in an overestimate.

5. The Impacts of Hydroelectric Operations on the Estuary are not Considered

Within the analysis, changes in hydroelectric operations are restricted to their impacts on flows and velocities within the area above Bonneville Dam. However, there is substantial reason to believe that hydroelectric operations have an appreciable effect on the estuary as well and, therefore, on fish survival. Flow management and the impact of dams on downstream movement of material are known to have changed the timing and magnitude of the spring freshet and the type and quantity of organic and inorganic material delivered to the estuary. The shape and dynamics of the ocean plume are known to have changed as the river has been developed. Biological linkages, especially to salmon production and survival, are difficult to demonstrate, in part because little effort has been directed at this topic. However, because the estuary represents a critical transitional stage for both adult and juvenile salmonids, it is reasonable to suspect a linkage between river development and operations and salmon production. These impacts are not considered in the present analysis because of the lack of quantitative relationships between estuarine physical conditions and salmon. However, some of the empirical relationships incorporated into the models (e.g. common year effects on upstream and downstream stocks in the Delta model, dependence of extra mortality on mainstem flow in the Alpha model) may implicitly capture such effects.

5 Results

5.1 Introduction

PATH analyses to date have made considerable progress towards defining alternative hypotheses about key elements of life-cycle survival and identifying the critical differences between alternatives. This advance in understanding and clarification of differences are described in Chapter 4 and Appendix A of this report.

In this section, we focus on the implications of these alternative hypotheses for decision-making. The objectives of this section are to:

1. Explore ways to summarize complex analyses and results into graphs that are easy to understand, interpret, and explain to decision-makers (Section 5.2).
2. Provide **preliminary** insights into the relative performance of alternative actions (Section 5.3), recognizing that we have only explored a partial set of actions and hypotheses, and that further refinements in our methods are likely to occur.
3. Identify the uncertainties that are most important to assign weights to (if possible) and/or resolve through either continued monitoring and research or deliberate adaptive management experiments (Sections 5.4 and 5.5).
4. Test the sensitivity of decisions to the weights that are placed on key uncertainties (Section 5.6), so as to focus the assessment of existing evidence, and the acquisition of additional evidence.
5. Summarize some other important performance measures (Section 5.7).

5.2 Ways to Summarize Results

We have generated predicted outcomes for alternative management actions (A1, A2, and A3; see Chapter 2 for a description) using each possible combination of models and hypotheses described in Section 4. Each one of these combinations is called an "aggregate hypothesis" or a "run". Because there are 5,148 unique aggregate hypotheses, there are 5,148 unique alternative futures that one could examine to evaluate alternative actions. This range of outcomes reflects the uncertainty associated with predicting future events from imperfect or incomplete information. We have attempted to boil down this information to reveal the critical uncertainties. Such summarization will be particularly necessary in the final decision analysis report that considers additional actions and species.

To illustrate some of the concepts that follow, example output for 20 out of the 5,148 aggregate hypotheses analyzed is shown in Table 5.2-1. The values shown in Columns 2 to 4 are the probability that the number of spawners for the sixth best stock will exceed a pre-defined "survival" level of spawning abundance for that stock, projected over the next 24 years (this is the 24-year NMFS survival standard; see Chapter 3). Higher numbers indicate larger projected spawning abundances. The 24-year survival standard was selected for this illustration arbitrarily; readers should be aware that there are several different jeopardy standards and performance measures that could be used (see Chapter 3 and Appendix D).

One way to select an action on the basis of the information in Table 5.2-1 would be to simply assume that the "correct" aggregate hypothesis is known with complete certainty, and then select an action based on the outcomes from this single aggregate hypothesis. For example, suppose we assume that aggregate hypothesis #9 is correct. The probability of exceeding the survival escapement level under this aggregate hypothesis is 0.82 for A1, 0.68 for A2, and 0.63 for A3. If we apply the informal criterion that this probability must exceed 0.70 for an action to be considered acceptable (see Chapter 3 for a description of the NMFS jeopardy standards), then the only action that is acceptable is A1, and that is the action that should be taken **(Figure 5.2-1)**.

Table 5.2-1: Example outputs. Probabilities of exceeding Survival escapement level over 24 years, under three management actions, for 20 randomly selected aggregate hypotheses. Weights were assigned randomly for illustrative purposes.

	Probability that the number of spawners exceed the survival escapement level			Is probability greater than 0.7? 1 if yes, 0 if no			
Aggregate Hypothesis	**A1**	**A2**	**A3**	**A1**	**A2**	**A3**	**Weight**
1	0.35	0.65	0.45	0	0	0	0.05
2	0.35	0.65	0.47	0	0	0	0.08
3	0.84	0.85	0.78	1	1	1	0.09
4	0.78	0.85	0.78	1	1	1	0.08
5	0.33	0.63	0.46	0	0	0	0.01
6	0.32	0.65	0.47	0	0	0	0.02
7	0.84	0.84	0.78	1	1	1	0.07
8	0.74	0.85	0.78	1	1	1	0.08
9	0.82	0.68	0.63	1	0	0	0.05
10	0.83	0.71	0.63	1	1	0	0.07
11	0.51	0.53	0.67	0	0	0	0.05
12	0.51	0.53	0.67	0	0	0	0.06
13	0.66	0.60	0.75	0	0	1	0.01
14	0.67	0.61	0.76	0	0	1	0.00
15	0.53	0.52	0.66	0	0	0	0.07
16	0.53	0.52	0.66	0	0	0	0.07
17	0.67	0.60	0.75	0	0	1	0.06
18	0.68	0.60	0.74	0	0	1	0.04
19	0.53	0.50	0.67	0	0	0	0.03
20	0.53	0.51	0.67	0	0	0	0.01
# of aggregate hypotheses resulting in probability >0.7				6	5	8	
Proportion of aggregate hypotheses resulting in probability >0.7				0.3	0.25	0.4	
Expected ability to meet 0.7 standard				0.44	0.39	0.43	

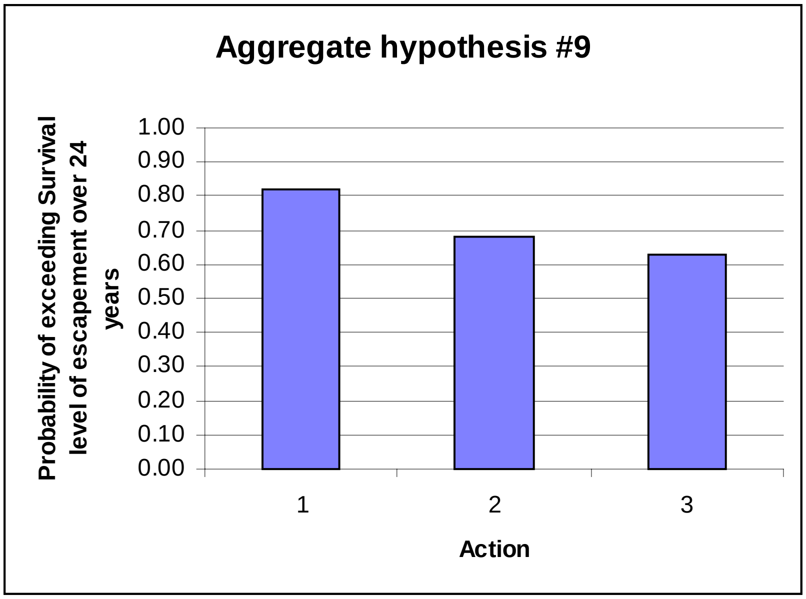


Figure 5.2-1: Display of example results for aggregate hypothesis #9 in Table 5.2-1.

The problem with this approach is that we don't know for certain which aggregate hypothesis is correct, so we have no way of selecting a single hypothesis. Moreover, different aggregate hypotheses can lead to different conclusions. For example, A3 is the only acceptable action under aggregate hypothesis #18 (i.e., the only action with a probability greater than 0.7), while no actions are acceptable under aggregate hypothesis #1 (Table 5.2-1). Therefore, the choice of a particular action will depend on which of the 5,148 aggregate hypothesis is assumed to be correct.

What we need is a way to summarize the results for all aggregate hypotheses in a way that is easy to understand and accurately captures the entire range of outcomes that are possible. There are at least three potential approaches to do this:

1. a frequency distribution of outcomes,
2. calculate the fraction of aggregate hypotheses that meet some criterion
3. calculate the expected ability of an action to meet some criterion

Brief explanations of these approaches are provided below for readers who are unfamiliar with basic principles of probabilities and probability distributions.

5.2.1 Frequency Distribution of Outcomes

One approach is to count the number of aggregate hypotheses that produce an outcome in a given range (or "bin") for a given action. This can be done for a range of bins that covers the entire range of outcomes for that action and shown as a bar chart, which is called a frequency distribution. An example frequency distribution is shown in Figure 5.2-2, based on the outcomes in Table 5.2-1. There are two main features of this frequency distribution. First, the height of the bars reflects the relative frequency with which an outcome in a particular bin is produced. In the example, A1 outcomes in the 0.51 to 0.60 bin are produced with greater frequency than those in other bins. Second, the frequencies show the maximum and minimum limits on the range of outcomes, which provides an indication of the amount of uncertainty in the outcomes. In the example frequency distribution for A1, there are no values less than 0.31 and no outcomes greater than 0.90. Hence, in the four most pessimistic aggregate hypotheses (worst-case scenario), action A1 has only a 0.31 to 0.4 chance of exceeding the pre-defined spawning abundance threshold. The four most optimistic aggregate hypotheses have a 0.81 to 0.9 chance.

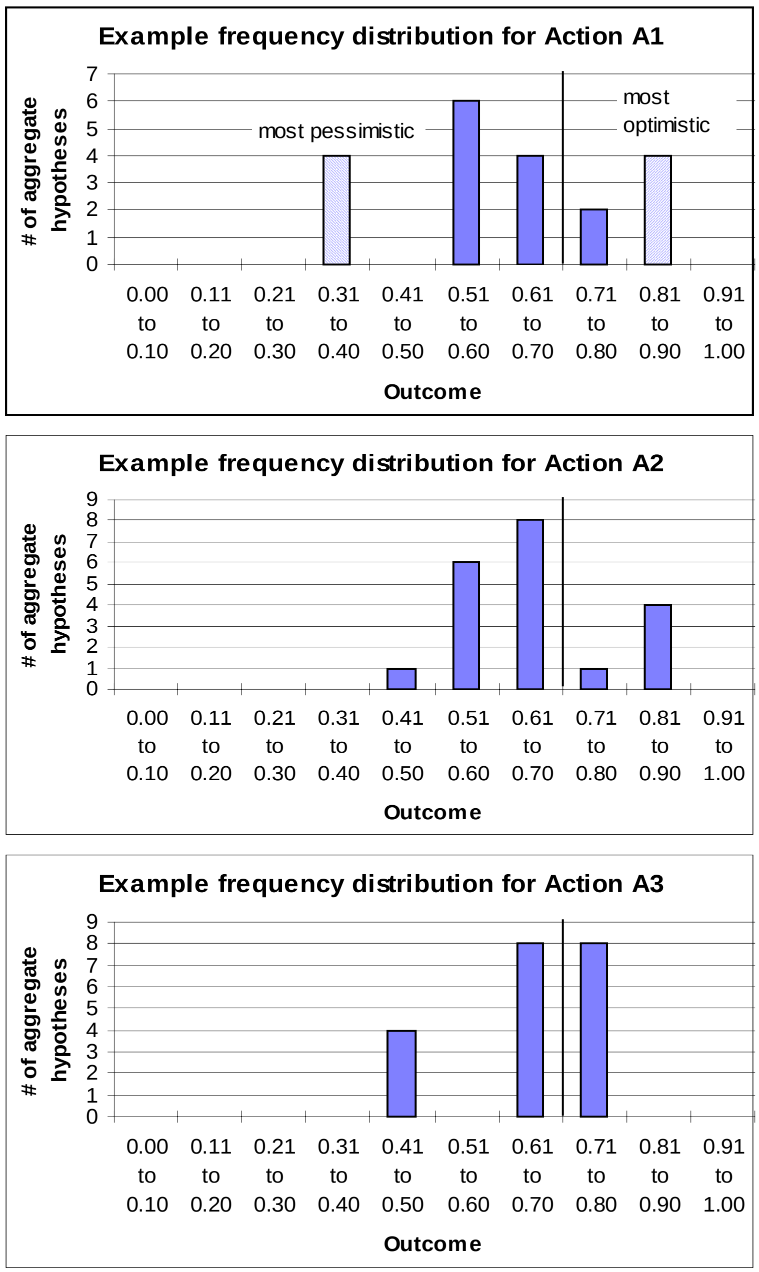


Figure 5.2-2: Example frequency distributions for A1, A2, and A3, based on Table 5.2-1.

5.2.2 Fraction of aggregate hypotheses that meet some criterion

Frequency distributions such as those in Figure 5.2-2 are effective in communicating the degree of uncertainty, but do not clearly show which actions are preferred if an absolute criterion (such as the 0.7 level defined by the NMFS jeopardy standards) is applied. One way to apply this criterion is to calculate the fraction of aggregate hypotheses that result in a probability of greater than 0.7 (i.e., the fraction of aggregate hypotheses that are to the right of the vertical line at 0.7 in Figure 5.2-2). For the example data, six out of 20 (0.30) aggregate hypotheses meet this criterion for action A1, five out of 20 (0.25) for A2, and eight out of 20 (0.40) aggregate hypotheses meet the criterion for action A3. The calculation is also shown in tabular form in Table 5.2-1.

On a relative basis, A3 would be preferred to action A1 and A2. One could also apply some absolute criterion to this fraction to assess the acceptability of actions. For example, an action might be considered acceptable if the NMFS standard were met in more than half of the "possible futures", that is, if the fraction of aggregate hypotheses producing 0.7 probability of exceeding survival escapement levels was 0.5 or greater. Under this criterion, none of the actions in the above example would be considered acceptable.

5.2.3 Expected ability of actions to meet some criterion

One of the drawbacks of a frequency distribution approach is that it implicitly assumes that all aggregate hypotheses are equally likely to be correct, even though there may be information to suggest that some aggregate hypotheses are more likely to correctly represent the way things actually occur in nature. One way to incorporate the relative "belief" in different hypotheses (i.e., the relative possibility that one hypothesis or another most accurately represents the actual conditions) is to weight each aggregate hypothesis by a probability between 0 and 1. The weighting that is placed on a particular aggregate hypothesis reflects the relative belief that the hypothesis best represents the way things work in nature. For example, if one were absolutely certain that a particular aggregate hypothesis represented the way things worked, that aggregate hypothesis would be assigned a weighting of 1, while all of the others would be assigned a weighting of 0. If two aggregate hypotheses were considered to be equally possible, each would be assigned a value of 0.5.

Once these weightings are assigned, we can calculate the weighted fraction of aggregate hypotheses in which the probability of exceeding the survival (or recovery) escapement level is 0.7 (0.5 for recovery). This is also referred to as the "expected ability of an action to meet the survival or recovery standard". In columns 5-7 of Table 5.2-1, we place a "1" if the aggregate hypothesis in the row meets the 24-year survival standard (i.e., probability in columns 2-4 is greater than 0.7), and a "0" if it does not. Then, the expected ability to meet the 24-year survival standard in Table 5.2-1 is calculated by multiplying the weighting for an aggregate hypothesis by the "1" or "0" for that hypothesis, then summing over all aggregate hypotheses (this is equivalent to adding up the weights of only those aggregate hypotheses that meet the standard). If all hypotheses meet the standard, the expected ability to meet the standard is 1.0. Doing this calculation using the example data and weights yields an expected ability to meet the survival standard of 0.44 for A1, 0.39 for A2, and 0.43 for A3. These values imply that given the level of uncertainty indicated by the relative weights, there is less than a 50/50 chance that the actions will meet the 24-year survival standard.

Note that the expected ability to meet a particular standard is a function of both the number of aggregate hypotheses in which that standard is met and the relative weighting placed on that hypothesis. For example, high expected abilities can arise either when a large number of aggregate hypotheses result in that standard being met or when a small number of highly-weighted aggregate hypotheses meet the standard. Because the weighted average probability and the expected ability to meet survival or recovery standards include all possible outcomes and their relative weights, they effectively capture the uncertainty associated with predicting future outcomes from imperfect or incomplete information.

