

PATH Weight of Evidence Report

Plan for Analyzing and Testing Hypotheses (PATH)
Weight of Evidence Report

Prepared by

David R. Marmorek and Calvin N. Peters (editors)
ESSA Technologies Ltd.
1765 West 8th Avenue, Suite 300
Vancouver, BC V6J 5C6

Contributors:

James J. Anderson, CBR
Larry Barnthouse, McH
Lou Botsford, UC (Davis)
Nick Bouwes, ODFW
Phaedra Budy, USFWS
Tom Cooney, NMFS
Rick Deriso, IATTC
Jim Geiselman, BPA
Al Giorgi, BioAnalysts Inc.
Rich Hinrichsen
Josh Hayes, CBR
Mike Jones, MSU
Lynne 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
Howard Schaller, ODFW
Steve Smith, NMFS
Chris Toole, NMFS
Earl Weber, CRITFC
John Williams, NMFS
Paul Wilson, CBFWA
Richard W. Zabel, UW

August 21, 1998

Citation: **Marmorek, D.R. and C.N. Peters (eds.)** 1998. PATH Weight of Evidence Report. Prepared by ESSA Technologies Ltd., Vancouver, BC, 116 pp. and appendices.

© 1998 ESSA Technologies Ltd.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission from the PATH Planning Group (D. Marmorek, C. Peters, H. Schaller, J. Geiselman, E. Weber, C. McConnaha, C. Toole).

Table of Contents

List of Tables	v
List of Figures	vii
1.0 Introduction	1
1.1 <i>The PATH Weight of Evidence Process</i>	1
1.2 <i>Contents of this Report</i>	2
2.0 Criteria for Weight of Evidence Assessment	5
3.0 Results and Sensitivity Analyses	9
3.1 <i>Brief Description of Modeling Approach</i>	9
3.2 <i>Changes Since the March 1998 Preliminary Decision Analysis Report</i>	15
3.3 <i>Hypotheses that Affect the Ability of Actions to Meet the Standards</i>	15
3.4 <i>Hypotheses that Affect the Relative Ranking of Actions</i>	18
3.5 <i>Implications for the Weight of Evidence Process</i>	20
4.0 Evaluation of Key Hypotheses	25
4.1 <i>Aggregate Hypotheses</i>	25
4.1.1 <i>Description</i>	25
4.1.2 <i>Evaluation of Aggregate Hypotheses</i>	26
4.2 <i>Evaluation of Individual Hypotheses</i>	35
4.2.1 <i>Evaluation of Passage/Transportation Models</i>	35
4.2.1.1 <i>Passage models</i>	36
4.2.1.2 <i>TURB assumptions</i>	45
4.2.1.3 <i>Transportation Models</i>	46
4.2.2 <i>Life Cycle Models</i>	52
4.2.3 <i>Extra Mortality Hypotheses</i>	67
4.2.3.1 <i>Definition of Extra Mortality</i>	67
4.2.3.2 <i>The Historical Pattern of Extra Mortality</i>	67
4.2.3.3 <i>Description of Extra Mortality Hypotheses Modeled to Date</i>	70
4.3 <i>Evaluation of Drawdown Hypotheses</i>	87
4.3.1 <i>Equilibrated Juvenile Survival Rate</i>	87
4.3.2 <i>Transition Period – Duration and Juvenile Survival Rate</i>	90
5.0 Summary of New Hypotheses	93
5.1 <i>Multi-factor Extra Mortality Hypothesis</i>	93
5.1.1 <i>Description</i>	93
5.1.2 <i>Effects on Results</i>	94
5.1.3 <i>Summary of Evidence</i>	94
5.1.4 <i>Issues to Resolve</i>	94

5.2	<i>Hatchery Extra Mortality Hypothesis</i>	94
5.3	<i>Reformulation of the Hydro Extra Mortality Hypothesis</i>	96
5.4	<i>Juvenile Survival Rate During the Transition Period</i>	97
5.4.1	Description.....	97
5.4.2	Effects on Results.....	97
5.4.3	Summary of Evidence.....	99
5.4.4	Issues to Resolve	99
5.5	<i>Adult Survival Rate Under Drawdown</i>	100
5.5.1	Description.....	100
5.5.2	Effects on Results.....	100
5.5.3	Summary of Evidence.....	100
5.5.4	Issues to Resolve	101
5.6	<i>Juvenile Survival Rate at Equilibrium</i>	101
5.6.1	Description.....	101
5.6.2	Effects on Results.....	101
5.6.3	Summary of Evidence.....	101
5.6.4	Issues to Resolve	102
5.7	<i>Productivity as a function of marine nutrients</i>	102
5.7.1	Description.....	102
5.7.2	Effects on Results.....	102
5.7.3	Summary of Evidence.....	102
5.7.4	Issues to Resolve	103
5.8	<i>Additional Habitat Hypothesis</i>	103
5.8.1	Description.....	103
5.8.2	Effects on Results.....	103
5.8.3	Summary of Evidence.....	103
5.8.4	Issues to Resolve	103
Bibliography (incomplete)		105
Glossary of Terms, Acronyms, Variables and Parameters		111
Appendix A: Detailed Results		A-1
A.1	<i>Overall Results</i>	A-1
A.2	<i>Effects of Hypotheses on Ability of Actions to Meet Standards</i>	A-5
A.2	<i>Effects of Interactions Among Hypotheses on Relative Ranking of Actions</i>	A-11
Appendix B: Changes since the March 1998 Preliminary Decision Analysis Report		B-1
	<i>Changes since the Preliminary Decision Analysis Report</i>	B-1
Appendix C: Sensitivity Analysis of FLUSH TRANS 2 Transportation Model		C-1
Appendix D: Additional Diagnostic Analyses of passage / transportation models.....		D-1
D.1	<i>Introduction and Definition of Terms</i>	D-1
D.2	<i>Comparison of ln(prospective:retrospective system survivals)</i>	D-2
D.3	<i>Differences in Retrospective System Survival</i>	D-5

D.4 Differences in Prospective A1 and A2 System Survival D-7

D.5 Differences in Prospective A3 System Survival D-8

D.6 Differences in System Survival – Summary..... D-8

D.7 Differences in Transportation Component..... D-9

Appendix E: Using Categorical Regression Trees to Identify Key Uncertainties E-1

Appendix F: Sensitivity Analysis of D ValuesF-1

Introduction.....F-1

MethodsF-1

ResultsF-3

Appendix G: Sensitivity Analyses of Delta Model G-1

Appendix H: Sensitivity Analysis of Alternative Formulation of the Hydro Extra Mortality Hypothesis.....H-1

Appendix I: Sensitivity Analysis of Alternative FLUSH D ValuesI-1

List of Tables

Table 1-1:	Summary and schedule for Weight of Evidence approach.....	2
Table 3-1:	Hydro system management actions currently under consideration.	12
Table 3-2:	Set of uncertainties and alternative hypotheses considered in this analysis.....	13
Table 3-3:	Analysis of runs meeting all jeopardy standards – overall results.....	15
Table 3-4:	Analysis of runs by which action has a higher jeopardy probability than the other actions, and by how much.....	19
Table 3-5:	Sensitivity analysis of hypotheses that lead to 24-year survival probabilities $A1$ or $A2 \geq A3$ using CRiSP.....	20
Table 3-6:	Summary of relative importance of alternative hypotheses.....	21
Table 4-1:	Comparison of aggregate hypotheses to S-R data: Goodness of fit statistics.....	30
Table 4-2:	Comparison of aggregate hypotheses to observed SAR data.....	33
Table 4-3:	Mean of median projected escapements for 1996-2090 with management action $A1$, and under various aggregate hypotheses.....	35
Table 4-3:	Specific Issues and pieces of evidence relating to Clarity of hypotheses.	37
Table 4-4:	Specific issues and pieces of evidence relating to Mechanisms of hypotheses.....	37
Table 4-5:	Summary of linear regression fits (r^2) to empirical reach survival estimates for CRiSP and FLUSH model estimates.....	41
Table 4-6:	Specific issues and pieces of evidence relating to Clarity of hypotheses.....	47
Table 4-7:	Specific issues and pieces of evidence relating to Mechanisms of hypotheses.....	48
Table 4-8:	Values of STEP estimated for Snake River and lower Columbia River stocks, for three different assumptions about turbine/bypass survival.....	56
Table 4-9:	Mechanistic arguments in support of the delta model representation.....	57
Table 4-10:	Mechanistic arguments in support of the alpha model representation and regional differences in extra mortality.....	58
Table 4-11:	Comments on tests of the alpha and delta model to spawner-recruit and SAR data.....	60
Table 4-12:	Effect of alpha and delta approaches on projected escapements for Johnson Creek under scenario $A1$	67
Table 4-13:	Potential problems with the clarity of how extra mortality hypotheses are currently implemented in the models (EM1 to EM3), or how they might be (EM4 and EM5).....	75
Table 4-14:	Key mechanisms proposed for each extra mortality hypothesis, and source of more detailed descriptions.....	76
Table 4-15:	EM1 . Empirical evidence and arguments in support of the hydrosystem EM hypothesis (a), and evidence/arguments against (b).....	81
Table 4-16:	EM3 . Empirical evidence and arguments in support of the regime shift EM hypothesis (a), and evidence/arguments against (b).....	84
Table 4-17:	Pre-dam and current estimates of survival of wild spring/summer chinook smolts in free-flowing reaches of the Snake River.....	89
Table 5-1:	Adult survival rates for 1991-1993 using radio tags.....	100
Table A-1:	Effects of individual hypotheses on ability of actions to meet all of the standards (CRiSP passage model).....	A-6
Table A-2:	Effects of individual hypotheses on ability of actions to meet all of the standards (FLUSH passage model).....	A-6
Table A-3:	Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of $A1$ to meet all jeopardy standards (CRiSP/T4).....	A-7
Table A-4:	Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of $A2$ to meet all jeopardy standards (CRiSP/T4).....	A-8