Ideally, weightings assigned to aggregate hypotheses should be based on whatever empirical evidence is available at the time of the analysis. However, in cases where empirical evidence is unavailable or is interpreted differently by different groups, weightings will have to be based on the personal experience and judgement of individuals. In the future, PATH will attempt to reach consensus on the assignment of weightings to alternative hypotheses for those critical uncertainties that drive the results. To facilitate this process, we attempt in this chapter to identify what those critical uncertainties are (Sections 5.4 and 5.5), and to conduct some preliminary sensitivity analyses of the effects of placing different weightings on critical hypotheses (Section 5.6).

Once the critical uncertainties are identified, and weightings assigned, the focus should be on defining specific combinations of hypotheses (i.e., prospective aggregate hypotheses) that are both internally consistent (i.e., the individual hypotheses are based on common assumptions and logic) and are consistent with specific retrospective aggregate hypotheses (Table 4.1-2). For example, a prospective aggregate hypothesis that is consistent internally and with retrospective aggregate hypothesis H1 in Table 4.1-2 is the combination of FLUSH, FGE1, TURB5, PREM3, T1 or T2, Delta model, Hydro-related extra mortality, and “Markovian” climate. A prospective aggregate hypothesis that is consistent internally and with retrospective aggregate hypothesis H2 in Table 4.1-2 is the combination of CRiSP, FGE1, TURB4, PREM3, T3, Alpha model, Regime shift extra mortality, and cyclical climate.

5.3 Relative Performance of Alternative Actions

PATH can provide only very preliminary results on the performance of alternative actions from analyses completed to date. There are three reasons for this. First, not all proposed hypotheses and management actions have been formally evaluated. Second, because we have not yet assigned relative weights to alternative hypotheses, we assume equal weighting for all aggregate hypotheses. Third, further refinements in our methods will likely occur following peer review. **Because of these limitations, the results we present in this section should not be interpreted as implying that one action is better than another. Instead, they should be seen only as an illustration of how these kind of results might be displayed.**

We separate results generated with the CRiSP-T3 passage model and transportation assumptions from those generated with the FLUSH-T1/T2 model. We did this because these two models represent fundamentally different approaches to estimating mortality through the juvenile migration corridor (see Sections 4.2.1 and A.2.1), and because they are each associated with different assumptions about the relative survival of transported and non-transported fish in the ocean (T3 for CRiSP, T1/T2 for FLUSH; see Sections 4.3.1 and A.3.1). Note that although we refer to CRiSP and FLUSH as alternative hypotheses for convenience, it is really their respective underlying assumptions and mechanisms with which we are concerned. Comparison of some diagnostic outputs of the passage models (in-river survival and total direct survival) are shown in Section 5.8. Both passage models were used in conjunction with the BSM life cycle model to project spawning abundances.

To examine the relative performance of the different actions, we use the three official NMFS jeopardy standards[1]. Survival standards are met when the spawning escapement of the sixth best[2] Snake River index stock exceeds the pre-determined survival escapement level an average of 70% of the time over 24 and 100 years. The recovery standard is met when the geometric mean of projected escapement for the sixth best Snake River index stock over the last 8 years of a 48-year period exceeds the pre-determined recovery escapement level an average of 50% of the time (see Section 3 for a description of the Jeopardy Standards). On the frequency distributions below, the jeopardy standards are represented by the vertical lines at 0.7 for survival measures, and 0.5 for recovery. The fraction of runs to the right of these vertical lines indicate the fraction of aggregate hypotheses that meet the survival and recovery standards. Readers should be aware that the probability thresholds defined above (0.7 for survival, 0.5 for recovery) have been debated by some regional entities. Therefore, we explore the sensitivity of our results to different probability thresholds in Section 5.3.2.

5.3.1 Frequency Distributions of Performance Measures

First, to show the amount of uncertainty in outcomes we show frequency distributions of the probabilities that the sixth best stock will have spawner abundances greater than defined survival and recovery escapement. We do this for A1, A2, and A3 (Figures 5.3-1 to 5.3-3). Frequencies are expressed as the proportion of the total number of runs (aggregate hypotheses) for a given passage model rather than the absolute number of runs because there were an unequal number of CRiSP-T3 and FLUSH-T1/T2 runs. This was the case because some hypotheses were specific to certain passage models (e.g., T1 and T2 transportation models in FLUSH), and because a smaller set of passage hypotheses was run by CRiSP. Both passage models ran both best and worst case set of combinations of passage hypotheses that spanned the range of possible outcomes, but not all intermediate combinations were run by CRiSP.

The results show that there is a large variation in outputs, even within models and actions. For example, probabilities of being above the recovery escapement level generated with the FLUSH-T1/T2 passage model range from 0.15 to 0.85 under A1, while CRiSP-T3 probabilities range from 0.05 to 0.9. For both models, there is greater variation in probabilities associated with recovery escapement levels than in probabilities of exceeding survival escapement levels. This is because the recovery standard averages the number of spawners over only one 8-year period at the end of 48 years. While the ability to meet the survival standard is strongly affected by current stock levels and short-term projections (i.e., all runs begin from the same starting point), the ability to meet the recovery standard reflects projected escapement levels in an 8-year period 10 to 12 generations from now. These levels are more affected by the management action and associated hypotheses, and less affected by the starting point. Secondly, with alternative climate hypotheses, these 8 years may contain good conditions for fish in some runs, and bad conditions in others. Ranges do not appear to be sensitive to time periods, except for FLUSH-T1/T2-derived outputs under A3. There, the range of survival and recovery probabilities over the longer time periods (100 years for survival, 48 years for recovery) is smaller than the range of probabilities over 24 years. This is because the implementation and effectiveness of A3 measures are delayed and would have a larger influence on a 24-year probability distribution. Also, the probabilities for the longer time periods are high, and by definition cannot exceed 1.

[1] . Another jeopardy standard, the 24-year recovery standard, was recommended by the BRWG (1994) but has not been officially adopted by NMFS. Although we have not presented results for the 24-year recovery standard in this chapter, some results for this standard can be included in future drafts of this report.

[2] i.e., 5 out of the seven Snake River index stocks performed better than this sixth stock

CRiSP-T3 results generally have a greater range than FLUSH-T1/T2 runs, particularly under A3. Under this action, CRiSP-T3 probabilities of exceeding the survival level of escapement range from 0.15 to 0.95, and probabilities of exceeding the recovery level of escapement range from 0.0 to 0.85. In contrast, FLUSH-T1/T2 probabilities of exceeding the survival escapement level for A3 range from 0.5 to 1 and the probability of exceeding the recovery escapement level ranges from 0.7 to 1.0.

Again, we note that these results are preliminary and will likely change as hypotheses are modified and assigned weights, and analyses are refined. However, they do show that there is considerable uncertainty in the outcomes of alternative management actions. Probabilities can range anywhere from very low to very high values, depending on the underlying aggregate hypothesis. Because this introduces considerable uncertainty into which decision should be made, it is important to identify the individual components of an aggregate hypothesis that have the greatest effect on decisions. This is the objective of Sections 5.4 and 5.5.

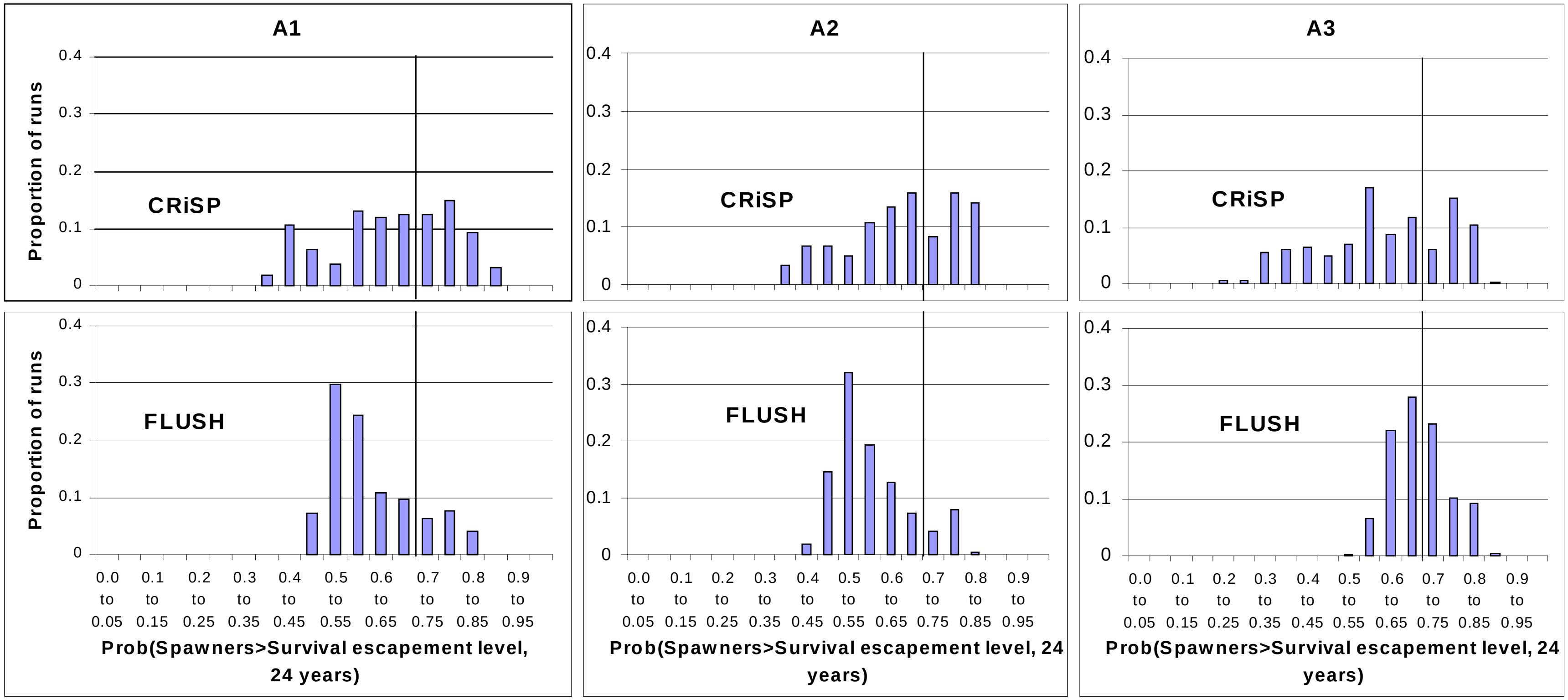


Figure 5.3-1: Frequency distributions of probability of spawners for the sixth best Snake River index stock exceeding survival levels over 24 years. The vertical line at 0.7 represents the criterion associated with the NMFS jeopardy standards.

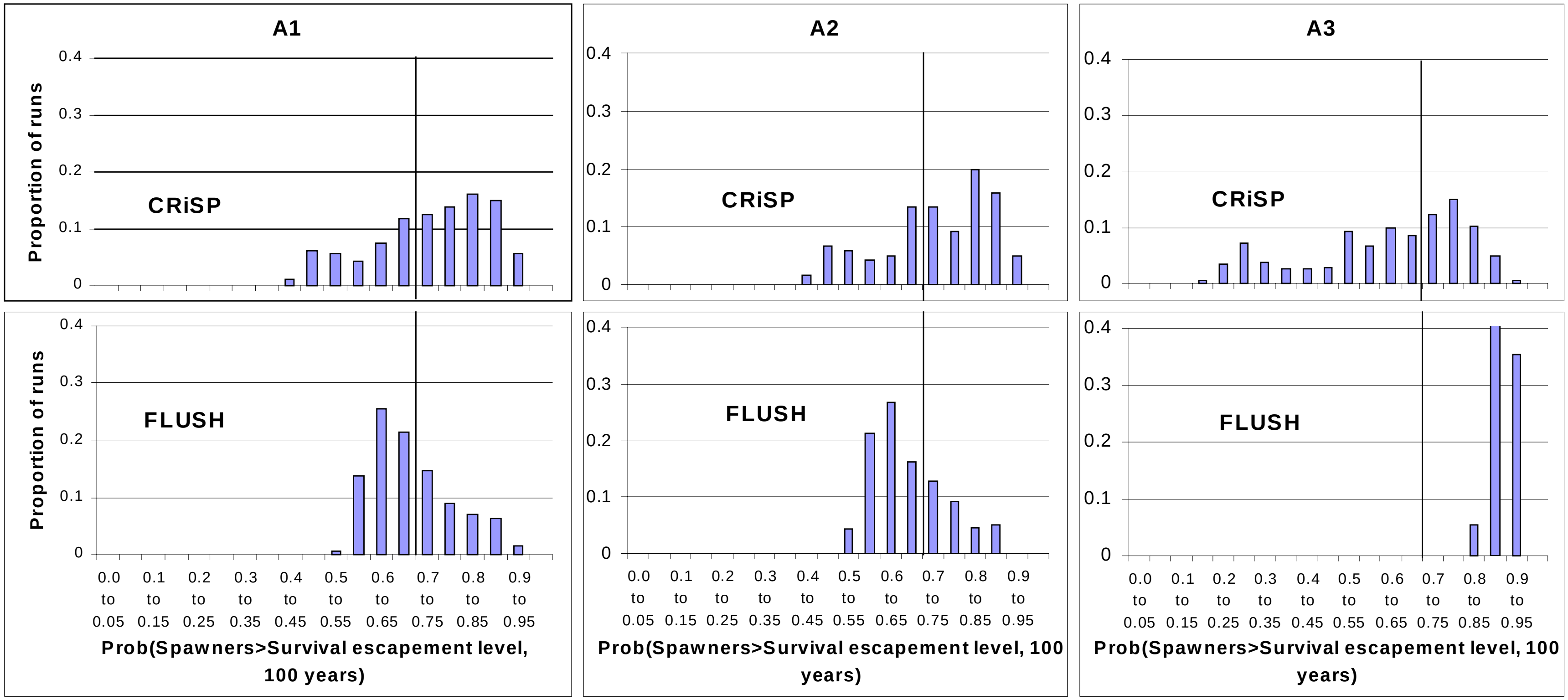


Figure 5.3-2: Frequency distributions of probability of spawners for the sixth best Snake River index stock exceeding survival escapement levels over 100 years. The vertical line at 0.7 represents the criterion associated with the NMFS jeopardy standards.

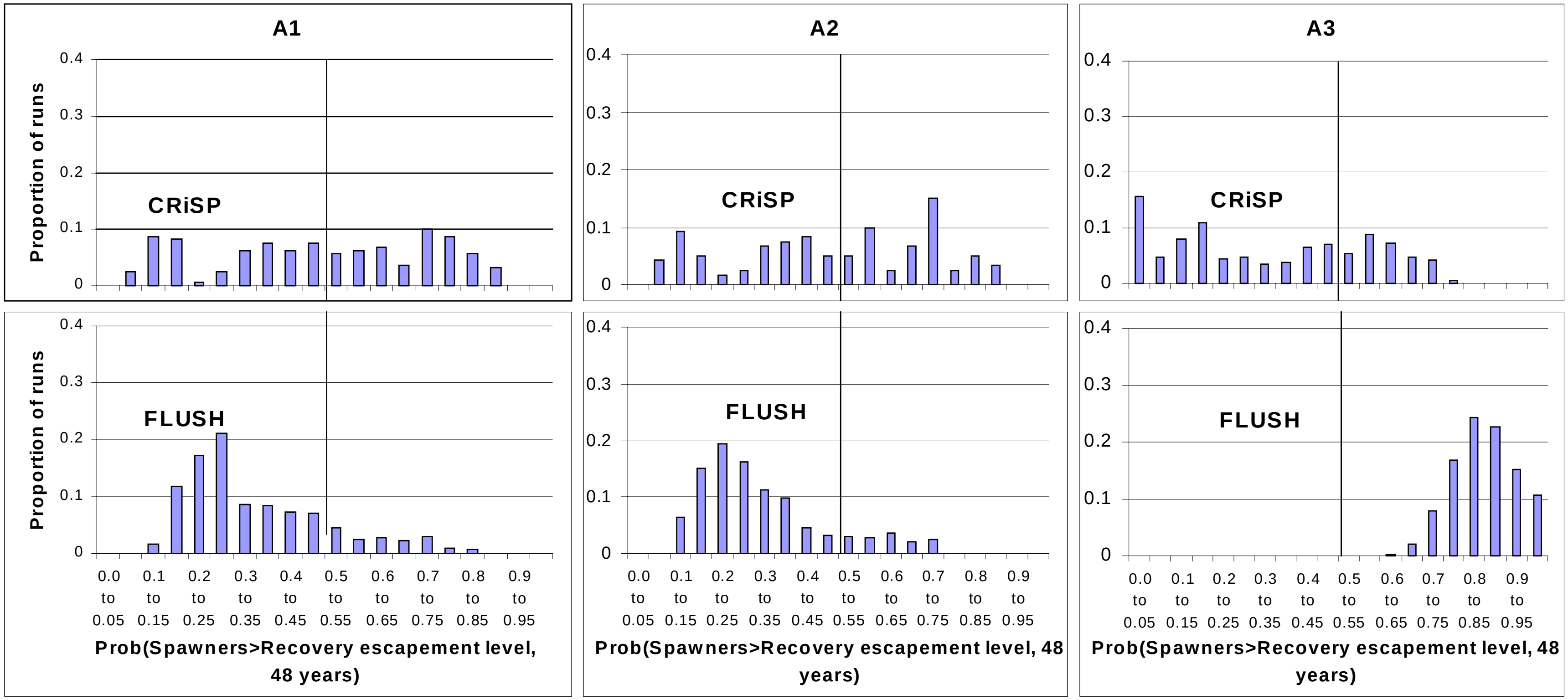


Figure 5.3-3: Frequency distributions of probability of spawners for the sixth best Snake River index stock exceeding recovery escapement levels over 48 years. The vertical line at 0.5 represents the criterion associated with the NMFS jeopardy standards.

5.3.2 Expected ability to meet survival and recovery standards

The weighted fraction, or expected ability of actions to meet survival or recovery standards are compared across actions and passage models in Figures 5.3-4 to 5.3-7.[1] Obviously, higher expected abilities are better than lower values, because they indicate a higher chance of survival and recovery. Using the CRiSP-T3 model, A1 or A2 have very similar expected abilities, while A3 always is the lowest. For FLUSH-T1/T2, A3 is always the preferred option. A1 is next, and A2 always has the lowest expected ability. These patterns hold for all of the jeopardy standards. The expected ability to meet the 100 year survival standard is greater than the 24 year survival standard because there is more time for populations to increase from their current low levels.

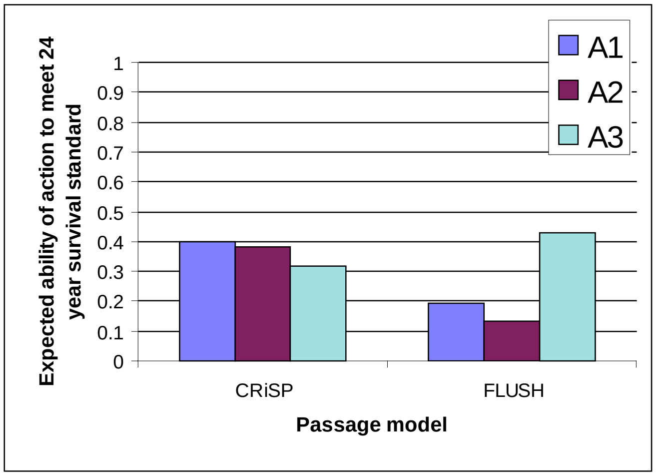


Figure 5.3-4: Expected ability to meet the 24-year survival standard.

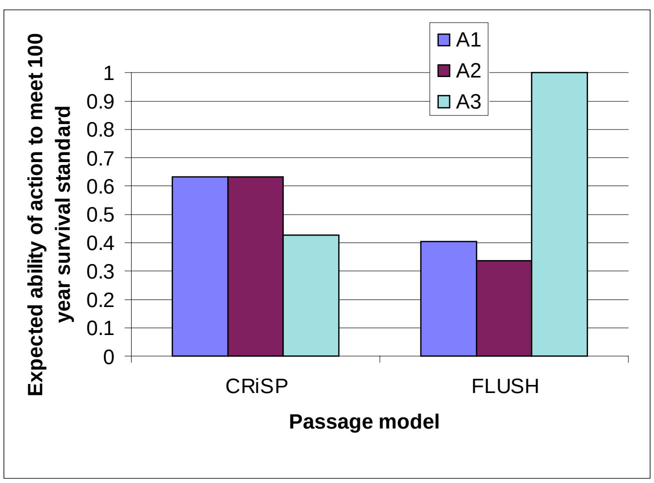


Figure 5.3-5: Expected ability to meet the 100-year survival standard

[1] Within a given passage model, all hypotheses were weighted equally. The weightings were also adjusted to correct for situations where individual hypotheses were not represented equally in the full set of aggregate hypotheses. For example, not all combinations of extra mortality and future climate hypotheses were used. As a result, there were more aggregate hypotheses containing the "cyclical" climate hypotheses than there were containing the "markov" hypothesis

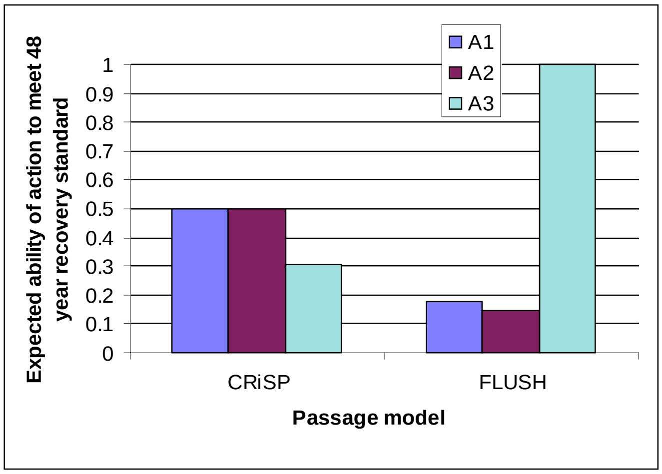


Figure 5.3-6: Expected ability to meet the 48-year Recovery standard

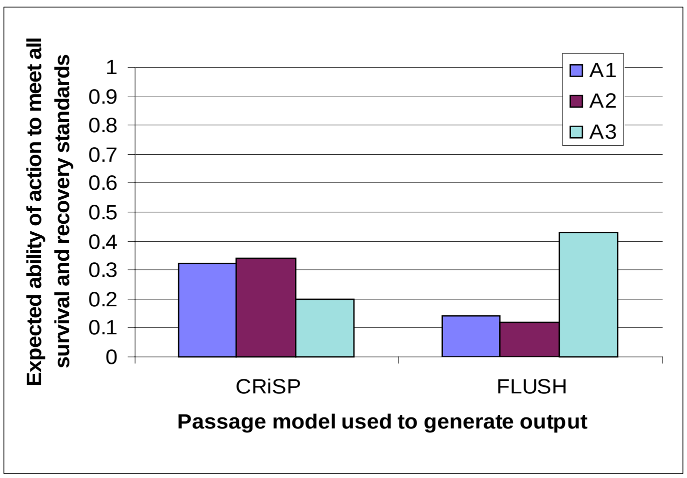


Figure 5.3-7: Expected ability to meet all survival and recovery standards

Sensitivity to weaker and stronger jeopardy standards

The NMFS jeopardy standards are somewhat binary in that an action either meets the standard (e.g., results in a probability of exceeding the survival escapement level of greater than 0.7) or it does not. Therefore, a result that just barely misses the standard (e.g., results in 0.69 average probability of exceeding survival escapement level) is not distinguished from a result that misses the standard by a wide margin (e.g., results in 0.0 average probability of exceeding survival escapement). The result is that the determination of whether an action meets the standard may be quite sensitive to the average probability that is defined as the threshold. We explored the effects of using weaker (i.e., easier to meet) and stronger (more difficult to meet) jeopardy standards than the informal NMFS definition (0.70 probability of exceeding survival escapement levels, 0.50 probability of exceeding recovery escapement levels). As a weaker standard, we assumed that the sixth best stock must exceed the survival escapement level an average of

0.60 of the time, and the recovery escapement level 0.40 of the time. For a stronger standard, we used 0.80 probability of exceeding survival escapement levels and 0.60 probability of exceeding recovery escapement levels. Results are shown in Figure 5.3-8.

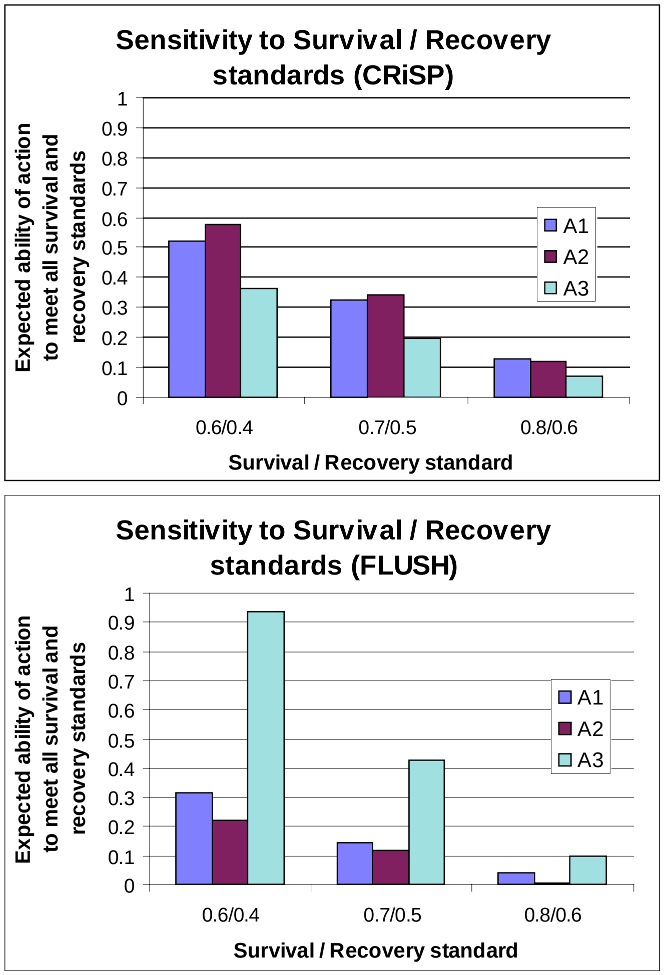


Figure 5.3-8: Sensitivity of the expected ability to meet survival and recovery standards to a weaker (0.6/0.4) and stronger (0.8/0.6) jeopardy standard. The guideline probabilities for the NMFS-defined standard (0.7/0.5) is shown for comparison. A: CRiSP-T3; B: FLUSH-T1/T2.

An alternative approach to assessing the sensitivity of outcomes to the survival and recovery probability thresholds would be to assign each outcome a graduated score from 0 to 1 based on its margin in meeting or missing the standard. For example, an outcome that just barely meets a standard (has a probability of exceeding the survival escapement level of 0.71) might be given a score of 0.5, an outcome that exceeds the probability threshold by a wide margin (e.g., has a probability of 0.90 or greater) might be given a score of 1.0, and an outcome that misses the standard by a wide margin (e.g., has a probability of exceeding the survival escapement level of 0.50 or less) might be given a score of 0.0. The expected ability of an action to meet a standard would then be calculated as a weighted average of these scores over all aggregate hypotheses.

Sensitivity to alternative harvest rate schedules

We also tested the sensitivity of outcomes to lower harvest rates. We considered two additional harvest scenarios. In the first, the harvest rates are reduced by one-third from their values in the current harvest rate schedule (see Tables 4.5.7-1 and 4.5.7-2). In the second, we consider a hypothetical scenario in which harvests of spring-summer chinook are eliminated (i.e., all harvest rates are set to 0). We stress that this "no harvest" scenario is hypothetical only, and is merely intended as a further sensitivity analysis of outcomes to harvest rates.

The results show that the expected ability of actions to meet the survival and recovery standards has either no effect or is only marginally improved when harvest rates are reduced by one third (Figure 5.3-9). The only action and passage model combination in which the expected ability under this harvest scenario differs from the base case (i.e., assuming the current harvest rate schedule) is CRiSP-T3 A3, where the expected ability increases from 0.1 to 0.3, and CRiSP-T3 A2, where surprisingly the expected ability decreases under lower harvest rates. We're not sure why this anomalous result occurs. The effect on outcomes is greater with the "no harvest" scenario. The largest difference is seen for FLUSH-T1/T2 A3, where the expected ability increases from 0.5 under the base case harvest scenario to 0.7 under the "no harvest" scenario.

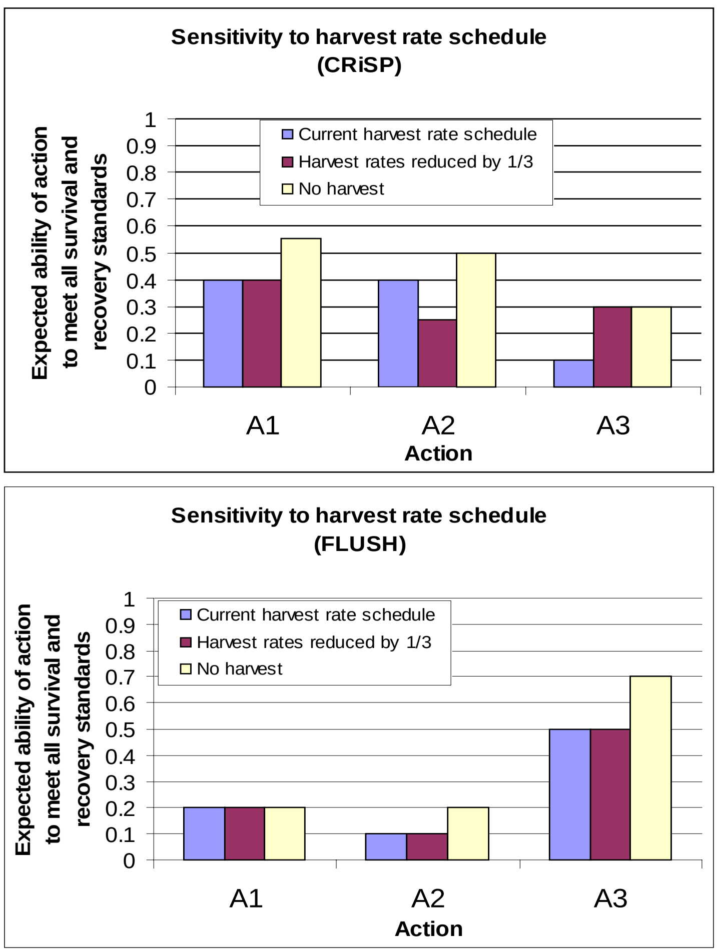


Figure 5.3-9: Sensitivity of outcomes to alternative harvest rate scenarios.

Sensitivity to passage models and associated transportation assumptions

The rank order of actions in this preliminary analysis depends most heavily on the differences in assumptions made by the passage models about mortality through the juvenile migration corridor and the relative survival of transported and non-transported fish in the ocean. FLUSH-T1/T2 favors A3 while CRiSP-T3 favors either A1 (using survival performance measures) or A2 (recovery performance measures). This confirms our general expectations based on the structure and application of these models, but our result explicitly quantifies the differences. Such a quantification is extremely important for focussing debates over their differences.

Sensitivity to survival and recovery standard

The ranking of this set of actions is relatively insensitive to the jeopardy standard that is chosen. Under FLUSH-T1/T2, the ranking of actions was the same regardless of the standard considered. There was a slight change in ranking between survival standards and recovery standards in CRiSP-T3, with either A1 or A2 having the highest expected ability to meet the standards. Although A1 (Status quo) and A2 (maximize transportation) are separate actions, we note that there is relatively little difference between the two in terms of the proportion of fish transported (as calculated by the passage models). For this reason, and because the transportation vs. drawdown question seems to be of most interest in the region, we are concerned primarily with the ranking of A3 relative to A2 and A1 in this sensitivity analysis, and less so with the ranking of A1 relative to A2.

Summary of Results

The results suggest that with this set of actions, there are few instances in which **all** of the survival and recovery standards are met with a high expected ability when all hypotheses are given equal weighting (Figure 5.3-7). The highest expected ability to meet both survival and recovery standards using the CRiSP-T3 model is around 0.35 (obtained with action A2), and around 0.4 using FLUSH-T1/T2 (obtained with action A3). Although a “satisfactory” level has not yet been defined, we would assume that decision-makers would want the expected ability to meet the recovery and survival standards to be high, because that implies a high degree of certainty that these standards will be met.

Actions have a greater ability to meet the longer-term (100 and 48-year) standards than the 24-year standard. The expected ability of A1 or A2 to meet the 100-year survival standard is about 0.6, and about 0.5 for the 48-year recovery standard. Under FLUSH-T1/T2, the expected ability of A3 to meet the 100-year survival and 48-year recovery standard is 1.0.