Table A-5:	Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A3 to meet all jeopardy standards (CRiSP/T4).....	A-8
Table A-6:	Effects of Interactions between Prospective hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (CRiSP/T4).....	A-8
Table A-7:	Effects of Interactions between Prospective hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (CRiSP/T4).....	A-9
Table A-8:	Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A1 to meet all jeopardy standards (FLUSH/T1).....	A-9
Table A-9:	Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A2 to meet all jeopardy standards (FLUSH/T1).....	A-9
Table A-10:	Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A3 to meet all jeopardy standards (FLUSH/T1).....	A-10
Table A-11:	Effects of Interactions between Prospective hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (FLUSH/T1).....	A-10
Table A-12:	Effects of Interactions between passage hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (FLUSH/T1).....	A-10
Table A-13:	Effects of interactions between Prospective hypotheses and passage hypotheses on ranking of A1/A2 vs. A3 for CRiSP/T4 and the 24-year survival standard.....	A-11
Table A-14:	Effects of interactions between Prospective hypotheses and drawdown hypotheses on ranking of A1/A2 vs. A3 for CRiSP/T4 and the 24-year survival standard.....	A-12
Table A-15:	Effects of interactions between passage hypotheses and drawdown hypotheses on ranking of A1/A2 vs. A3 for CRiSP/T4 and the 24-year survival standard.....	A-12
Table D-1:	Differences between T1 (FLUSH) and T4 (CRiSP) transportation models.....	D-10
Table F-1:	CRiSP Retrospective D Values for different TURB hypotheses.....	F-2
Table F-2:	Median D values using TURB4 and TURB5 assumptions	F-3
Table F-3:	D values for various assumptions about estimates of central tendencies (median, wtd. Arithmetic mean, geomean), break points for defining pre/post time periods, and inclusion or exclusion of 1973 and 1976 values.....	F5
Table F-4:	Effects of assumptions about estimates of central tendencies (median, wtd. arithmetic mean, geomean), break points for defining pre/post time periods, and inclusion or exclusion of 1973 and 1976 values on average 24-year survival probabilities for Marsh Creek (Alpha model).....	F-6
Table G-1:	Sensitivity of average probabilities to spawner-recruit data and X-dam parameter (CRiSP).....	G-1
Table G-2:	Sensitivity of average probabilities to spawner-recruit data and X-dam parameter (FLUSH)...	G-2
Table I-1:	Alternative FLUSH D values	I-1
Table I-2:	Geometric mean alternative FLUSH D values.....	I-1

List of Figures

Figure 2-1:	Assessing the credibility of model projections and assumptions by looking at their consistency with empirical evidence from the real world.....	5
Figure 3-1:	Diagram of analytical approach used in the decision analysis.....	11
Figure 3-2:	Regression tree for the 24-year survival probability for all factors, truncated at R-squared > 0.95 (37 splits).....	16
Figure 3-3:	Regression tree for the 100-year survival probability for all factors, truncated at R-squared > 0.95 (18 splits).....	17
Figure 3-4:	Regression tree for the 48-year recovery probability for all factors, truncated at R-squared > 0.95 (16 splits).....	18
Figure 3-6:	Ability of actions to meet all standards as a function of the weights placed on the passage/transportation model and the hydro vs. regime shift extra mortality hypothesis.....	23
Figure 4-0a:	Spawning escapements (normalized to average escapements) for seven Snake River spring/summer chinook index stocks (upper graph) and six lower Columbia River index stocks (lower graph).....	27
Figure 4-0b:	Comparison of normalized $\ln(R/S)$ for Snake River and Lower Columbia River index stocks, and MLE estimate of m (incremental life-cycle mortality of Snake River stocks to John Day dam).....	28
Figure 4-0c:	Wild Snake River spring chinook salmon smolt to adult return (catch + escapement) rates to the upper dam on the Snake River broken out by periods.....	31
Figure 4-1:	Wild Snake River spring and summer chinook salmon smolt to adult return (catch + escapement) rates to the upper dam on the Snake River (SCK-SAR2, two methods); and steelhead (STH).....	31
Figure 4-2:	Projected spawners for Johnson Creek with best (CRiSP/T4, Delta, Hydro) and worst-case (FLUSH/T1, Alpha, BKD) aggregate hypotheses.....	34
Figure 4-3:	Comparison of observed and predicted reach survival estimates.....	41
Figure 4-5:	SAR's of wild Snake River spring/summer transported fish. DATA are from NMFS.....	51
Figure 4-4:	Projected spawners for Johnson Creek with various combinations of passage/transportation models, life-cycle models, and the BKD extra mortality / Markov climate hypothesis.....	52
Figure 4-6:	Common year effects derived from the Lower Columbia and Snake River spawner-recruit data (Deriso et al. 1996).....	61
Figure 4-7:	Comparison of stock productivities estimated by the delta and alpha models.....	66
Figure 4-8:	Extra mortality estimated from the delta model ($m-M$), where m is the total passage plus extra mortality rate, and M is the passage instantaneous mortality rate for all wild smolts (both transported and non-transported) from Lower Granite pool to below Bonneville Dam.....	68
Figure 4-9:	Pattern of extra mortality $\{1-\exp(-a_n)\}$, as estimated by the alpha model, for the same passage model assumptions as in Fig. 4-8.....	69
Figure 4-10:	Pattern of extra mortality $\{1-\exp(-a_t)\}$, as estimated by the alpha model, for the same passage model assumptions as in Fig. 4-8.....	69
Figure 4-11:	Value of $slope = (1-I_{n,r})/(1-V_{n,r})$ over time from the FLUSH model.....	82
Figure 4-12:	Variation in post-Bonneville mortality factor $(1-I_n)$ over time under the Delta life cycle model, for FLUSH/T1-TURB5 and CRiSP/T4-TURB4 passage model inputs.....	83
Figure 4-13:	Projected escapements for Johnson Creek under hydrosystem-related and regime shift extra mortality hypotheses, relative to historical data.....	86
Figure 5-1:	Sensitivity of the 24-year survival probability to weight placed on the BKD hypothesis.....	96
Figure 5-2:	Trajectory of juvenile survival rates under different scenarios of survival rate response following drawdown.....	98
Figure 5-3:	Sensitivity of 24-year survival probabilities for Marsh Creek to alternative scenarios of survival rate response following drawdown.....	99

Figure A-1: Proportion of runs meeting all standards, by passage/transportation model..... A-1

Figure A-2: Proportion of runs meeting the 24-year recovery indicator, by passage/transportation model..... A-1

Figure A-3: Cumulative Probability Distribution of jeopardy probabilities. Jeopardy standards are indicated by vertical lines. A-5

Figure B-1: Old (TRANS3) and new (TRANS4) transportation assumptions used with CRiSP. B-1

Figure B-2: FLUSH and CRiSP (revised) in-river survival under A3, for both the high (0.96) and low (0.85) hypothesized values of equilibrated juvenile survival rates through the drawdown reach. B-2

Figure B-3: Proportion of runs meeting all standards, by passage/transportation model. Top panel – results from Preliminary Decision Analysis Report. Bottom..... B-3

Figure B-4: Changes in CRiSP A3 results due to changes in D and Vn values. B-4

Figure B-5: Ratio of prospective:retrospective CRiSP/T4 D values..... B-5

Figure B-6: Ln(prospective:retrospective system survival) using Old and New CRiSP/T4 B-6

Figure B-7: Probabilities of selecting a water year from 1977-1992 for prospective analysis. B-6