These preliminary results suggest that significantly greater improvements in survival are required beyond those provided by the management actions analyzed here. This is because none of the current set of actions are able to meet **all** of the standards with any degree of certainty. The 24-year survival standard is the most difficult to achieve, while actions have a higher degree of certainty of meeting the longer-term standards.

This is the case when the aggregate hypotheses are weighted equally. In the final analysis, weightings on aggregate hypotheses will reflect our best joint understanding of the way things work and will not necessarily be equal. Therefore, we explore the sensitivity of the performance of the actions in meeting the standards to the weightings placed on key uncertainties in Section 5.6. First, though, we need to identify what those key uncertainties are. This is the focus of the next section.

5.4 Sensitivity of Outcomes and Decisions to Effects of Uncertainties

We have already noted that the relative outcomes of management actions are greatly affected by the usage of either the CRiSP-T3 or the FLUSH-T1/T2 passage model (Figure 5.3-4 to 5.3-6). In this section we look at the effects of the rest of the uncertainties in Table 4.1-2 on the results of the decision analysis. The primary consideration is the relative effect of each uncertainty on the decision to be made (i.e., how does each factor affect the choice of action to ensure survival and recovery of listed stocks).

Decision Criteria

The process for making decisions about which hydrosystem action or actions to undertake is still being developed, in consultation with many agencies and groups. Here, we use two possible bases for decision-making for the sensitivity analyses in this section, both based on the NMFS Jeopardy Standards. The first is a **relative criterion**, in which the preferred action is the one that simply maximizes the expected ability to meet all three NMFS survival and recovery standards. Actions that result in a large expected ability to meet survival and recovery standards are better.

The second possible basis for decision-making is based on an **absolute criterion**. We assume that some minimum expected ability to meet all of the survival and recovery standards is required for an action to be considered acceptable. That is, we assume that decision-makers will want to be reasonably certain that the action they choose to implement will achieve the 24-year survival standard, the 100-year survival, and the 48-year recovery standard in spite of the presence of uncertainty. Because it is not clear at the moment what the minimum expected ability should be, we use 0.7 for illustrative purposes. The actual threshold minimum may be higher than 0.7, but using a lower value provides a more sensitive test for the significance of the effects of uncertainties.

Although we are looking at a combination of all three of the jeopardy standards in these analyses, we note that the 24-year survival standard is the most difficult to achieve. Therefore, looking only at the longer time periods (100-year survival standard and 48-year recovery standard) will show different patterns in responses. Results of sensitivity analyses for individual standards are shown separately in Appendix B of this report.

Sensitivity of decision to alternative hypotheses

The sensitivity of the decision (using both relative and absolute criteria) to alternative hypotheses can be tested by:

1. assigning weights of 1 and 0 to the alternative hypotheses under consideration (with equal weightings applied to all other hypotheses);
2. correcting for unequal representation of some hypotheses; and
3. comparing the outcomes.

For example, we examine the sensitivity to FGE hypotheses by comparing (Figure 5.4-1):

a) The expected ability to meet survival and recovery standards when FGE1 is assigned a weighting of 1 (with a corresponding weighting on FGE2 = 0, and equal weighting on all TURB, PREM, and other hypotheses)

versus

b) the expected ability when FGE2 is assigned a weighting of 1 (with a corresponding weighting on FGE1 = 0, and equal weighting on all other hypotheses).

Sensitivity to passage-related hypotheses are explored in Section 5.4.1. Sensitivity to other (non-passage) hypotheses are explored in Section 5.4.2. We note that the sensitivity of some hypotheses may be dependent on which other hypotheses are assumed. For example, the ranking of actions may be sensitive to which FGE hypothesis is used only when a particular TURB hypothesis is assumed. Such situations are not be immediately apparent by weighting all other uncertainties equally as we have done in this section, although we are already partially testing for joint sensitivities by doing separate sensitivity analyses for CRiSP-T3 and FLUSH-T1/T2-derived outputs. We explore the sensitivity to combinations of hypotheses in Section 5.4.3.

Summary of Results

Sensitivity analyses are summarized in Table 5.4-1. Details and supporting graphs are provided below.

Sensitivity to passage-related uncertainties

Results presented in Section 5.3 show that the ranking of actions is highly sensitive to which passage model is used in generating the outputs. CRiSP-T3, in general, tends to favor A1 or A2, while FLUSH-T1/T2 favors A3. This pattern holds true regardless of what other hypotheses are assumed, suggesting that the differences in assumptions inherent in the passage model are the main determinants of which action performs the best. Again, this is probably not a surprising result, but these results allow us to explicitly quantify the implications of differences in the passage models.

Within each model, it appears that very few of the uncertainties (at least when looked at independently) have significant effects on either the relative ranking of actions or on the ability of these actions to exceed our decision criterion of 0.7. Although some passage-related hypotheses cause large differences in the ability of some actions to achieve survival and recovery standards, in no cases are the expected abilities significantly greater than 0.5. As we noted earlier, it may be that certain combinations of passage-related hypotheses have effects large enough to change the ranking of actions or to boost the expected ability above our assumed minimum criterion of 0.7. Sensitivity to some of these combinations are explored further in Section 5.4.3.

Sensitivity to other uncertainties

Besides the passage model assumptions, the only other uncertainty that affects the decision is the uncertainty about the source of extra mortality. Under the ”BKD” and the regime shift hypothesis, the expected ability of all of the actions to meet survival and recovery standards falls short of our assumed criterion of 0.7. In other words, if post-Bonneville extra mortality remains regardless of hydrosystem actions, the stocks will have a poor ability to recover. However, if extra mortality is related to the hydrosystem, both A1 and A2 (under CRiSP-T3) and A3 (under FLUSH-T1/T2) exceed this minimum level.

Table 5.4-1: Summary of results of sensitivity analyses. Relative rankings and the absolute criterion are described in the text.

	Sensitivity of decision to uncertainty	
Uncertainty	**Decision based on relative ranking of actions**	**Decision based on absolute criterion of 0.7**
Passage Model	**Sensitive**	Not sensitive
Fish Guidance Efficiency	Not sensitive	Not sensitive
Turbine/Bypass Survival	Not sensitive	Not sensitive
Predator Removal	Not sensitive	Not sensitive
Pre-Removal Period	Not sensitive	Not sensitive
Equilibrated juvenile survival	Not sensitive	Not sensitive
Transition juvenile survival	Not sensitive	Not sensitive
Prospective model	Not sensitive	Not sensitive
Extra mortality	Not sensitive	**Sensitive**
Future climate	Not sensitive	Not sensitive

5.4.1 Effects of uncertainties related to downstream passage

In this analysis we considered four uncertainties related to downstream passage:

1. the effectiveness of extended-length bypass screens in increasing Fish Guidance Efficiencies (FGE1 and FGE2);
2. different hypotheses about the causes of bypass-related mortality during some historical years (TURB1, 4, 5, and 6);
3. the effectiveness of the predator removal program in reducing reservoir mortality (PREM1 and PREM3); and
4. uncertainties related to drawdown
   a) length of the pre-removal period (time between when a decision is made and when removal of dams begins)
   b) juvenile survival rate after river has returned to an equilibrated state
   c) juvenile survival rates during the transition period (time between removal of dams and achievement of equilibrated state).

Analyses for one additional passage-related uncertainty — spill efficiency at Lower Granite, Little Goose, and Lower Monumental dams — were not completed for this draft. Initial results suggest that results may be more sensitive to this uncertainty than to the other passage-related uncertainties. In addition, there may be additional hypotheses for 4 c) (juvenile survival rates during transition period).

Fish Guidance Efficiency

In general, FGE1 (the assumption that FGEs improve using extended-length screens) results in higher overall abilities to meeting survival and recovery standards than FGE2. The relative ranking of Snake River drawdown (i.e., A3 vs. A1/A2) is insensitive to the particular FGE assumption under both FLUSH-T1/T2 and CRiSP-T3.

In terms of an absolute criterion, FGE1 more than doubles the performance of A2 under CRiSP-T3. Even under this higher FGE, however, the expected ability to meet survival and recovery standards is less than 0.7. Expected abilities under FLUSH-T1/T2 are also less than 0.7. If the decision were to be made on the basis of this absolute criterion, no actions would be considered acceptable regardless of the FGE assumption.

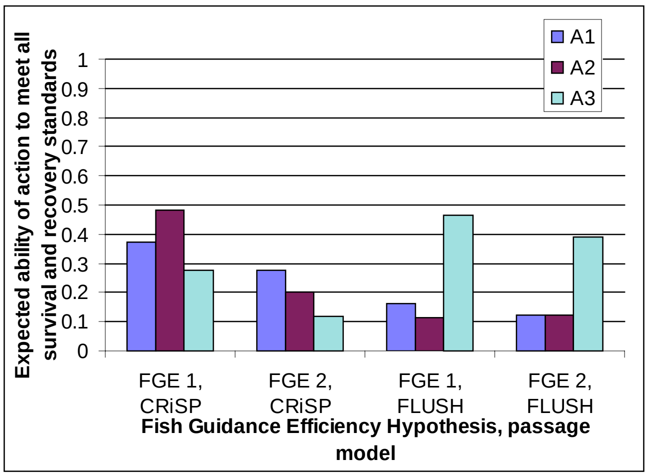


Figure 5.4-1: Expected ability to meet survival and recovery standards under different FGE hypotheses.

Turbine/Bypass Survival

The sensitivity of the model results to the TURB hypotheses was examined in a similar way. Results are shown in Figure 5.4-2. The relative ranking of A3 vs. A1 or A2 is insensitive to TURB hypotheses under both FLUSH-T1/T2 and CRiSP-T3. A1 or A2 have the highest expected ability to meet survival and recovery standards for CRiSP-T3 under all TURB hypotheses, while A3 is always ranked highest under FLUSH-T1/T2.

Expected abilities do not appear to be greatly affected by different TURB hypotheses, and are less than 0.7 in all cases. A decision based on an absolute criterion is therefore also insensitive to the TURB hypothesis.

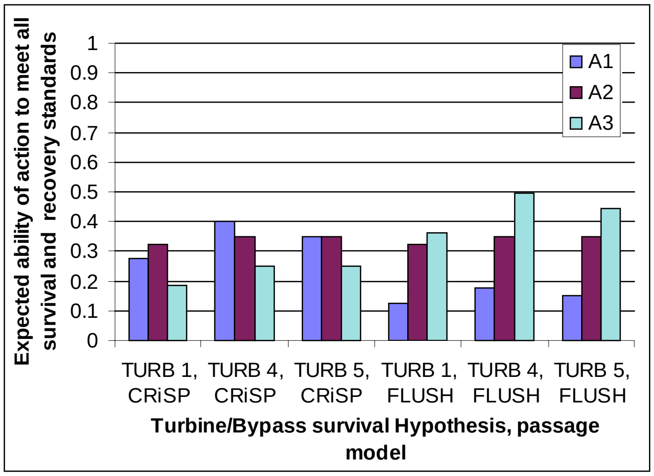


Figure 5.4-2: Expected ability to meet survival and recovery standards under different TURB hypotheses.

Predator Removal

Results using the different Predator Removal (PREM) hypotheses are shown in Figure 5.4-3. Alternative hypotheses have little effect on rankings using FLUSH-T1/T2 (A3 is highest in both cases). With CRiSP-T3, PREM hypotheses affect the rank order of A1 and A2, but do not affect the relative performance of A3.

PREM3 (25% reduction in reservoir mortality due to predator removal program) does tend to result in an increased expected ability to meet the survival and recovery standards, but even this higher PREM assumption does not result in expected abilities greater than 0.7.

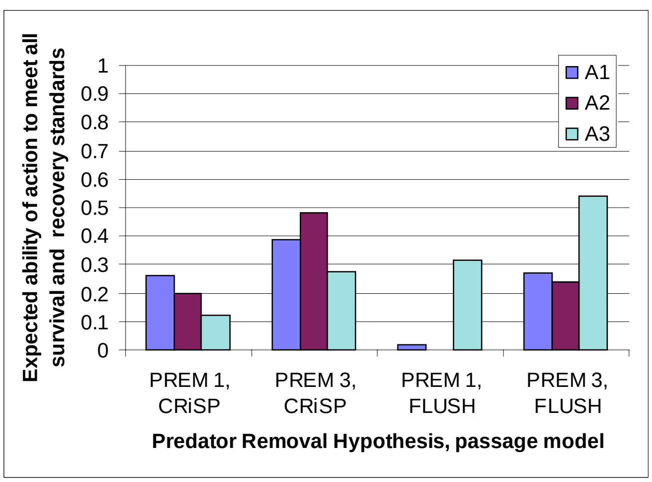


Figure 5.4-3: Expected ability to meet survival and recovery standards under different predator removal hypotheses.

Drawdown

The sensitivity of ranking of actions to uncertainties related to drawdown (A3) is shown in Figures 5.4-4 to 5.4-6. Relative rankings of actions are insensitive to each individual uncertainty. For CRiSP-T3, the expected ability to meet survival and recovery standards under A3 is always lower than the expected ability of A1 or A2, although the higher equilibrated juvenile survival hypothesis (EJUV2) pushes the performance of A3 to close to that of A2 and A1. For FLUSH-T1/T2, A3 is always the best, even under the most pessimistic of assumptions analyzed (e.g., 10-year transition period under TJUVb).

Making pessimistic drawdown assumptions (e.g., long pre-removal (PRER2) or transition (TJUVb) periods), tend to result in substantially reduced performance of A3 under FLUSH-T1/T2. However, the difference is not enough to change a decision based on an absolute criterion of 0.7, because no actions meet this criterion in any case.

Drawdown assumptions are an example where the cumulative effect of certain combinations of drawdown hypotheses may be significant enough to change decisions even though effects of individual hypotheses are not. Sensitivity to combinations of drawdown hypotheses are explored in Section 1.4.3.

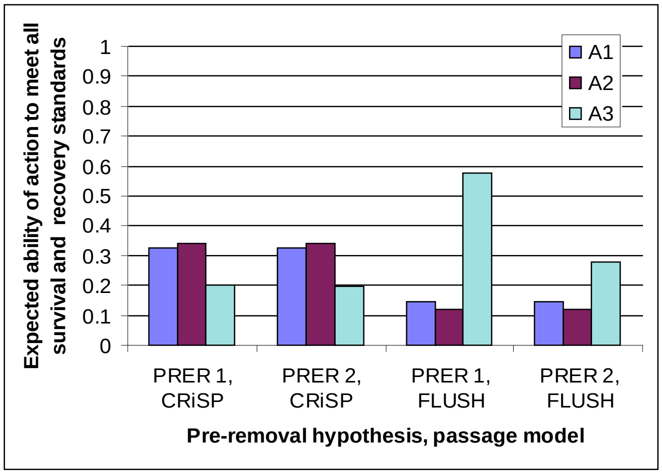


Figure 5.4-4: Expected ability to meet survival and recovery standards under different hypotheses about the length of the Pre-Removal period.

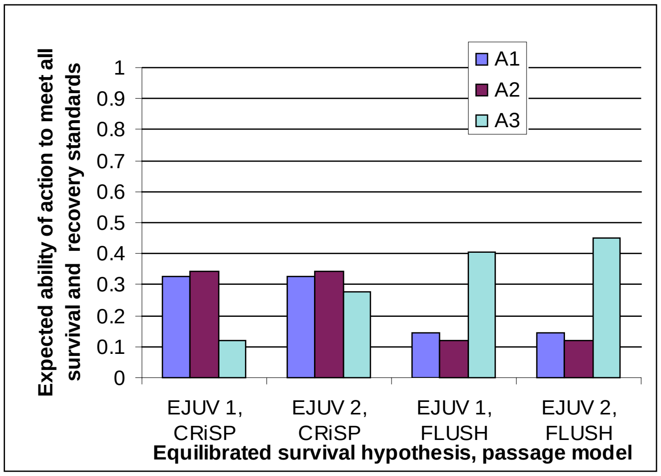


Figure 5.4-5: Expected ability to meet survival and recovery standards under different hypotheses about equilibrated juvenile survival rates after drawdown.