Figure B-8: Weighted average ln(prospective system survival:retrospective system survival)..... B-7

Figure B-9: Contribution of differences in individual water years to difference in weighted average ln(system survival ratio) for A1 (top) and A3 (bottom) over all water years. B-8

Figure C-1: Effects of FLUSH Transportation models on proportion of runs meeting 24-year survival standard. C-1

Figure D-1: ln(prospective system survival : retrospective system survival) (FGE1, TURB4, PREM3)..... D-3

Figure D-2: Retrospective and prospective system survival and its components (FGE1, TURB4, PREM3).. D-4

Figure D-3: Retrospective and prospective system survival vs. fish travel time (FGE1, TURB4, PREM3)... D-5

Figure D-4: Sensitivity of system survival to M and D (TURB4)..... D-6

Figure D-5: Retrospective M, in-river survival (Vn), and Pbt..... D-6

Figure D-6: Prospective (A1 and A2) M, Pbt, and Vn (FGE1, TURB4, PREM3)..... D-7

Figure D-7: Sensitivity of prospective system survival to M and D (FGE1, TURB4, PREM3)..... D-8

Figure D-8: Prospective (A3) Vn, under the hypothesis that the equilibrated juvenile survival rate through the drawdown reach is 0.96 (FGE1, TURB4, PREM3). D-8

Figure D-9: Retrospective and prospective D’s (TURB4)..... D-9

Figure D-10: Unadjusted T:C’s..... D-11

Figure D-11: Effects of earlier T:C studies on CRiSP/T4 D estimates..... D-11

Figure D-12: Adjusted T:C’s..... D-12

Figure D-13: Comparison of two methods for estimating retrospective D in years with no transportation studies. D-13

Figure D-14: Effects of using two methods for projecting prospective D values (only A2 is shown). D-13

Figure D-15: Effects of two different methods for calculating FLUSH D’s on 24-year survival probabilities..... D-14

Figure D-16: Weighted geometric mean T:C vs. survival of control fish function..... D-15

Figure D-17: In-river survival of control fish used to calculate D values..... D-16

Figure D-18: In-survival of control fish vs. Vn..... D-16

Figure E-1: R-squared vs. # of splits with the optimal number of splits indicated with vertical linesE-3

Figure E-2: Regression tree for geometric mean of jeopardy standards 1 (24-year survival) and 4 (48-year recovery) for all factors, truncated at R-squared > 0.95 (32 splits).....E-4

Figure E-3: R-squared vs. number of splits for the geometric mean of jeopardy standards 1 and 4.....E-4

Figure F-1: Comparison of 24-year survival probabilities using different TURB assumptions (Alpha model)..... F-4

Figure F-2: Comparison of 24-year survival probabilities using different TURB assumptions (Delta model)..... F-4

Figure F-3: Comparison of 24-year survival probabilities using different TURB assumptions, shown for each extra mortality/future climate hypothesis. F-5

Figure F-4: Comparison of 24-year survival probabilities using different methods of calculating D..... F-6

Figure F-5: Comparison of 48-year recovery probabilities using different methods of calculating D.....F-7

Figure F-6: Comparison of 24-year survival probabilities using different methods of calculating D, shown for each extra mortality/future climate hypothesis.F-7

Figure H-1: Sensitivity of A3 jeopardy probabilities to an alternative formulation of the hydro extra mortality hypothesis.....H-2

Figure I-1: Sensitivity to alternative FLUSH Ds, 24-year survival probability.....I-2

Figure I-2: Sensitivity to alternative FLUSH D's for different extra mortality/climate hypotheses (Delta model).....I-3

Figure I-3: Sensitivity to alternative FLUSH Ds, 48-year recovery probability.....I-3

1.0 Introduction

1.1 The PATH Weight of Evidence Process

Limitations in data on past and present conditions, and uncertainties about future conditions (e.g., climate), and differences in the interpretation of existing data, create uncertainties in the response of salmon stocks to management actions and environmental conditions. PATH retrospective analyses have helped to narrow these uncertainties and clarify their relative importance in determining past trends in stocks. Some uncertainties remain and will not be resolved without additional evidence. The existence of uncertainties does not however imply paralysis in decision-making. Decisions must be made with full consideration of uncertainties; otherwise the risks to stocks may be estimated incorrectly.

Techniques such as decision analysis help to clarify these risks by quantifying the effects of uncertainties. Decision analysis will not provide a single answer about stock responses to specific actions; rather, it will show which actions are most robust to the uncertainties. That is, it will show which actions are safest or most risk-averse, given the range of hypotheses and uncertainties in future climate. This report provides revised estimates of the effects of different management actions, characterises the effects of key uncertainties on projected outcomes, and assesses the evidence for and against alternative hypotheses for these key uncertainties. The purpose of the PATH Weight of Evidence process is to build on these insights by using expert judgement to assess the relative degree of belief in key alternative hypotheses about the effects of management actions on the performance of spring/summer chinook. PATH is analyzing the effects of management actions on fall chinook, but this report focuses only on spring/summer chinook. Other hydro system actions remain to be examined (i.e., B1 – Snake River and John Day drawdown; A2' – transportation with surface collectors), and some recently proposed hypotheses have not yet been fully examined or modeled. The impacts of uncertainties related to the management of hatcheries, habitat and harvest are also not fully explored in this document.

It is important to recognize that weighing alternative hypotheses does not provide the same types of insights as strong management experiments. Instead, the Weight of Evidence process provides a way to quantitatively incorporate the best available scientific information into decisions that must be made now, before additional evidence is available. This process is not intended to replace research, monitoring, or experimental management actions that could, if carried out properly, produce data that would in time narrow these uncertainties further. The process does however provide a more credible basis for designing adaptive management actions, and evaluating their risks and benefits relative to other proposed management actions. Our final report will examine the risks, benefits and opportunities for learning provided by each action.

Time and logistical constraints require that the Weight of Evidence process be as streamlined as possible. To that end, we have developed a structured process for exchanging and synthesizing information, with **strict deadlines** (Table 1-1). We also have focused the analyses on key uncertainties, as identified through detailed sensitivity analyses. The Weight of Evidence process is a closed process within the PATH group and external review panels (no public consultation). The final PATH report will of course be made available to the public.

Table 1-1: Summary and schedule for Weight of Evidence approach.

Step	Task	Who (Lead)	When
1	Form weighting panels and set up schedule for entire process.	PATH Planning Group, Barnthouse, McConnaha, PATH Planning Group, Promislow	April 6-May 15
2	Sensitivity analysis to refine questions as much as possible.	Peters, Deriso, Hinrichsen, Paulsen, with support from passage modelers	March 30 – May 15
3	Short period of time to allow 3-5 page summary of new hypotheses, and associated evidence.	Marmorek, PATH Planning Group (can be delegated)	April 6-May 31 –draft June 5 - final
4	Prepare list of sources of evidence.	Marmorek and Peters	April 6 - May 31
5	Identify critical components of evidence.	Marmorek and Peters, PATH Planning Group	April 6 - May 31
6	Compile first draft of Evidence Report.	Marmorek and Peters	May 31 - July 3
7	Review of draft Evidence Report by PATH participants. Specific questions passed on to subset of ISAB members for arms-length comment. Rerun models and compare to data as required.	Marmorek, Peters and Barnthouse	June 5 – July 27 (review) July 27 (comments distributed) July 30-Aug. 21 st (meetings/ revisions). Aug. 21 st : distribution to SRP
8	Weighting panels assign weights	Marmorek, Peters, Gregory, Barnthouse, SRP members	Sept. 8 – Sept. 10th
9	Documentation		Sept. 15 – Sept. 30

1.2 Contents of this Report

This report covers steps two to six of the Weight of Evidence approach in Table 1-1. The audience for this report are PATH participants (who have been involved in its refinement through two draft versions), and the SRP who will use it as a reference source in assigning weights. The report draws from all previous PATH reports and SRP reviews, as well as other literature, though by no means exhaustively. We have tried to cite the key points from past PATH reports rather than repeat their contents, so as to keep a reasonable length to this document. Submissions for previous drafts of this report are included separately. The main body of this report summarizes the key points of these submissions in a structured format, and refers to specific sections of the submissions for more detail.

We anticipate that this report will be summarised and referred to in the final report on Spring/Summer Chinook. We recommend that the final Spring/Summer report be concise and oriented towards decision makers, describing the potential benefits, risks, and opportunities for learning associated with each of the management options. The final report will have a much wider audience than this report. We stress that this report is only one step in PATH’s progress toward a final report; it is not the final report. Further work is required by PATH to assess the opportunities for learning associated with each option, and adaptive management options; this work will be a major PATH priority in the coming year.