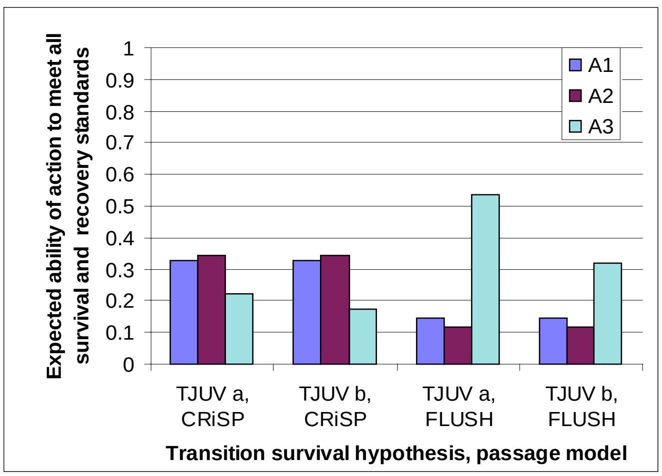


Figure 5.4-6: Expected ability to meet survival and recovery standards under different hypotheses about juvenile survival rates during the transition period between dam removal and equilibrated conditions.

5.4.2 Effects of other uncertainties

In this section, we look at the sensitivity of decisions to “other” uncertainties (i.e., those that are not related to downstream passage). The method for assessing the sensitivity follows that used for passage-related uncertainties in Section 5.4.1. As before, we show CRiSP-T3 and FLUSH-T1/T2 outputs separately. Uncertainties considered in this section are:

1. which prospective model is used: the Alpha model or the Delta model. The two models differ in their assumptions about inherent productivity and the extent to which climate has common effects on upstream and downstream stocks;

2. hypotheses about the sources of "Extra mortality" (mortality that occurs outside of the hydrosystem but is not captured in productivity parameters); and
3. hypotheses about future climate conditions.

Although uncertainties in the response of stocks to future habitat management (i.e., HAB hypotheses in Table 4.1-2) were included in the decision analysis (see Section 4.5.5), we do not look at the sensitivity to habitat effects here because these effects are stock-specific. Because the results presented here are for the sixth best stock, the effects of habitat assumptions are best seen when looking at results for individual stocks. Some of those comparisons are included in Section 5.5.

Prospective Model

Sensitivities of the ranking of management actions to the choice of prospective model (Alpha model vs. Delta model) are shown in Figure 5.4-7. Ranking of A2 vs. A3 is insensitive under FLUSH-T1/T2 and CRiSP-T3. The prospective model seems to have relatively large effects on absolute outcomes, with probabilities generated with the Delta model being higher overall than with the Alpha model, particularly for actions A2 and A3. However, these effects are not large enough to increase the expected ability to meet survival and recovery standards under any passage mode, action, or prospective model to 0.7.

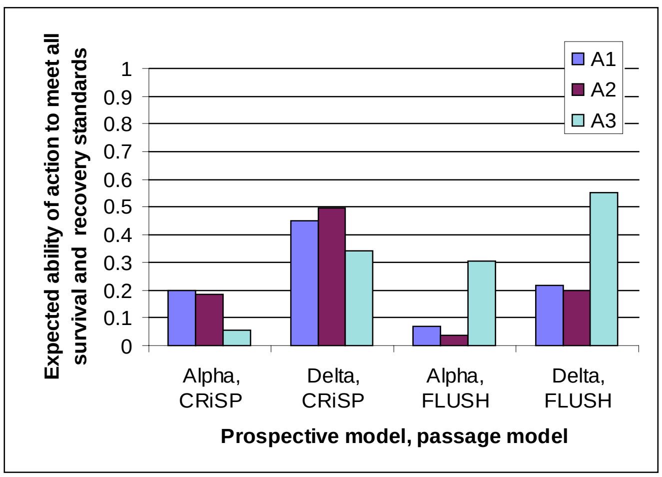


Figure 5.4-7: Expected ability to meet survival and recovery standards using different prospective models.

Extra Mortality

Sensitivity of the ranking of actions to extra mortality hypotheses is shown in Figure 5.4-8. Rankings of actions are insensitive with both passage models. A3 is preferred regardless of what is assumed about extra mortality with FLUSH-T1/T2, while either A1 or A2 is always preferred with CRiSP-T3.

However, extra mortality hypotheses have dramatic effects on decisions that are based on our assumed 0.7 minimum. With both the "BKD" and the regime shift hypotheses, all actions are clearly incapable of meeting this criterion. In fact, with the "BKD" hypothesis there is no chance of meeting the standards with all actions under CRiSP-T3, and with A1 and A2 under FLUSH-T1/T2. The same is true with the Regime shift hypothesis under FLUSH-T1/T2 A1 and A2, and none of the expected abilities for A3 exceed 0.25. Under the Hydro hypothesis, however, both A1 and A2 with CRiSP-T3 and A3 with FLUSH-T1/T2 have a greater than 0.7 expected ability to meet survival and recovery standards.

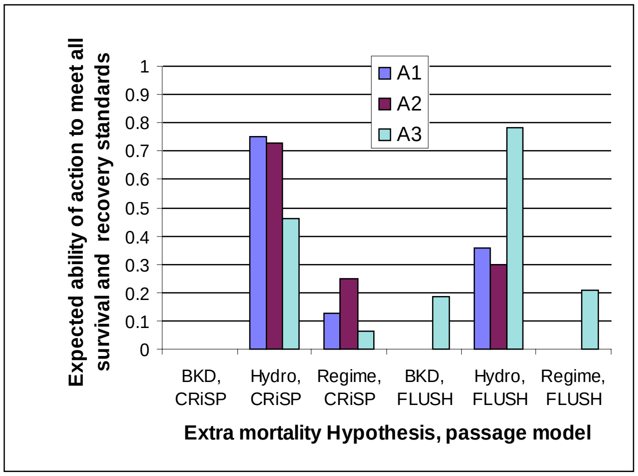


Figure 5.4-8: Expected ability to meet survival and recovery standards under different Extra Mortality hypotheses (A) CRiSP-T3 (B) FLUSH-T1/T2.

Future Climate

Sensitivity of results to uncertainty in future climate effects is shown in Figure 5.4-9. The results suggest that the alternative hypotheses about future climatic effects would not change a decision based on either a relative ranking or on an absolute basis.

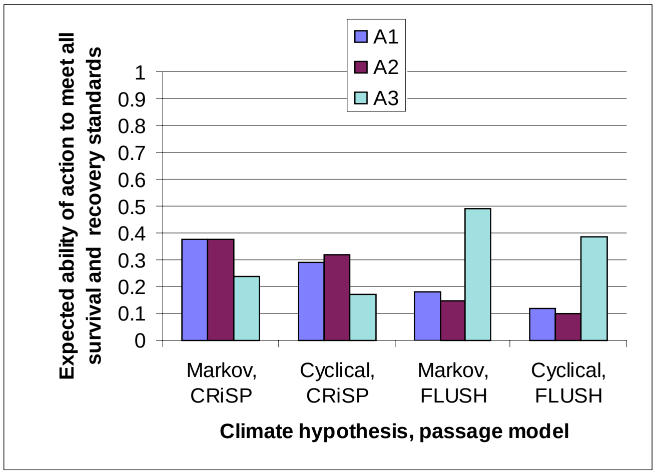


Figure 5.4-9: Expected ability to meet survival and recovery standards under different Future Climate hypotheses.

5.4.3 Effects of combinations of uncertainties

The results in the previous section have looked at the effects of different uncertainties independently of the effects of others. It is possible that uncertainties that independently have minor effects on the results, have significant effects when combined with assumptions about other uncertainties. In this section, we explore the effects of some of these combinations.

We look at the effects of "best-case" and "worst-case" combinations of uncertainties in three major categories (Table 5.4-1). Selection of "best" and "worst" cases are based on the observed effects of each hypothesis on results in Section 5.4.2. The three categories are 1) passage-related hypotheses not associated with drawdown (FGE, TURB, and PREM hypotheses); 2) passage-related hypotheses that are associated with drawdown (PRER, EJUV, and TJUV hypotheses); and 3) hypotheses that are not related to downstream passage, excluding extra mortality hypotheses and habitat hypotheses. Extra mortality hypotheses are excluded because we have already concluded that their individual effects are significant. Habitat hypotheses are excluded because these effects are stock specific and are not likely to show up in results for the 6th best stock. This leaves two uncertainties that are unrelated to downstream passage: the prospective model (Alpha or Delta) and future climate assumptions.

Table 5.4-1: Combinations of hypotheses tested in this section

Category	"Best" Case	"Worst" Case
Passage-related, not associated with drawdown	FGE1 (FGE with extended length screens better than FGE with standard length) TURB4 (relatively high historical dam mortality / relatively low reservoir mortality) PREM3 (25% reduction in reservoir mortality due to predator removal program)	FGE2 (FGE with extended length screens equal to FGE with standard length) TURB1 (relatively low historical dam mortality / relatively high reservoir mortality) PREM1 (0% reduction in reservoir mortality due to predator removal program)
Passage-related associated with drawdown	PRER1 (3 years between decision and dam removal) EJUV2 (96% juvenile survival rate at equilibrated conditions) TJUVa (2 years between dam removal and equilibration)	PRER2 (8 years between decision and dam removal) EJUV1 (85% juvenile survival rate at equilibrated conditions) TJUVb (10 years between dam removal and equilibration)
Not related to downstream passage	DELTA prospective model "Markov" Future climate (climate factors sampled from historical distribution with autoregressive properties)	ALPHA prospective model "Regime shift" future climate (future climate follows a cyclical pattern)

Combination of Passage Hypotheses

Best and worst case passage hypotheses have predictably large effects on results, but they do not significantly affect the relative ranking of actions (Figure 5.4-10). A1 or A2 is still always the best with CRiSP-T3, and A3 is still always the best under FLUSH-T1/T2. In terms of meeting an absolute criterion, the expected ability of actions to meet the survival and recovery standard is below 0.7 for all cases.

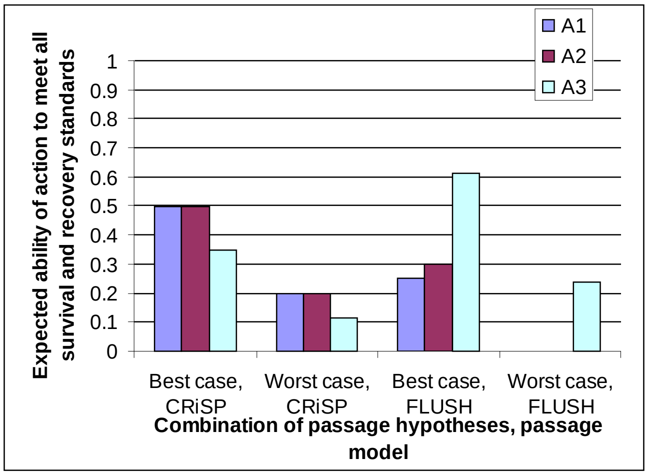


Figure 5.4-10: Expected ability to meet survival and recovery standards under best and worst case combinations of passage-related hypotheses.

Combination of Drawdown Hypotheses

Results are also relatively insensitive to drawdown hypotheses (Figure 5.4-11). For CRiSP-T3, even the best-case set of assumptions about drawdown are insufficient to bring the results for A3 up to A1 or A2, although they are much closer than was seen for any of the drawdown assumptions independently. For FLUSH-T1/T2, even the worst-case set of assumptions is not enough to bring A3 to below A2 or A1. However, the best-case drawdown scenario does elevate the expected ability to meet survival and recovery standards for A3 to around 0.8 under FLUSH-T1/T2. We explore this further in Section 5.6, where we test the sensitivity of meeting the survival and recovery standards to different weights placed on the drawdown assumptions.

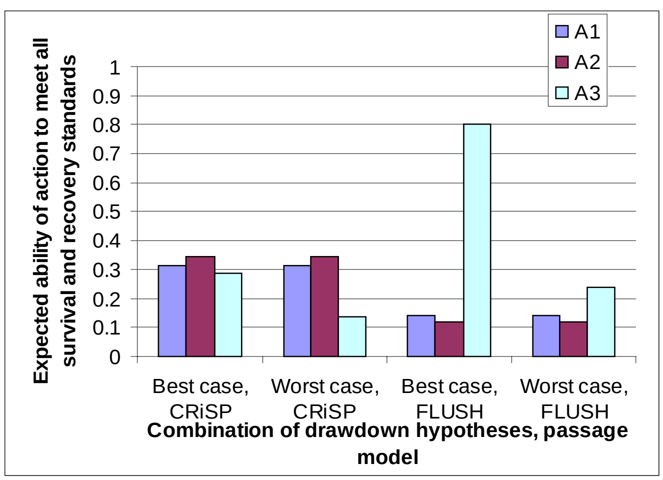


Figure 5.4-11: Expected ability to meet survival and recovery standards under best and worst case combinations of drawdown hypotheses.

Combination of other hypotheses

Results of the sensitivity analysis of best-case and worst-case sets of other (prospective model and future climate) hypotheses are shown in Figure 5.4-12. Rankings are insensitive to non-passage hypotheses for both passage models. A decision based on an absolute criterion of 0.7 is also insensitive, because expected values of meeting survival and recovery standards for any scenario do not exceed 0.7.

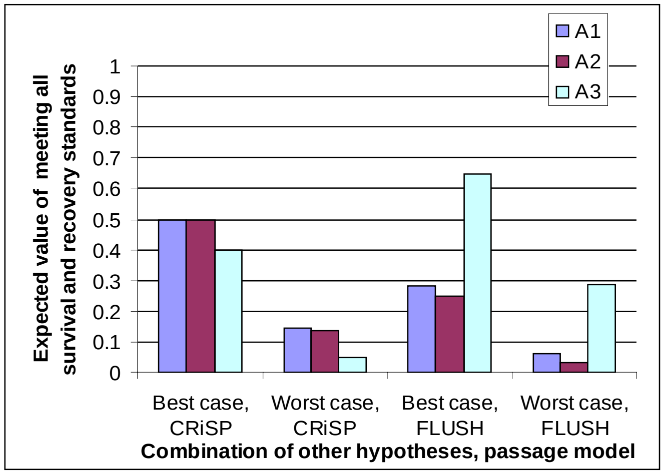


Figure 5.4-12: Expected ability to meet survival and recovery standards under best and worst case combinations of other (non-passage-related) hypotheses.

Summary of Results

Even when comparing best-case and worst-case conditions, a decision based on either the relative performance of alternative actions or on the ability to achieve our 0.70 standard are not greatly affected by passage or by other (non-passage) hypotheses. In all cases, A1 or A2 is still best under CRiSP-T3 while A3 is still best under FLUSH-T1/T2. The results suggest that drawdown assumptions would have to be worse than our worst-case scenario before A2 or A1 was ever better than A3 under FLUSH-T1/T2, and would have to be better than our best-case scenario before A3 was ever better than A2 or A1 under CRiSP-T3. Drawdown assumptions do affect the ability of A3 to meet an absolute criterion of 0.7 under FLUSH-T1/T2, but not under CRiSP-T3.

5.5 Sensitivity of Outcomes for a Single Stock (Marsh Creek) to Effects of Uncertainties

Methods

Previous sections have shown that relatively few combinations of actions and hypotheses actually meet all recovery and jeopardy standards that are based on outcomes for the sixth best stock. This raises the question: “What do you need to do (actions), and what do you need to believe (hypotheses) for a single, weaker stock to meet the NMFS standards?” This section examines actions and hypotheses for a particularly sensitive index stock (Marsh Creek). As a performance measure, we use the probability of Marsh Creek spawning abundances exceeding the survival escapement level (150 spawners) over 24 years.

We chose Marsh Creek because it is one of the weaker index stocks (along with Sulphur Creek). For these analyses, we defined a standard where the average probability of exceeding 150 spawners in any one year must be at least 0.75 (note that this is lightly higher than the NMFS-defined standard, where the average probability must be 0.7 or greater). Although the results of this analysis should be similar to those presented in the previous section, some differences are expected because in these analyses we look at a single stock rather than the sixth best stock, and because we are using a different standard.

Outcomes depend on a combination of three management actions (A1, A2, and A3), and many aggregate hypotheses. Aggregate hypotheses are specific combinations of the individual hypotheses described in Table 4.1-1. Because we needed to use balanced pair-wise comparisons for the analysis, we only looked at aggregate hypotheses that were run for both CRiSP-T3 and FLUSH-T1/T2. The result is a set of 1360 combinations of actions and aggregate hypotheses. More than half of these combinations are for action A3 because there are more uncertainties associated with A3 than with the other actions (i.e., PRER, EJUV, and TJUV hypotheses are only relevant to drawdown). Therefore, the 1360 runs were weighted so that each management action appears in 1/3 of the combinations.

Of the 1360 combinations of actions and hypotheses, only 21.8% (297) meet our 0.75 standard. The question we address in this section is “Of the 297 combinations of actions and hypotheses in which this standard was met, what percentage contained a particular action or hypothesis?” We determine whether this percentage is significant or not by comparing it to the percentage of the entire 1360 combinations that contained that action or hypothesis. If a particular action or hypothesis has no effect on the ability of Marsh Creek stock to meet our 0.75 standard, we would expect that its representation in the 297 “successful” combinations (those that meet the standard) would be the same as its representation in all of the combinations.

Results

Results are shown in Tables 5.5-1 to 5.5-4. For each hypothesis or action, the tables show:

a) the percentage of the total of 1360 combinations that include that hypothesis or action. For example, 50% of the combinations included the FGE1 hypothesis, and 50% included in the FGE2 hypothesis.

b) the percentage of the combinations that meet the 0.75 standard (297) that include that hypothesis or action. For example, out of the 297 combinations that met the 0.75 standard, 220 (74%) included FGE1.

If a) and b) differ substantially, then this indicates that a particular uncertainty has a relatively large effect on the ability to meet our standard of 0.75 average probability of exceeding 150 spawners.

Effects of Passage Models and Actions on Results for Marsh Creek

Effects of different passage models and actions are shown in Table 5.5-1. Overall results suggest that no management action “dominates” the set of results that meet our absolute criterion for Marsh Creek. Of all of the combinations that met the standard, 32% were A1, 31% were A2, and 37% were A3. CRiSP-T3 was considerably more optimistic than FLUSH-T1/T2. Of the 297 combinations that met the standard, 198 (67%) were CRiSP-T3, and 99 (33%) were FLUSH-T1/T2 (versus 50/50 for runs overall). Within CRiSP-

T3, the standard-meeting combinations were broken down relatively evenly into A1 (34%), A2 (34%), and A3 (32%). FLUSH-T1/T2 results were more heavily skewed towards A3; combinations that met the standard broke down into 28% for A1, 25% for A2, and 47% for A3.

What does this mean for decision-making purposes? Essentially, if one places all of the weight on CRiSP-T3 in a weighted analysis, there would be no reason to believe that any action would be more likely than the others to meet the standard (in Marsh Creek). On the other hand, if one placed all of the weight on FLUSH-T1/T2, A3 has a somewhat higher ability to allow the Marsh Creek stock to meet the standard than does A1 or A2. This is consistent with the results presented in Sections 5.3 and 5.4. There, we noted that A3 was consistently better than the other actions under FLUSH-T1/T2, but consistently worse than the other actions under CRiSP-T3.

Table 5.5-1: Percentage of combinations that cause Marsh Creek stock to meet the 24-year survival standard (i.e., spawning escapement is above the survival threshold at least 0.75 of the time), for each of three management actions.

Action	% of all 1360 combinations	% of 297 combinations meeting standard	% of 198 CRiSP-T3 combinations meeting standard	% of 99 FLUSH-T1/T2 combinations meeting standard
A1	33.3	32	34	28
A2	33.3	31	34	25
A3	33.3	37	32	47

Effects of Particular Hypotheses on Results for Marsh Creek

Results for the other hypotheses are shown in Table 5.5-2. Critical uncertainties are those where the representation in the runs that meet the standard is different from the representation in all runs. For example, FGE1 and FGE2 hypotheses are distributed equally (50/50) among all runs, but FGE1 is more highly represented in the runs that meet the survival standard than FGE2 (74% vs. 26%). This suggests that the FGE hypotheses are important in determining whether or not the Marsh Creek stock meets the 0.75 survival standard used here. Distribution of uncertainties is generally not dependant on the action. For example, of the 75% of the runs meeting the standard that included FGE1, 22% were A1, 24% were A2, and 28% were A3 runs. Uncertainties that make a substantial difference are the passage models (PMOD), fish guidance efficiency (FGE), predator removal effectiveness (PREM), the prospective model (Alpha or Delta), and the extra mortality hypotheses. Among those where the results are insensitive are turbine/bypass mortality (TURB), duration of the pre-removal period (PRER), equilibrated juvenile survival (EJUV), juvenile survival during transition (TJUV), future climate, and habitat enhancement (HAB).

It is important to keep in mind what is meant by “insensitive” in this context. If a result is insensitive to equilibrium juvenile survival (EJUV), for example, this means that EJUV hypotheses contribute as much to the 21.8% of the model runs that meet our 0.75 standard as they do to all 1360 runs. In other words, the different hypotheses for EJUV have little if any effect on whether or not the standard for Marsh Creek will be met. This suggests that EJUV is probably not worth worrying about as we consider further model runs (in the short term) or monitoring, small-scale experiments, or large-scale adaptive management experiments (in the longer term).

Table 5.5-2: Distribution of Hypotheses for all combinations and for those meeting the 0.75 24-year survival standard for Marsh Creek. Hypotheses for which the distributions are quite different are highlighted in **Bold**.

Uncertainty	Hypothesis	% of all model runs	% of runs meeting standard			
			Overall	A1	A2	A3
Passage Model & transportation assumption	**PMOD1 – CRiSP-T3**	**50**	**66**	**22**	**23**	**21**
	PMOD2 – FLUSH-T1/T2	**50**	**34**	**10**	**8**	**16**
Fish Guidance Efficiency	**FGE1**	**50**	**74**	**22**	**24**	**28**
	FGE2	**50**	**26**	**10**	**8**	**8**
Turbine/Bypass Survival	TURB1	54	53	14	10	29
	TURB4	23	26	10	11	5
	TURB5	23	22	8	10	4
Predator Removal	**PREM1**	**50**	**25**	**9**	**8**	**8**
	PREM3	**50**	**75**	**23**	**24**	**28**
Pre-Removal Period	N/A (A1 and A2)	67	63	32	31	n/a
	PRER1	17	24	n/a	n/a	24
	PRER2	17	13	n/a	n/a	13
Equilibrated juvenile survival	N/A (A1 and A2)	67	63	32	31	n/a
	EJUV1	17	15	n/a	n/a	15
	EJUV2	17	21	n/a	n/a	21
Transition juvenile survival	N/A (A1 and A2)	67	63	32	31	n/a
	TJUVa	17	22	n/a	n/a	22
	TJUVb	17	15	n/a	n/a	15
Prospective Model	**Alpha**	**50**	**29**	**9**	**8**	**12**
	Delta	**50**	**71**	**23**	**23**	**25**
Extra Mortality	**"BKD"**	**40**	**1**	**0**	**0**	**1**
	Hydro	**40**	**99**	**32**	**31**	**36**
	Regime shift	**20**	**0**	**0**	**0**	**0**
Future climate	"Markov"	40	50	16	16	18
	Cyclical	60	50	16	16	18
Habitat effects	0	50	50	16	16	18
	B	50	50	16	15	19

These results are generally consistent with those in Section 5.4, where we looked at the sensitivity of meeting the NMFS-defined jeopardy standards to the same uncertainties. In that section, we found that the passage models and the extra mortality hypotheses were the most important uncertainties. Here, these uncertainties are also important, particularly the extra mortality hypotheses. FGE, PREM, and prospective model alternatives were also important in these results, but were not identified as such in the previous section. However, we did note in Section 5.4 that these three uncertainties did cause a rather large change in absolute outcomes, but that these changes did not affect either the ranking of actions or the ability of actions to achieve a 0.70 expected ability to meet **all** survival and recovery standards. Because it is probably easier to achieve a single standard such as that used here than it is to achieve all three of the NMFS jeopardy standards, it is not surprising that there are more uncertainties that affect whether this easier standard is achieved.

In terms of future analyses and monitoring, the results suggest that many of the sensitivities could be ignored, because they appear to have relatively modest effects on the model results. Obviously, this interpretation applies only so long as the standard used here (spawning escapement for Marsh Creek exceeds 150 spawners 0.75 of the time over 24 years) is in fact a reasonable facsimile of what managers believe to be an acceptable performance standard. However, additional analyses (not reported here) suggest that other standards (100-year survival, 24-year and 48-year recovery) show similar patterns, and that Marsh Creek is indeed representative of the weaker index stocks. In addition, somewhat higher or lower standards (e.g., 0.8 or 0.7 probability of exceeding 150 spawners) do not seem to make much difference in the patterns noted above.

Effects of Interactions Among Uncertainties on Results for Marsh Creek

The results presented above show the effects of individual uncertainties, and thus do not consider the effects of possible interactions between uncertainties. To investigate this further, we separate the FGE, prospective model, and extra mortality results into CRiSP-T3 and FLUSH-T1/T2 combinations (Table 5.5-3). These results show whether these hypotheses are more sensitive under one passage model than the other. Column 2 of Table 5.5-3 displays the contribution of each hypothesis x passage model combination to the CRiSP-T3 or FLUSH-T1/T2 combinations (680 combinations for each), column 3 displays it for the 198 CRiSP-T3 combinations meeting the 0.75 standard, and column 4 displays it for the 99 FLUSH-T1/T2 combinations that met the 0.75 standard.

Results suggest that the sensitivities noted in Table 5.5-2 to FGE, prospective model, and extra mortality apply regardless of which passage model is used to generate results. The percentage of FLUSH-T1/T2 combinations that meet the standard appears to be more sensitive to FGE and prospective model (Alpha vs. Delta) hypotheses than CRiSP-T3 combinations.

Table 5.5-3: Distribution of passage model x FGE, Prospective model, and extra mortality hypotheses combinations.

Hypothesis	% of CRiSP-T3 or FLUSH-T1/T2 combinations	% of 198 CRiSP-T3 combinations that met 0.75 standard	% of 99 FLUSH-T1/T2 combinations that met 0.75 standard
FGE1	50	67	87
FGE2	50	33	13
Alpha prospective model	50	33	21
Delta prospective model	50	67	79
"BKD" extra mortality	40	0	1
Hydro extra mortality	40	99	99
Regime shift extra mortality	20	1	0

At least one uncertainty that makes a difference – FGE – is probably amenable to relatively short-term, inexpensive experiments. While the measured effectiveness of extended submersible screens differs depending on the method used (PIT tags, fyke nets, etc.), it should be possible to resolve the effectiveness of extended screens by comparing PIT tag and fyke net results in a low-spill year. Given relatively low snow pack so far this year, the spring of 1998 appears to be a year in which such comparisons can be done fairly easily.

The extra mortality hypothesis result is especially disturbing. Results presented in this section and in Section 5.4 indicate that there are several extra mortality hypotheses that cannot be true (or must have a weight of zero) in order to meet a 24-year standard standard. This seems to be true across passage models as well. In particular, if the "BKD" and the regime shift hypotheses are true, then this will prevent the short-term standard being met regardless of the hydro management decision taken. Note that this applies regardless of FGE, predator removal effects, and other sensitivities.

In contrast to FGE hypotheses, it seems unlikely that the extra mortality question can be resolved with conventional, small-scale experiments or monitoring. Perhaps the only way to resolve the extra mortality uncertainty is via large-scale adaptive management experiments, with large contrasts among experimental conditions. For example, one can envision experiments where large-scale management "variables" such as hatchery releases, transportation, or other factors are "switched" on and off in even and odd numbered years. We plan to have a workshop in 1998 to explore the possibilities of such experiments.

5.6 Sensitivity of Outcomes and Decisions to Weightings on Alternative Hypotheses

Research and adaptive management experts should focus on the uncertainties which have the greatest effects on decisions (based either on the ranking of actions or on their ability to meet some minimum criterion). We have scheduled a workshop in 1998 to talk about what kinds of research and experiments might be possible. PATH will also attempt to assign weights to those key uncertainties based on direct empirical evidence from the retrospective analysis, the stated biological rationale, and ecological principles. Because in many cases uncertainty exists because data are either lacking or are interpreted differently, weights will have to be assigned using a structured elicitation of the professional judgment and experience of PATH scientists and/or other experts. However, where strong evidence is lacking, such an elicitation may be unlikely to be able to significantly shift the weights of a set of alternative hypotheses. Once the weights are assigned, focus can then shift to defining those combinations of hypotheses that are internally logical and are consistent with specific retrospective aggregate hypotheses (Table 4.1-2).

The first step in assigning these weights is to establish just how sensitive the decision is to the weightings that are placed on alternative hypotheses. For example, we have already shown that the expected value for action A2 under CRiSP-T3 achieves a 0.7 threshold when the weighting placed on the hydro extra mortality hypothesis is 1 but does not meet this criterion when all of the extra mortality hypotheses are weighted equally. What we would like to know further is how the expected ability to meet survival and recovery standards changes as you change the probabilities within those two extremes. For example, is the 0.7 threshold reached when the weighting on the hydro hypothesis is 0.5? What is the critical weighting that must be placed on the hydro hypothesis before the 0.7 threshold is reached? This information can help to frame the assignment of weights by identifying what the critical weights are. If the expected value is greater than 0.7 as long as the weighting on the hydro hypothesis is greater than 0.6, then the key question to ask when assigning weights to the extra mortality hypotheses is "Is the weight on the hydro hypothesis relative to the "BKD" and regime shift hypothesis greater than or less than 0.6?" This is a much more specific question to answer than "What is the relative weight on the hydro hypothesis for extra mortality?" and can therefore help to focus the discussion when weights are assigned to alternative hypotheses. Precise framing of this discussion will be particularly important where there is disagreement among PATH scientists and agencies over what these relative weights should be.

The purpose of this section is to show how these kinds of analyses can be carried out for the combination of weightings on passage model and transportation assumptions (CRiSP-T3 or FLUSH-T1/T2) and extra mortality. We also look at the effects of different weightings assigned to drawdown assumptions on the expected ability of A3 to meet survival and recovery standards under FLUSH-T1/T2. These uncertainties were selected for these analyses because they were shown to affect either the relative ranking of actions or the ability of these actions to meet some absolute criterion. Moreover, these uncertainties will likely be the most difficult to assign weightings to because of firmly-held beliefs about interpretation of historical data and because extra (post-Bonneville) mortality and drawdown effects are the most difficult to measure. Again, we focus here on the effects of different weightings on the ability of actions to meet all three survival and recovery standards. Effects on individual jeopardy standards are shown in Appendix B.

Sensitivity to weightings on passage models and extra mortality hypotheses

Relative weightings on passage models (and their associated transportation assumptions) and extra mortality hypotheses are jointly represented in Figure 5.6-1. There has to be a great deal of certainty about passage model/transportation assumptions and the hydro extra mortality hypothesis before any action achieves an expected ability to meet survival and recovery standards greater than 0.7. For A1, this criterion is only met if we are absolutely certain (i.e., weight=1.0) that CRiSP-T3 and the hydro extra mortality hypothesis are correct. A2 only satisfies this criterion if CRiSP-T3 is assigned a weighting of 1, and the hydro hypothesis is assigned a weighting of at least 0.9. For A3, the 0.70 criterion is met when FLUSH-T1/T2 is assigned a weight of at least 0.8 and the hydro hypothesis is assigned a weighting of 1.0, or when FLUSH-T1/T2 is assigned a weighting of 1.0 and the hydro hypothesis is assigned a weighting of 0.8.

Sensitivity to weightings on drawdown assumptions

Results for the drawdown assumptions suggest that the length of the pre-removal period (PRER) and the length of the transition period (TJUV) are the most important effects in determining whether the expected ability to meet survival and recovery standards exceeds 0.7 for A3 under FLUSH-T1/T2 (Table 5.6-1). The most optimistic assumptions about these two factors (PRER1 – three-year pre-removal period, TJUVa – two-year transition period) must be assumed to be true with almost complete certainty before A3 meets the 0.7 criterion. If the weights placed on these hypotheses are high, then the assumptions about the equilibrium juvenile survival rate have virtually no effect. For example, if the weighting placed on PRER1 is 1.0, and the weighting assigned to TJUVa is also 1.0, then A3 meets the 0.7 criterion even if the optimistic value for equilibrated juvenile survival rate (EJUV2) has no chance of occurrng.

Table 5.6-1: Combinations of weights placed on PRER, EJUV, and TJUV hypotheses for which A3 under FLUSH-T1/T2 exceeds 0.7 expected ability to meet survival and recovery standards.