The structure of the document is as follows. Step 2 (a sensitivity analysis to refine questions) is summarised in Section 3, with additional information in Appendices A to E. The objective of the sensitivity analysis is to determine which hypotheses are most critical in determining: 1) the ability of the actions to meet jeopardy standards, and 2) the relative ranking of actions. The evidence for and against the key (previously proposed) hypotheses is summarized in Section 4. Various submissions included in the Appendices (and referenced in section 4) describe this evidence in more detail.

The sensitivity analysis has taken much longer than anticipated for several reasons. CRiSP runs had to be re-done to incorporate the new set of CRiSP/T4 D values that were submitted in January, and the corrected set of CRiSP A3 Vn (in-river survival) estimates that were submitted in May. Also in May, we discovered an error in the PAPA flow input file, and had to re-run all of the Alpha model runs. The sensitivity analysis is based on these new results, which are quite different from those in the preliminary report (that is for the CRiSP and associated transportation assumptions). We have attempted not only to understand which assumptions are driving the new results, but also the reasons for changes from the preliminary results.

Several new hypotheses have been submitted under Step 3 of the WOE process. The new extra mortality hypotheses are summarized in Section 4 (together with previously proposed extra mortality hypotheses), while the remaining new hypotheses are summarized in Section 5. Details of the new hypotheses are included in various submissions, which are referenced in Sections 4 and 5. None of the new hypotheses has been quantified sufficiently to be run through the passage models and/or BSM. However, the rationale and some supporting evidence has been provided in all cases. We have included some preliminary analyses of evidence and, where feasible, sensitivity analyses to assess the degree to which these new hypotheses could affect the results. These analyses of evidence and sensitivity analyses should be used as a guide to determining which new hypotheses should be included in the final decision analysis for spring/summer chinook. We recognize the importance of maintaining the balance of hypotheses in the decision analysis; that is, new hypotheses which are similar to existing hypotheses should replace them, rather than diluting the analysis with several similar hypotheses. It is intended that the Weighting Panel, in addition to assessing the relative likelihood of existing hypotheses in the decision analysis, can provide PATH with some guidance on the relative strengths and weaknesses of new hypotheses. Guidance is also welcome on the relative priority of these new hypotheses in future investigations, and the logical way to incorporate them into the analysis (e.g., as sensitivity analyses or as branches of the decision tree).

Finally, in recognition of the jargon-filled nature of PATH documents, we have included a glossary of acronyms, terms, variable names and parameters. We hope this eases your journey through this report.

2.0 Criteria for Weight of Evidence Assessment

The models used in PATH involve necessary simplifications of the real world (i.e., the areas at the top and bottom of Figure 2-1). Examples of processes not explicitly modeled are changes in the life history and genetic diversity of populations, changes in estuarine habitat, and interactions with hatchery fish. Due to data limitations, the analysis focuses on factors which are believed to have affected Snake River spring/summer chinook since the early 1950s. It is neither feasible nor desirable to attempt to model all processes which can currently affect fish populations, or may have affected them over the last century. In this assessment of evidence, however, we need to consider if some of these non-modeled factors and data limitations might significantly affect the outcome of the modeling analysis (i.e., alter the relative or absolute abilities of each action to meet jeopardy standards). The sensitivity analysis (Section 3) describes the key assumptions which have the greatest impact on the preferred management actions. The Weight of Evidence Assessment seeks to determine the credibility of alternative sets of assumptions. In particular, are some of the key assumptions in the models unrealistic, or do they lead to projections which are unrealistic given both past experience and our best scientific judgement (i.e., the zones at the left and right of Figure 2-1)?

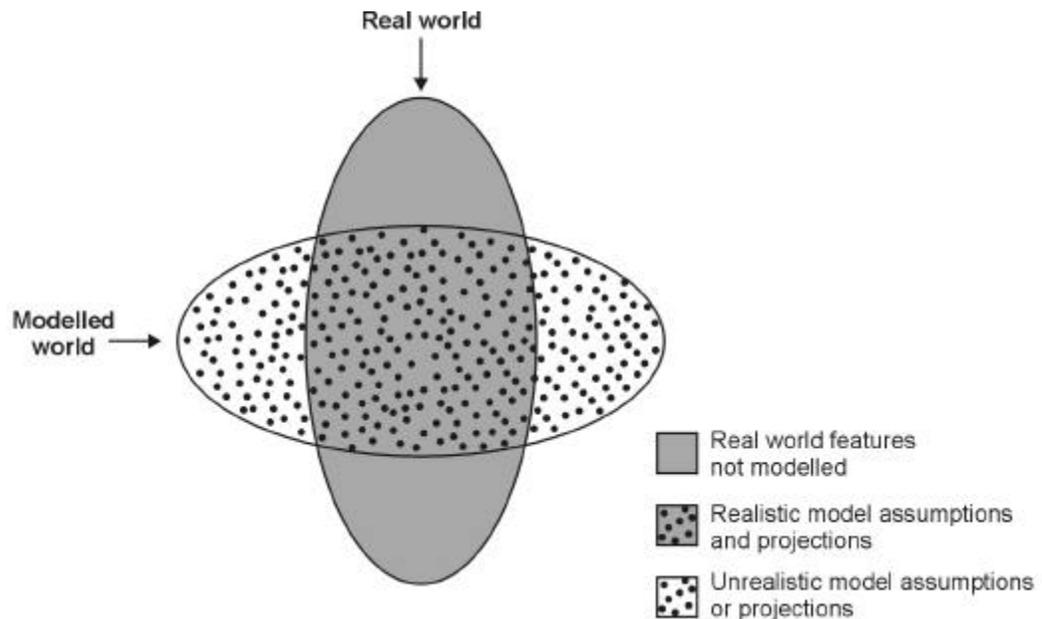


Figure 2-1: Assessing the credibility of model projections and assumptions by looking at their consistency with empirical evidence from the real world.

There are two sets of criteria to be specified:

- A) *Overall criteria* for assessing aggregate or component hypotheses; and
- B) *Evidential criteria* for assessing the applicability, clarity, and rigor of individual pieces of evidence potentially relevant to particular aggregate or component hypotheses.

We propose four *overall criteria* for assessing aggregate and component hypotheses. *Evidential criteria* are implicitly considered for criterion 2, and explicitly considered for criterion 3.

Criterion 1) the clarity of the hypothesis

The intent of this criterion is to assess whether the hypothesis as implemented in various models clearly represents the effects that are intended. Are the stressors claimed to be driving the hypotheses represented by state variables which reflect those stressors and not other stressors that are part of competing hypotheses? The clarity criterion does not favour single factor hypotheses. But where multiple factor hypotheses are proposed, they should be structured in a way that clearly separates the effects of different factors.

Criterion 2) the existence of a reasonable mechanism or set of mechanisms by which the hypothesis operates

The hypothesis must propose a reasonable mechanism by which a given stress is converted into a change in survival. There should be evidence from physiological studies or direct survival measurements in the field to clearly associate a proposed stress with a reduction in survival. See evidential criteria in Table 2-1.

Criterion 3) The consistency with empirical evidence

Measures of stock performance should vary inversely with the magnitude of the stressor across contrasts in space and time. Various measures of stock performance should be examined. These may include: recruits per spawner; smolt to adult return rates (SARs); productivity (Ricker 'a'); transport-to-control ratios; and various survival measurements (dam/bypass survival; recruitment anomalies from stock-recruitment curves, reach survival estimates, etc.).

We repeat Randall Peterman's caution here that comparison of models and hypotheses to historical data is not necessarily a strong test of how those models will perform in the future (see Appendix A of Peters et al. 1997, in FY97 report). This is because in most cases, the actions that we are evaluating will create conditions that are outside of the range of historical observations. That is, while consistency with historical conditions is a necessary condition for a hypothesis to be considered, such consistency is not a sufficient condition for that hypothesis to be a reasonable predictor of future conditions.

However, comparisons with historical data can still provide useful information for qualitative assessments of how reasonable (or unreasonable) is it for nature to behave in a way that would make this hypothesis valid. This involves assessing the degree to which the hypothesis has consistent impacts on survival in different times and places, all other things being equal. The degree of consistency of a given aggregate hypothesis with empirical data will vary with the type of statistical analyses used.

Box 2-1. Evidential criteria, used implicitly with *overall criterion 2* (mechanism) and explicitly with *overall criterion 3* (consistency with empirical evidence). For criterion 3, each of the three evidential criteria is assessed on a 4-point scale, where '1' is best, and '4' is worst; sometimes a range is provided.