Weight placed on PRER1	Weight placed on TJUVa	Minimum weight that must be placed on EJUV2 before 0.7 criterion is met
0.8	1	0.9
0.9	0.9	0.8
	1	0.3
1.0	0.8	0.7
	0.9	0.2

	1.0	0.0

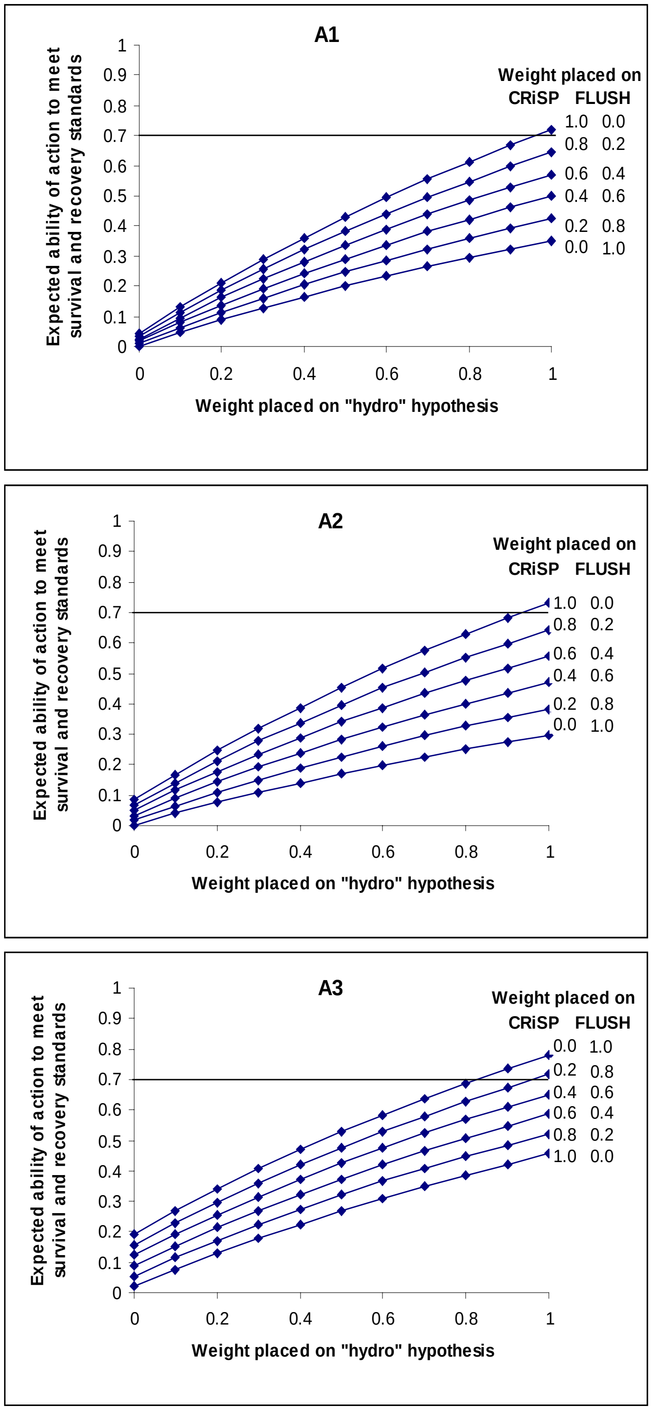


Figure 5.6-1: Sensitivity of expected ability to meet survival and recovery standards to relative weights placed on the "hydro" extra mortality hypothesis and the passage models. Note that the remaining weight placed on the extra mortality hypotheses (i.e., 1 –weight placed on hydro hypothesis) is divided evenly between the "BKD" and the regime shift hypotheses. For example, when the weight placed on the hydro hypothesis is 0.8, the weights placed on the "BKD" and the regime shift hypotheses are both 0.1.

Summary of Results

These results suggest that there will have to be considerable agreement on these key issues before one of the actions is clearly able to meet all of the survival and recovery standards with any degree of certainty. Given the strongly-held beliefs and the lack of data that give rise to the uncertainties, this consensus is not likely to be achievable without a well-planned experimental design. That is, the experimental design must provide a large enough signal to be able to differentiate among hypotheses. The design should be specifically directed towards answering questions about extra mortality and passage model assumptions. Furthermore, it suggests that significantly greater improvements in survival are required beyond those provided by the management actions analyzed here.

5.7 Other Performance Measures

The NMFS jeopardy standards are only one of a number of different measures of performance produced by PATH modeling analyses (see Table 3-1). Other measures include projections of spawner abundances, harvest rates and catch, Smolt-Adult survival rates, and various diagnostic outputs such as survival rates associated with passage through dams and Fish Travel Times. In this section, we briefly report on two of these: projected harvest rates, and Smolt-to-Adult survival rates.

Harvest Rates

Projected harvest rates are important for determining the economic implications of the actions under consideration. This is because harvest rate restrictions for ESA-listed stocks affect the ability to harvest non-listed co-migrating stocks. PATH modeling analyses produce mainstem and tributary harvest rates for all seven Snake River spring/summer chinook index stocks in every 5th year of the 100-year simulation period. Similar computations of harvest statistic for lower Columbia River stocks are planned, but have not been done. These values can be computed for each of the 5,148 aggregate hypotheses, but we have only done so for optimistic and pessimistic aggregate hypotheses in each action. The optimistic aggregate hypothesis was the one that maximized the average spawner abundance for that stock over the 100-year simulation period, while the pessimistic aggregate hypothesis minimized the average number of spawners. Although these are only a small subset of the entire set of aggregate hypotheses, a comparison of optimistic and pessimistic scenarios does give an indication of the range of outputs we can expect given the uncertainties in the analyses. The results presented here are based on the current harvest rate schedules for spring and summer chinook.

We present in this section only a few examples of the types of harvest statistics we could report. First, we show an example of the trends in harvest rates over time (Figure 5.7-1). These results are for mainstem harvest rates for a single stock (Imnaha), a single action (A1), and an optimistic aggregate hypothesis. Because the Imnaha stock is a mixed spring/summer stock, the harvest rate for that stock is estimated by averaging the spring and summer run harvest rates. Time trends in mainstem and tributary harvest rates for Imnaha and Marsh Creek stocks for all three actions are shown in Appndix B. We also show projected time trends in spawning escapement for these two stocks in Appendix B.

The figure can be interpreted as follows. In each year, we produce a frequency distribution of harvest rates rather than a single estimate. This distribution arises from the uncertainty and randomness in the biological and environmental processes underlying salmon population dynamics. To capture this uncertainty, we ran the life-cycle model one thousand times, with each run using a randomly selected value for factors such as future water flows and salmon productivity parameters (this approach is described in more detail in Chapter 3, Section 4.1, and Appendix A.1). Because each simulation run can result in a different harvest rate, the result is that we have a frequency distribution of 1000 possible harvest rates in each year.

This frequency distribution is represented in the figures below by a "box and whisker" plot. The bottom end of the lower line (i.e., "whisker") in each year represents the 10th percentile of the distribution, which means that 10% of the 1000 possible harvests produced by the model for that one year are below that value, and 90% are above that value. For example, the 10th percentile for the CRiSP-T3 output in Simulation Year 35 is 0.12. Therefore, 10% of the 1000 harvest rates simulated for year 35 were below 0.12. The lower end of the box in each year indicates the 25th percentile, the upper end of the box represents the 75th percentile, and the top of the upper line is the 90th percentile. For the CRiSP-T3 output in year 35, this means that 25% of the 1000 harvest rates are below 0.23, 75% of the 1000 harvest rates are below 0.32, and 90% of the harvest rates are below 0.37.

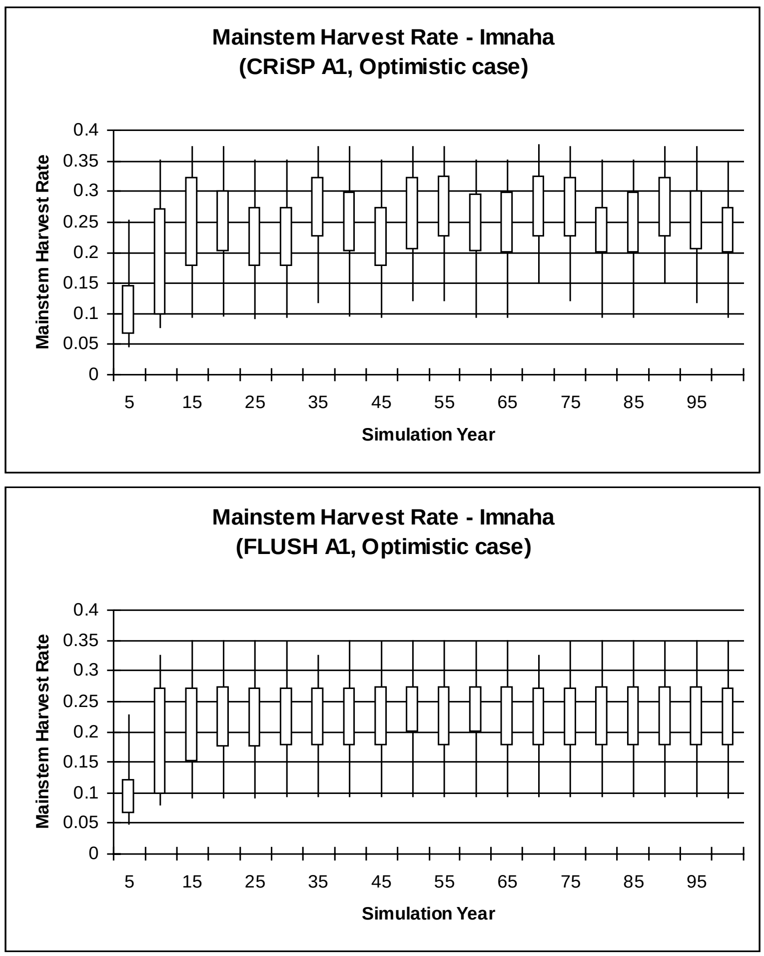


Figure 5.7-1: Mainstem harvest rates under A1 for the Imnaha stock of spring-summer chinook over 100-year simulation period for an optimistic aggregate hypothesis.

The dominant feature in these figures is the amount of variability in annual harvest rates. In most years, harvest rates can range from below 0.1 to above 0.35 for this particular scenario. Note also that these results are only for a single optimistic aggregate hypothesis, and that the degree of variability will be different for different aggregate hypotheses. To illustrate this, Figure 5.7-2 shows the mainstem harvest rates for Imnaha under A1, assuming a pessimistic aggregate hypothesis.

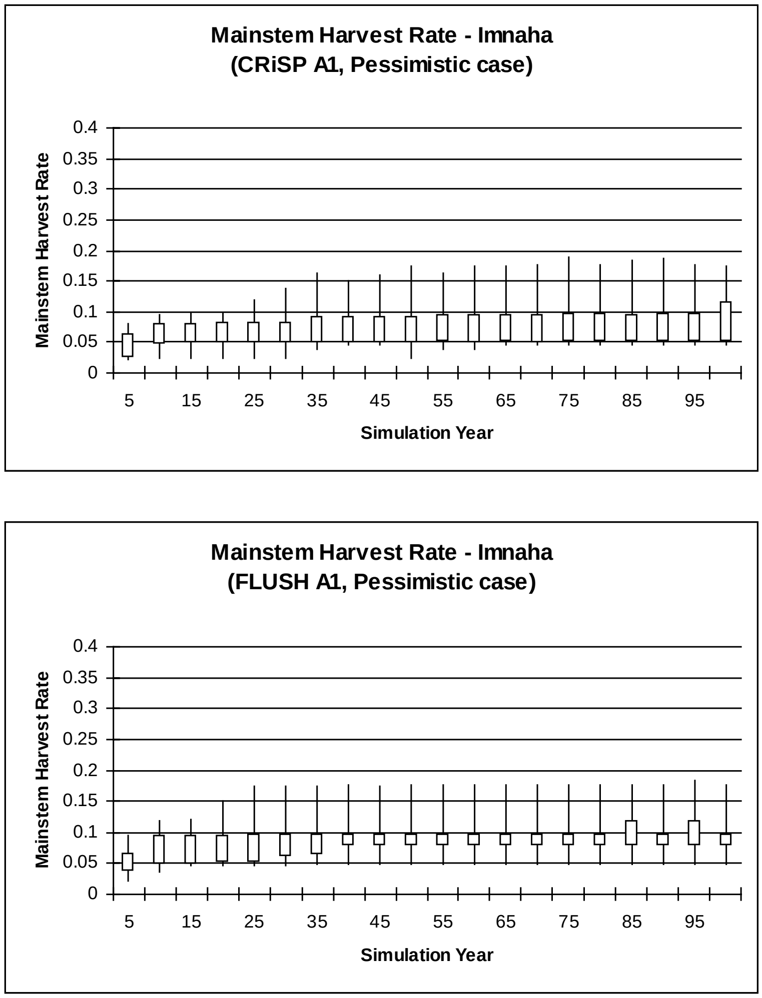


Figure 5.7-2: Mainstem harvest rates under A1 for the Imnaha stock of spring-summer chinook over 100-year simulation period for a pessimistic aggregate hypothesis.

Displays like Figures 5.7-1 and 5.7-2 clearly show trends over time and the amount of variability both within a particular aggregate hypothesis and across aggregate hypotheses. Such uncertainty is important to communicate to decision-makers and to others who will be using this information, such as the economic workgroup. However, these groups of people will also probably want some sort of summary statistic to allow quick comparisons of the harvest implications of different actions and uncertainties. As one example of a summary statistic, we simply calculate the average of the 50th percentile (median) harvest rates in each year over the first 50 years of the 100-year simulation period. Averages for Imnaha are shown in Figure 5.7-3. Another example would be to look at harvest rates in the first and second 25-year blocks. This would distinguish harvest rates experienced during any transition phase from those experienced when the stocks reach some equilibrium.

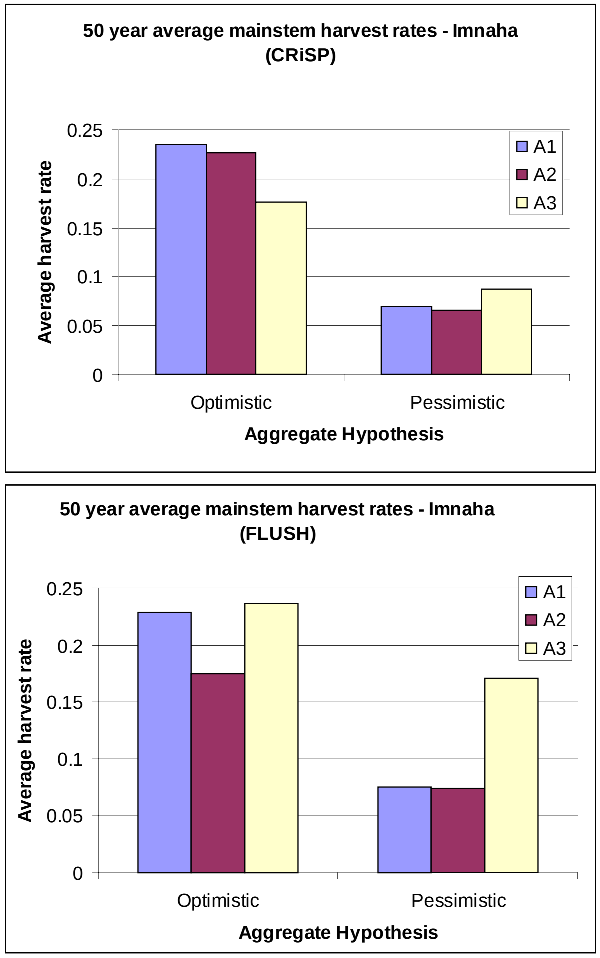


Figure 5.7-3: Fifty-year average mainstem harvest rates for Imnaha stock for CRiSP-T3 (top) and FLUSH-T1/T2 (bottom) model outputs.

Smolt to Adult Survival Rates

Smolt to adult survival rates (SARs) estimate survival rates of fish from the time they pass the upper-most dam as smolts to the time they return to that dam as adults. SARs are estimated in the PATH prospective analyses by relating model estimates of survival over some historical time period to empirical measurements of SARs during that time period, and then projecting that relationship into the future. The model calculates a median SAR over the 100-year simulation period.

Previous analyses by the PATH hydro workgroup suggested that an SAR of between 2 and 6% be used as an interim goal for evaluating whether alternative actions meet survival and recovery standards (Chapter. 6 in PATH FY1996 Retrospective Report). This interim goal was based on Snake River and Warm Springs SARs during periods when those stocks were believed to be healthy, and on theoretical SARs associated with a range of Snake River egg-smolt survival rates from the last three decades.