- i) *Applicability*: Is the evidence relevant to the hypothesis being evaluated (i.e., is it the right stock, monitored in the right place at the right time)?
- ii) *Clarity*: Is the evidence clear, and not contested or confounded by other information, or an absence of sufficient good quality measurements?
- iii) *Rigor*: Is the evidence: 1) well established, generally accepted, peer reviewed¹ empirical evidence from relevant experiments and observations; 2) strong evidence but not fully conclusive; 3) theoretical support with some evidence; or 4) speculation or conjecture?

Criterion 4) the validity of the method of projecting the hypothesis into the future

The algebraic formulations used in the models apply relative changes to particular water years in the form of ratios of future to past values of certain survivals (e.g., system survival (ω), post-Bonneville survival (λ_n), and in river survival (V_n)) for either certain water years or time-averaged periods. We need to assess whether the assumptions implied in these methods of projection are reasonable. For example, do the methods properly reflect the hypotheses and mechanisms they were intended to represent? Do the methods generate reasonable escapement estimates under scenarios which extend current operations into the future? In addition, we need to assess whether the input variables used retrospectively to calibrate a particular model are reasonably simulated when that same model is projected into the future.

As indicated above, we have assigned ratings (sometimes a range) to the evidential criteria, but have not assigned ratings to the overall criteria. We leave the judgements of alternative hypotheses against the overall criteria (or other criteria) to the Weight of Evidence panelists as they assign relative weights to alternative hypotheses.

¹ Peer review is recognized as having its limitations with respect to scrutiny of models (Starr et al. 1998).

3.0 Results and Sensitivity Analyses

The purpose of this section of the Weight of Evidence Report is to report on the key findings of our sensitivity analyses of PATH modeling results. The purpose of these analyses was to identify which of the various hypotheses were the most critical in determining:

1. the ability of management actions to meet the NMFS jeopardy standards. The NMFS Jeopardy Standards are defined in Appendix D of the Preliminary Decision Analysis Report.
2. the relative ranking of management actions.

This section describes only the key results of the sensitivity analysis; more details are provided in Appendix A. Section 4 lays out evidence for and against the key hypotheses identified in Section 3.

Before presenting the results, we provide a brief description of the modeling approach used. This approach was described in detail in the PATH Preliminary Decision Analysis Report.

3.1 Brief Description of Modeling Approach

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 (Beamesderfer et al. 1997) and environmental data (e.g., climate indicators) are used for calibration of the life cycle models' stock production functions and other parameters (Box 5). 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. 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, Table 3-1). 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). The hydrosystem operating requirements associated with each option are described in Appendix C of the Preliminary Decision Analysis Document.
- the expected flows associated with each 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 are evaluated by simulating their consequences using a linked set of models in a four-step process to generate performance measures:

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 compute the improvement in survival relative to the retrospective period. 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.
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 to that of 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).

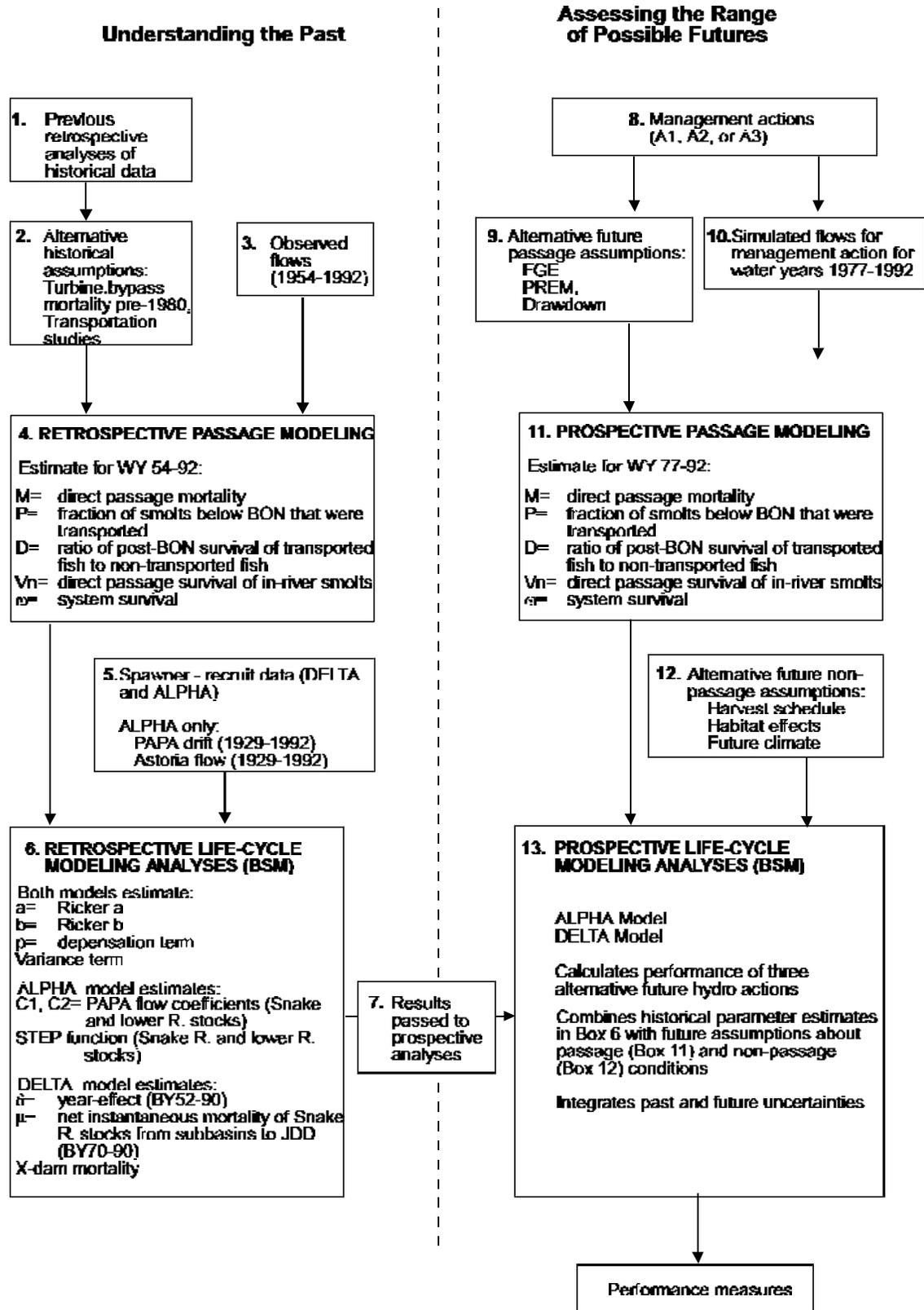


Figure 3-1: Diagram of analytical approach used in the decision analysis.

The set of actions currently under consideration for managing the hydrosystem is shown in Table 3-1. 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.

Table 3-1: Hydro system management actions currently under consideration.

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)			
A3	X	X	Natural River	-

There are many uncertainties that can potentially affect the responses of fish populations to management actions. We have focused on eleven important uncertainties, and have laid out a range of alternative hypotheses for each (alternative hypotheses are defined in Table 3-2). 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. *Life-cycle model*– uncertainty in the extent to which Snake River and lower Columbia stocks share common mortality effects.
7. *Extra mortality/Future climate* – Extra mortality is any mortality occurring outside of the juvenile migration corridor that is not accounted for by either: 1) productivity parameters in spawner-recruit relationships; 2) estimates of direct mortality within the migration corridor (from passage models); or 3) for the delta model only, common year effects affecting both Snake River and Lower Columbia

River stocks. Extra mortality can in theory occur either before or after the hydropower migration corridor.. Uncertainty in future climate relates to future patterns in climatic conditions. Extra mortality and future climate are coupled because they are closely linked with one another.

8. *Habitat effects* – uncertainty in the biological effects of future habitat management actions.
9. *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.
10. *Length of Transition Period* – duration of period between completion of dam removal and establishment of equilibrium conditions in the drawdown 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).
11. *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).

Uncertainties 9, 10, and 11 only apply when projecting the effects of drawdown to natural river of the four lower Snake River dams (option A3):

Table 3-2: 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	PMOD1	CRISP estimates of in-river survival (Vn) and proportion transported
	PMOD2	FLUSH estimates of in-river survival (Vn) and proportion transported
Fish Guidance Efficiency (FGE)	FGE1	FGE w/ESBS > FGE w/STS (values depend on project) (ESBS = extended length submersible bar screens). (STS = standard length submersible travel screens). e.g., LGR 1996-1997: FGE1 = 78%
	FGE2	FGE w/ESBS = FGE w/STS. e.g., LGR 1996-1997: FGE2 = 55%
Historical Turbine + Bypass Survival	TURB1	Turbine survival = 0.9. Bypass survival = 0.97 - 0.99, depending on the project. Used for post-1980 years in all runs.
	TURB4	Highest pre-1980 mortality; turbine and bypass mortality are due to descaling alone
	TURB5	Lowest pre-1980 mortality; turbine mortality rate = half descaling rate; bypass mortality rate = descaling rate
	TURB6	Medium pre-1980 mortality; turbine mortality = 0.9; bypass mortality rate at LGR, LGO = descaling rate
Predator removal efficiency	PREM1	0% reduction in reservoir mortality resulting from predator removal program.
	PREM3	25% reduction in reservoir mortality.
Duration of pre-removal period under drawdown	PRER1	3 years
	PRER2	8 years

Uncertainty (Report section with further details)	Hypothesis Label	Description
Equilibrated Snake River juvenile survival rate under drawdown	EJUV1	Survival rate through drawdown reach =0.85
	EJUV2	Survival rate through drawdown reach = 0.96
Transition Period: Juvenile survival	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	TRANS1 or T1 (FLUSH only)	Relationship between TCR and survival of control fish, 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). Prospective and retrospective D values lower than TRANS4 and relatively constant between pre-1980 and post-1980 periods.
	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.
	TRANS4 or T4 (CRISP only)	For pre-1980 retrospective analyses, relative post-BONN survival set at median D-value estimated from seven T:C studies in 1970's and associated CRISP in-river survival rate estimates. Post-1980 retrospective analyses use median 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 post-1980 values. Prospective and retrospective D values higher than TRANS1 and 2; increase in retrospective D values after 1980 reflects improved transport conditions.
Life-cycle model	PROSPA ALPHA	- Extra mortality is specific to each sub-region, and affected by climate variables.
	PROSPD DELTA	- Extra mortality is independent of the common year effects which affect several subregions.
Extra mortality / Future climate	EMCLIM1	Extra mortality is here to stay; prospective D values selected randomly from post-1980 values; future climate is sampled from historical distribution with autoregressive properties.
	EMCLIM2	Extra mortality is here to stay; prospective D values selected randomly from post-1980 values; future climate follows cyclical pattern.
	EMCLIM3	Extra mortality is proportional to hydropower-related mortality, with a different proportionality coefficient in each year. Prospective D values are selected according to water year. Future climate is sampled from historical distribution with autoregressive properties.
	EMCLIM4	Extra mortality is proportional to hydropower-related mortality, with a different proportionality coefficient in each year. Prospective D values are selected according to water year. 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. Prospective D values selected randomly from post-1980 values
Habitat Effects	HAB0	Same management as current.
	HABB	Implementation of all possible habitat restoration or protection.

3.2 Changes Since the March 1998 Preliminary Decision Analysis Report

A number of significant changes in the analysis have occurred since the Preliminary Decision Analysis Report for Spring/Summer Chinook was completed in March 1998. These changes and their implications are described in more detail in Appendix B.

Changes since the Preliminary Decision Analysis Report:

- CRiSP D values (ratio of post-Bonneville survival of transported fish: non-transported fish) have decreased based on Anderson 1998.
- CRiSP in-river survival rates of non-transported fish (V_n) under A3 have increased.
- Uses corrected set of Astoria flows. Previous results used the wrong values.
- Addresses imbalanced design and associated problems (described by Hinrichsen and Paulsen 1998).
- We have omitted HAB B hypothesis (effects of maximum protection / enhancement) because of problems with implementation in the preliminary draft (see Submission 4).

These changes have affected overall results from those presented in the Preliminary Decision Analysis Report (see Appendix B). They have also affected the relative importance of various hypotheses in terms of their effects on the projected responses of stocks to management actions.

3.3 Hypotheses that Affect the Ability of Actions to Meet the Standards

The proportion of runs in which all standards are met are shown in Table 3-3. 85 out of 240 A1 runs (35%) meet all of the jeopardy standards, 35% of A2 runs, 63% of A3 runs with 3-year delay, and 47% of A3 runs with 8-year delay. Passage/transportation assumptions affect the ability of actions to meet all of the standards - CRiSP results make up the majority of the runs that meet all of the standards for all actions.

Table 3-3: Analysis of runs meeting all jeopardy standards – overall results

Action	# of runs for each action	# of runs (%) which meet all of the jeopardy standards	Proportion of runs meeting the standards that are CRiSP/T4	Proportion of runs meeting the standards that are FLUSH/T1
A1	240	85 (35%)	0.85	0.15
A2	240	85 (35%)	0.86	0.14
A3 (3-year delay)	960	605 (63%)	0.64	0.36
A3 (8-year delay)	960	453 (47%)	0.75	0.25

We used two approaches for determining which hypotheses affect the ability to meet the standards. First, we looked at the contribution of individual hypotheses to the runs that met all of the standards (0.7 for survival standards, 0.5 for recovery standards). If a hypothesis has a strong effect on the results, its proportional representation in the runs that meet the standards will be different from its proportional representation in all of the runs. Detailed results for this analysis are shown in Tables A-1 and A-2 in Appendix A. The second approach was to subject the model results to a Regression Tree analysis to determine which factors accounted for most of the variability in jeopardy standard results. The details of this analysis are provided in Appendix E; regression trees for the 3 jeopardy standards are shown in Figures 3-2 to 3-4. The left side of the tree represents hypotheses that result in lower 24-year survival

probabilities; the right side represents hypotheses that result in higher jeopardy probabilities (average probabilities for a particular “branch” of the tree are shown at the bottom of the branch). Length of the vertical branches is proportional to the proportion of variance in results explained by that hypothesis. Interestingly, the actions themselves generally account for a small proportion of the variance in results for the 24-year survival standard (Figure 3-2). For some combinations of factors (e.g., Hydro extra mortality, CRiSP/T4 passage/transportation model, delta model), the effects of the actions are not significant at all. However, with the 100-year survival standard and the 48-year recovery standard the split between A1/A2 and A3 accounted for the majority of the variance (Figure 3-3, 3-4).

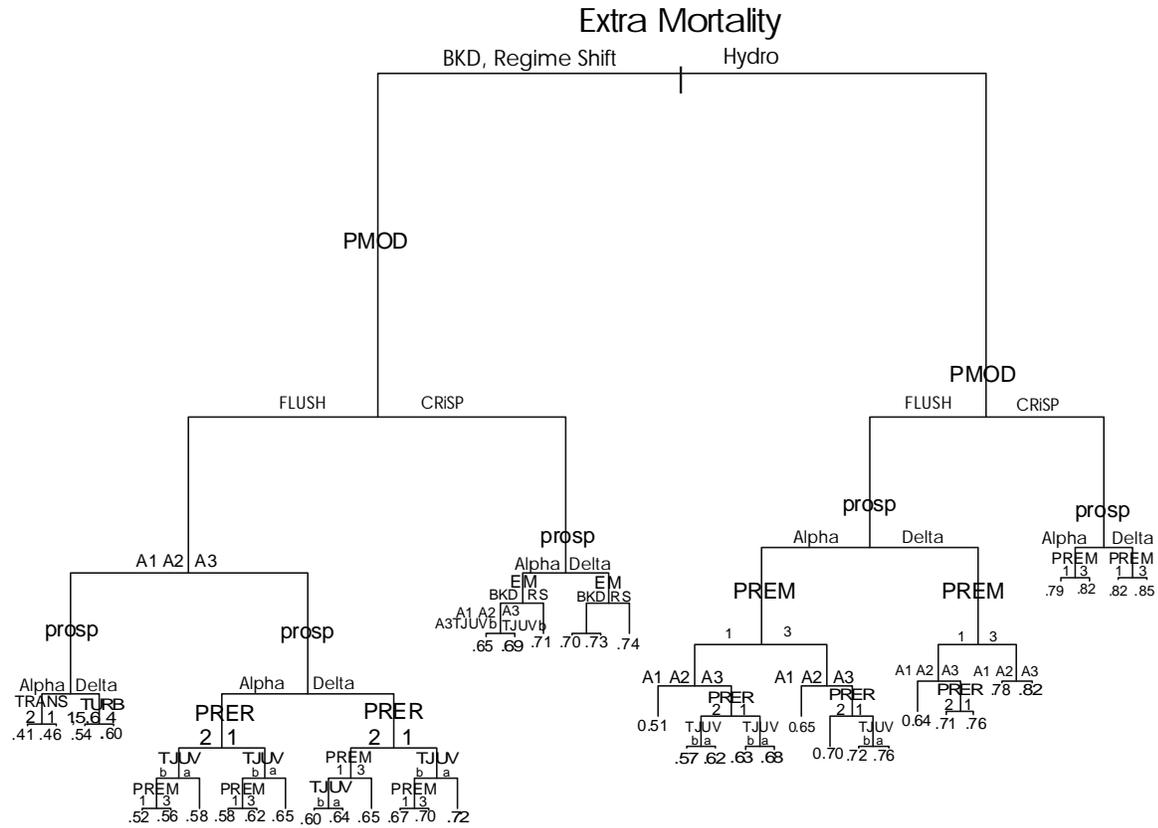


Figure 3-2: Regression tree for the 24-year survival probability for all factors, truncated at R-squared > 0.95 (37 splits).