The purpose of this section of the report is to determine the range of SARs associated with meeting survival and recovery standards. To do this, we show a frequency distribution of the 100-year median SARs for those aggregate hypotheses in which the 100-year survival standard is met (i.e, the 100-year average probability of exceeding the survival escapement level for the 6^{th} best stock is greater than 0.7). We use the 100-year survival standard for this comparison because this time period corresponds to the 100-year time period used to calculate median SARs in the model.

Results show that SARs between 2 and 7% are associated with meeting the 100 year survival goal (Figure 5.7-4). This is very close to the interim goal identified by the PATH Hydro group. In general, CRiSP-T3 SARs for runs that met the 100-year survival standard were lower than FLUSH-T1/T2 SARs. Further analyses of projected SARs are presented in Appendix B.

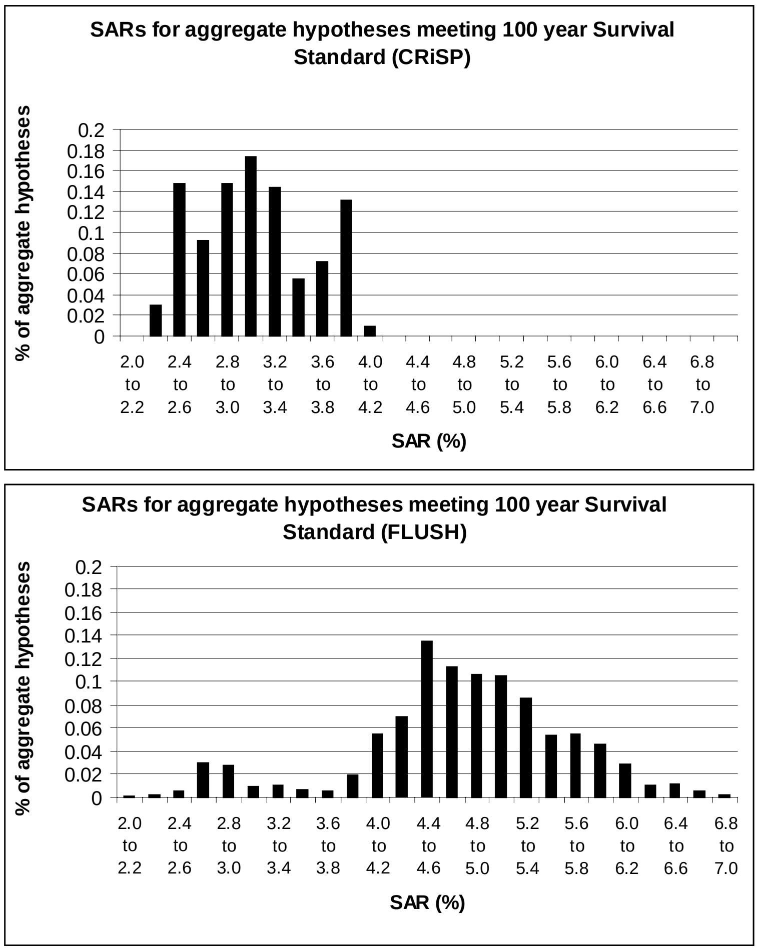


Figure 5.7-4: Frequency of distribution of average SARs for those aggregate hypotheses that met the 100-year survival standard.

Qualitative Performance Measures

In addition to quantitative performance measures, we would also like to look at how well the alternative management actions do in terms of qualitative measures of performance. Such qualitative measures can allow us to incorporate less quantitative but nonetheless important issues relating to the relative health of individual salmon populations, aquatic communities, and entire ecosystems. Many of these issues have been synthesized and discussed in the ISG report "Return to the River". One possible way for PATH to incorporate the Return to the River concepts is to construct a table like that shown in Table 5.7-1. For each indicator of a "normative river" identified in the Return to the River report, we (or members of the ISG) could use some sort of scale to score the conditions created by that action. For example, we can score the actions based on their effects on life history diversity in the Columbia River Basin, which was identified in the Return to the River report as a key element of a healthy ecosystem.

Table 5.7-1: Assessment of Alternative Management Actions based on Qualitative Performance Measures

(Scores of -1 to -3: management action leads to condition inconsistent with normative river (-3 is worst); +1 to +3: management action leads to condition consistent with normative river (+3 is best); 0: action has no effect on this measure).

Measures of Normative River **(Based on review of Return to the River)**	**Direction of positive effect**	**A1**	**A2**	**A3**
life history diversity within the basin	increased			
proportion of basin accessible to salmon	increased			
seasonal fluctuation in flow	increased			
daily fluctuation in flow	reduced			
smolt condition factor in estuary	increased			
mortality rate (in estuary)	reduced			
time for a downstream migrant to reach estuary	reduced			
mortality rate of downstream migrants	reduced			
water temperature - near shore in main channel	reduced			
area of emergent plant production in estuary	increased			
extent of the (marine) freshwater plume - spring	increased			
period of operation of bypass systems	increased			
mortality rate through bypass systems	reduced			
mortality rate in inter-dam reaches	reduced			
mortality rate in reservoirs	reduced			
fall water temperature in Snake river	reduced			
marine bycatch of immature spring chinook	reduced			
condition of mainstream rearing habitats	improved			
utilization of mainstream rearing habitats	increased			
proportion of stock artificially propagated	reduced			
number of hatchery fish released	reduced			
number of intakes with screens to reduce entrainment	increased			
entrainment mortality	reduced			

5.8 Passage Model Diagnostics

In Section 5.3, we noted that the two models used in these analyses represent fundamentally different approaches to estimating mortality through the juvenile migration corridor (see Sections 4.2.1 and A.2.1), and different assumptions about the relative survival of transported and non-transported fish in the ocean (see Sections 4.3.1 and A.3.1). This section compares diagnostic outputs for one of these differences (mortality through the migration corridor) to show how these assumptions differ between models.

In-River Survival

All passage model runs indicated that, for a given water year, in-river survival was higher under the drawdown scenario A3+EJUV2 at equilibrium than under the maximum transport scenario A2 (Figures 5.8-1 and 5.8-2). (Comparisons for A3+EJUV1 are not available at this time). For scenario A2, FLUSH predicted in-river survival ranging from 4-35%, depending upon historical TURB calibration assumptions, water year, PREM, and FGE assumptions, while the CRiSP model predicted in-river survival ranging from 28-50%. For scenario A3+EJUV2, CRiSP estimates ranged from 59-71%, while FLUSH model estimates ranged from 42-77%. The estimated improvement in reservoir survival with drawdown was greatest with FLUSH TURB1, TURB5, and TURB6 (38-53% in-river survival difference between A2 and A3+EJUV2) and lowest with the CRiSP model (19-26% in-river survival difference). FLUSH TURB4 predicted intermediate in-river survival improvements (approximately 34-40%). Results were relatively insensitive to FGE assumptions, but were quite sensitive to PREM assumptions, especially for FLUSH.

Total Direct Survival

Total direct survival of both transported and in-river migrants to below Bonneville Dam (i.e., not including post-Bonneville mortality of either transported or non-transported fish) was estimated using the PATH Hydro Work Group assumption that survival of transported fish in barges is 98% for scenario A2. Total direct survival under scenario A3+EJUV2 is identical to the in-river survival reported previously because no fish are transported. All passage model runs indicated that, for a given water year, total direct survival was higher under scenario A2 than under scenario A3+EJUV2 at equilibrium (Figures 5.8-3 and 5.8-4). (Comparisons for A3+EJUV1 are not available at this time). Total direct survival estimates of FLUSH (including all four TURB assumptions) and CRiSP overlapped, with both models estimating total direct survival between 70-95% under scenario A2 and 44-78% under scenario A3+EJUV2. CRiSP estimates of the difference between the two scenarios (9-25%) were lower than FLUSH estimates of the difference (14-40%) for nearly all water years.

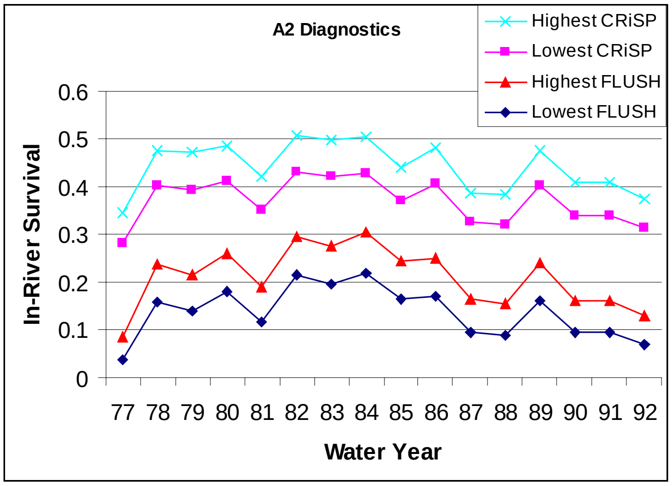


Figure 5.8-1: Range of in-river survival rates estimated by CRiSP and FLUSH for action A2.

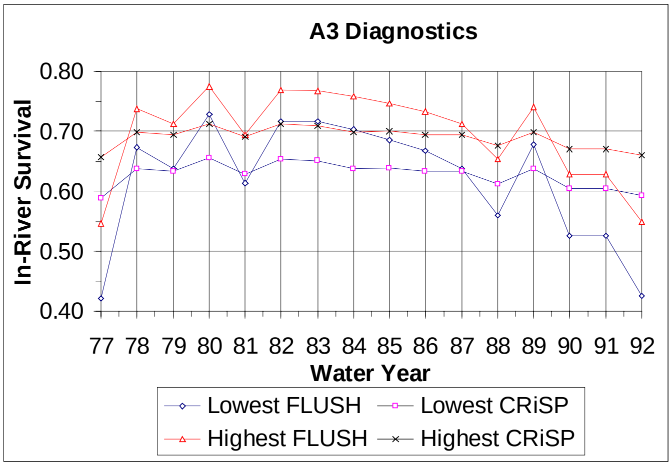


Figure 5.8-2: Range of in-river survival rates estimated by CRiSP and FLUSH for action A3.

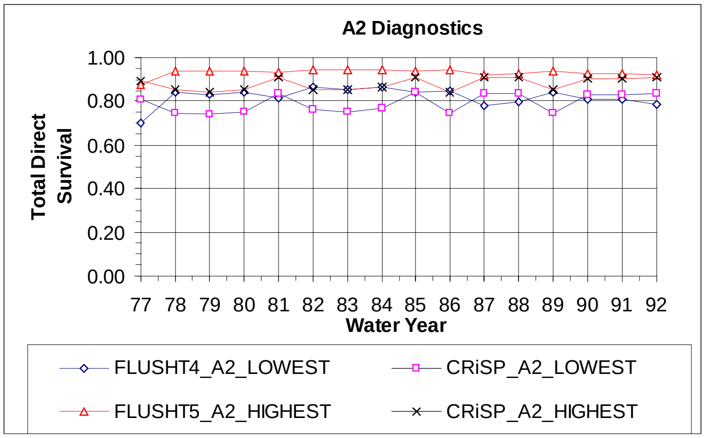


Figure 5.8-3: Range of total direct survival rates estimated by CRiSP and FLUSH for action A2.

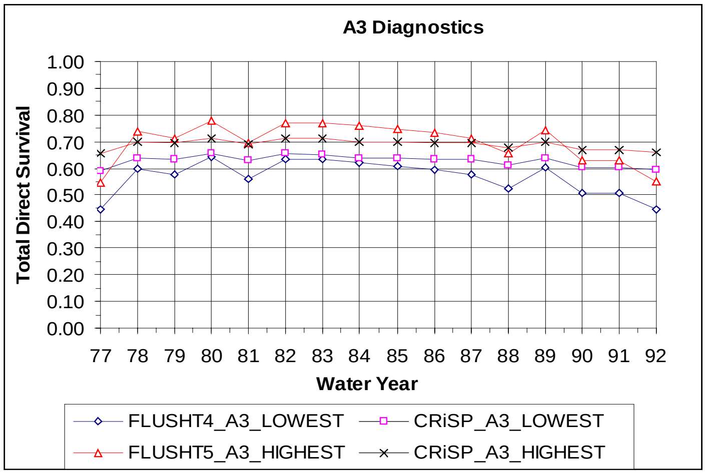


Figure 5.8-4: Range of total direct survival rates estimated by CRiSP and FLUSH for action A3.

6. Glossary

Aggregate hypothesis: A set of alternative hypotheses about all components of the system (stock productivity, downstream migration, marine survival, etc.).

Alpha Model: One of two models of salmon population dynamics used in the PATH prospective analyses. It is based on a Ricker stock-recruitment function, with additional terms for direct juvenile passage mortality and for remaining additional mortality from natural and anthropogenic causes. These two terms are assumed to be specific to the Snake River, Mid-Columbia, and Lower Columbia regions (see Delta Model).

Bacterial Kidney Disease (BKD): A serious salmonoid disease which can cause death or health impairment in both juveniles and adults.

Brood year (BY): The year in which a fish was propagated or spawned.

Coded wire tage (CWT): A tiny tag (1 x 0.25 mm) generally imbedded in the nose cartilage of fingerling or fry while the fish is still in the hatchery. The coded tag allows detailed data on brood year, date of release, and other information to be obtained when the fish is recaptured years later.

D: A parameter used in PATH modeling, equal to post-Bonneville survival of transported fish divided by post-Bonneville survival of in-river fish.

Delta Model: One of two models of salmon population dynamics used in the PATH prospective analyses. It is based on a Ricker stock-recruitment function, with additional terms for direct juvenile passage mortality, an extra mortality factor, and a common year effect. The direct and extra mortality terms are region-specific, while the common year effect acts on all regions (see Alpha Model).

Depensatory: A process that causes mortality rates to increase as abundance decreases. An example of a depensatory process is when the number of individuals removed by predation remains constant as the population abundance decreases.

Drawdown: Releasing water from a reservoir to lower its elevation, thereby reducing surface area and cross-section. This increases water velocity (at any given discharge) in comparison to velocities at higher water levels in the reservoir.

Evolutionary Significant Unit (ESU): A population or group of populations that is considered distinct (and hence a “species”) for purposes of conservation under the ESA. To qualify as an ESU, a population must: 1) be reproductively isolated from other conspecific populations; 2) represent an important component in the evolutionary legacy of the biological species.

Fish guidance efficiency (FGE): The percentage of juvenile fish approaching a turbine intake that are guided into facilities designed to bypass the turbine.

Fish Transit Time (FTT): The time it takes smolts to travel from the head of Lower Granite pool to the Bonneville tailrace.

In-river survival: Direct survival rate of non-transported smolts. The in-river survival rate is estimated from the top of the first reservoir encountered to below Bonneville Dam.

Jeopardy standards: Main performance measures used in this preliminary decision analysis to evaluate alternative management actions and assess sensitivity of outcomes to various uncertainties. The Jeopardy standards are a measure of spawning abundance relative to pre-defined thresholds that are associated with survival and recovery of endangered stocks (see Survival standard and Recovery standard).

Natural river: An option for implementing drawdown of dams where the reservoir is completely drained to create a free-flowing river. This is done either by removing the earthen embankments adjacent to the dam structure, or by building a channel around the dam. In either case, diversion of water around the dam structure results in loss of power-generating capability.

PIT tags: Passive Integrated Transponder tags are used for identifying individual salmon for monitoring and research purposes. The miniaturized tag consists of an integrated microchip that is programmed to include specific fish information. The tag is inserted into the body cavity of the fish and decoded at selected monitoring sites.

Recovery standard: The performance measure used to describe the effect of a certain hydrosystem action on the chance of a spawning stock for recovery; the fraction of simulation runs for which the average spawner abundance over the last 8 years of a 48-year simulation is greater than a specified level (different for each stream).

Ricker *a*: A measure of stock productivity at low levels of abundance.

Spillway crest: An option for implementing drawdown of dams where water levels in the reservoir are lowered to approximately 60-70% of the maximum level. Turbines could continue to operate under this drawdown configuration.

Smolt-to-adult return rate (SAR): Survival rates of fish from the time they pass the upper-most dam as smolts to the time they return to that dam as adults.

Survival standard: The performance measure used to describe the possibility of extinction; the fraction of time during many simulations that the spawning abundance of a stock is above a certain specified low threshold (150 or 300 spawners depending on the characteristics of the stock and the stream).

System survival: The number of in-river equivalent smolts below Bonneville Dam divided by the population at the head of the first reservoir.