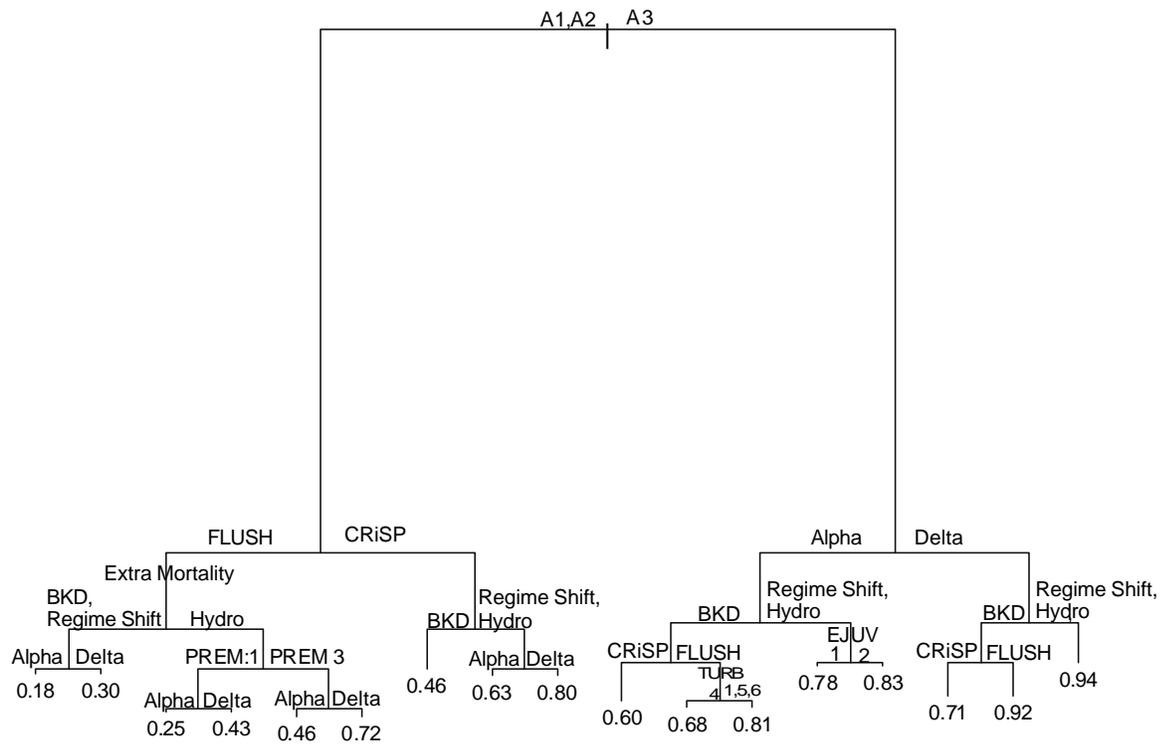


Figure 3-4: Regression tree for the 48-year recovery probability for all factors, truncated at R-squared > 0.95 (16 splits).

3.4 Hypotheses that Affect the Relative Ranking of Actions

This section describes the hypotheses that determine which action is preferred over the others. We use the term “preferred” or “favored” here NOT in the context of providing advice to managers on which action should be implemented, but in the context of assessing the sensitivity of ranking of actions to alternative hypotheses.

Overall results suggest that there are certain combinations of hypotheses for which A3 is favored over A1 and A2, while other combinations of hypotheses results in the three actions having similar results (Table 3-4). The passage models and their associated transportation models are strongly associated with this distinction. CRiSP/T4 tends to show little difference among actions for the 24-year survival standard, ranks A3 over A1 or A2 in a larger proportion of runs for the 100-year survival standard, and ranks A3 over A1 or A2 in most runs for the 48-year recovery standard. FLUSH/T1 tends to rank A3 over the other actions regardless of which jeopardy standard is used as a criterion. A sensitivity analysis of the margin by which A3 exceeds A1/A2 is shown in Table 3-4; the cumulative probability distributions in Figure A-3 also show these differences.

Table 3-4: Analysis of runs by which action has a higher jeopardy probability than the other actions, and by how much. Runs are broken down by passage/transportation model and by jeopardy standard. Margins of difference of 0.02, 0.06, and 0.10 were arbitrarily chosen as sensitivities.

# (%) of runs where:	CRiSP / T4			FLUSH / T1		
	24-year survival	100-year survival	48-year recovery	24-year survival	100-year survival	48-year recovery
A1/A2>=A3	106 (11%)	1 (0.1%)	6 (0.6%)	8 (1%)	0	0
A1/A2 exceeds A3 by 0.02 or greater	5 (0.5%)	0	2 (0.2%)	0	0	0
A1/A2 exceeds A3 by 0.06 or greater	0	0	0	0	0	0
A3>A1/A2	854 (89%)	959 (99.9%)	954 (99.4%)	952 (99%)	960	960
A3 exceeds A1/A2 by 0.02 or greater	463 (48%)	818 (85%)	946 (99%)	887 (92%)	960	960
A3 exceeds A1/A2 by 0.06 or greater	34 (4%)	156 (16%)	912 (95%)	645 (67%)	871 (91%)	960
A3 exceeds A1/A2 by 0.10 or greater	1 (0.1%)	7 (0.7%)	787 (82%)	352 (37%)	805 (84%)	960

For the purposes of the weight of evidence report, we are primarily interested in those hypotheses that result in clear differences in jeopardy probabilities between actions. To do this, we looked at the 106 combinations of hypotheses within CRiSP / T4 that resulted in A1 or A2 being equal or preferred to A3 using the 24-year survival probability (Table 3-5). We looked in detail at this particular jeopardy standard because it was the only one in which A1 or A2 was preferred over A3. Column 1 of this table shows the proportion of all CRiSP runs that include each hypothesis. Column 2 shows the proportion of the runs in which the 24-year survival probability for A1 or A2 \geq A3. Column 3 shows the percentage change in Column 2 relative to Column 1 as a means of scaling and comparing results for each hypothesis. Large changes in the proportion of runs implies that the hypothesis has a significant influence on the ranking of actions. Column 4 ranks the hypotheses in terms of their effects on the ranking of actions, as measured by the size of the change in their proportion of runs.

This table is useful for identifying the effects of single factors on the results, but does not show effects of interactions between factors. For example, the length of the transition period following drawdown (TJUV hypotheses) may only be important if coupled with a particular equilibrated juvenile survival rate (EJUV) hypothesis. Tables A-13 to A-15 in Appendix A explore the effects of combinations of hypotheses. Overall, the results presented in Appendix A suggest that the effects of those factors identified in Table 3-5 as being the most important (i.e., length of transition period, extra mortality hypotheses) are not dependent on other factors. For example, the length of the transition period in particular has a very strong effect (TJUVb – the 10-year transition period - makes up over 90% of the 106 runs in which A1 or A2 was preferred over A3) regardless of other factors. That is, using TJUVb instead of TJUVa lowers the 24-year survival probabilities of A3 sufficiently to be slightly below those of A2.

Table 3-5: Sensitivity analysis of hypotheses that lead to 24-year survival probabilities $A1$ or $A2 \geq A3$ using CRiSP. See Table 3-1 for more details on the individual hypotheses.

<i>Hypothesis</i>	Description	Proportion of all CRiSP runs	Proportion of 106 runs in which $\max(A1 \text{ or } A2) \geq A3$ (24-yr survival std)	% change from proportion of all CRiSP runs	Rank Order of Sensitivity
FGE 1	Higher FGE	0.5	0.69	+37.7	6
FGE 2	Lower FGE	0.5	0.31	-37.7	
TURB1	Standard turbine/bypass survival	0.333	0.27	-17.8	4
TURB4	Lower turbine/bypass survival	0.333	0.25	-26.3	
TURB5	Higher turbine/bypass survival	0.333	0.48	+44.5	
PREM 1	Predator removal ineffective	0.5	0.66	+32.1	7
PREM 3	Predator removal effective	0.5	0.34	-32.1	
Alpha	Extra mortality specific to sub-regions, no common year effects	0.5	0.65	+30.2	8
Delta	Common year effects affect several subregions	0.5	0.35	-30.2	
BKD/Markov	Extra mortality of non-transported fish is here to stay; future climate sampled from historical conditions with autoregressive properties	0.2	0.17	-15.1	2
Hydro/Markov	Extra mortality of non-transported fish proportional to survival in migration corridor; Markov climate	0.2	0.17	-15.1	
BKD/Cyclical	BKD extra mortality; future climate follows cyclical pattern	0.2	0.17	-15.1	
Hydro/Cyclical	Hydro extra mortality; Cyclical future climate	0.2	0.13	-34.0	
Regime Shift/Cyclical	Both extra mortality and climate follow cyclical pattern	0.2	0.36	+79.2	
PRER 1	3 years between time decision is made and initiation of dam removal	0.5	0.30	-39.6	5
PRER 2	8 years between time decision is made and initiation of dam removal	0.5	0.70	+39.6	
EJUV 1	Juvenile survival rate equilibrates following drawdown at 0.85	0.5	0.73	+45.3	3
EJUV 2	Juvenile survival rate equilibrates following drawdown at 0.96	0.5	0.27	-45.3	
TJUV a	Juvenile survival rates equilibrate 2 years after drawdown	0.5	0.08	-84.9	1
TJUV b	Juvenile survival rates equilibrate 10 years after drawdown	0.5	0.92	+84.9	

3.5 Implications for the Weight of Evidence Process

The above results show that a subset of the uncertainties described in Table 3-2 are highly influential in determining the relative ranking of actions and the ability of actions to meet the standards. The relative importance of these hypotheses is summarized in Table 3-6. The evidence in Section 4 is focussed primarily on those hypotheses that have “High” relative importance for either the ability of actions to meet the standards or the relative ranking of actions.

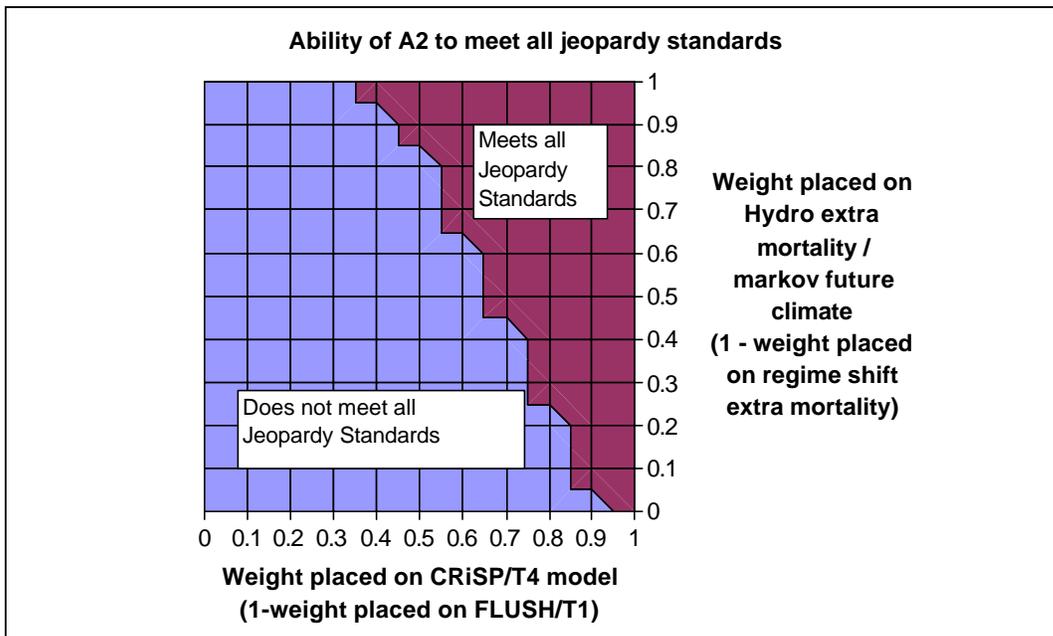
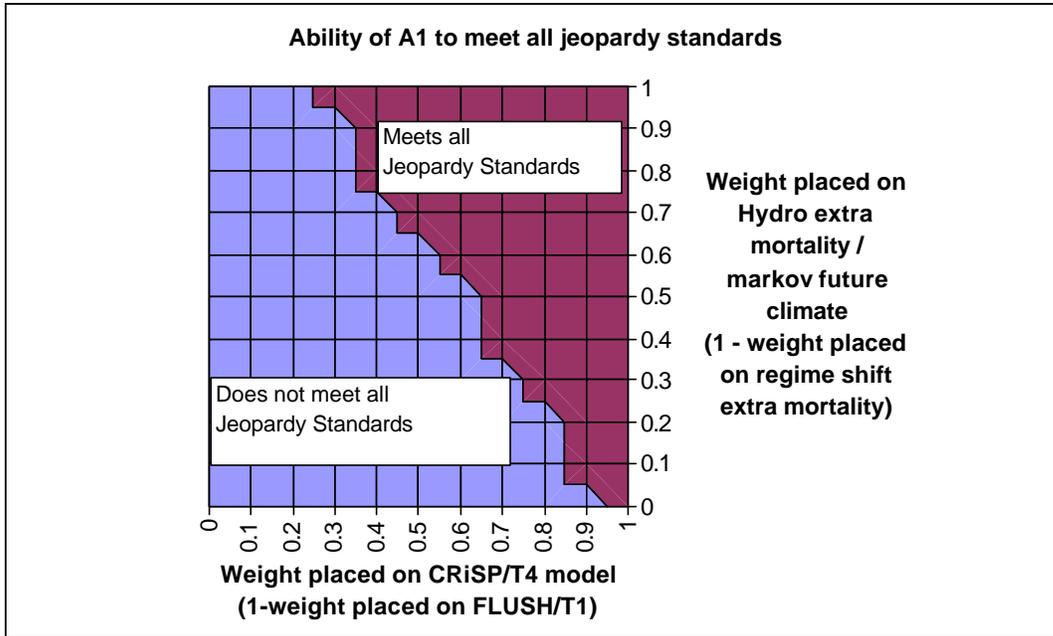
Table 3-6: Summary of relative importance of alternative hypotheses. The three most important hypotheses are in each column are ranked “high”, the middle three are ranked “moderate”, and the bottom three are ranked “low”. Rankings are based on results presented in Tables 3-3, A-6, A-7, 3-4, and 3-5.

	Relative importance of each hypothesis in determining:	
	Ability of actions to meet the standards	Relative ranking of actions
Passage/Transportation Models	High	High
FGE	Low	Low
Historical turbine/bypass mortality (TURB)	Moderate	Moderate
Predator Removal Effectiveness (PREM)	Moderate	Low
Pre-removal period (PRER)	Low	Moderate
Equilibrated juvenile survival rate (EJUV)	Low	Moderate
Length of transition period (TJUV)	Moderate	High
Life-cycle models	High	Low
Extra mortality hypotheses	High	High

Passage models are themselves complexes of hypotheses and sub-models, and we have determined that the most important difference between FLUSH and CRiSP in influencing the results are the models’ estimates of the survival rates of control fish (fish that are allowed to migrate in the river) in Transport:Control experiments (see Appendix D for detailed diagnostic analyses of the passage and transportation models). These survival rates are used in conjunction with the Transport:Control ratios from these studies to estimate D values ($D = \frac{\text{post-Bonneville survival of transported fish}}{\text{post-Bonneville survival of non-transported fish}}$).

The objective of the weighting panel deliberations will be to assign weights to these key uncertainties based on the evidence presented in this report. We explored the sensitivity of outcomes to these weights by looking at the sensitivity of the ability of actions to meet all of the standards to the weights placed on the different passage/transportation assumptions and the hydro vs. regime shift extra mortality hypotheses. We show these results as an illustration of the sensitivity of results to weights placed on key uncertainties. Although this is only one of many possible combinations of hypotheses, the passage model and extra mortality hypotheses were both very important in determining whether the actions met all of the standards (Table 3-3, A-6, A-7). The BKD extra mortality hypothesis is given a weight of 0 here; all other hypotheses are weighted equally in these results.

Higher weights on both the CRiSP model and the Hydro extra mortality hypothesis are required to meet all of the jeopardy standards (Figure 3-2). When higher weight is placed on CRiSP/T4 (the exact weight depends on the action), all standards are met regardless of whether the regime shift or the hydro extra mortality hypothesis is assumed. In contrast, when a high weight is placed on FLUSH/T1 the standards are not met for actions A1 and A2 regardless of the weight placed on the hydro vs. regime shift extra mortality hypothesis. A3 meets the jeopardy standards under most combinations of extra mortality and passage models if a three-year pre-removal period is assumed; if an eight-year pre-removal period is assumed, a high weight must be placed on either CRiSP/T4 or the Hydro extra mortality for all standards to be met.



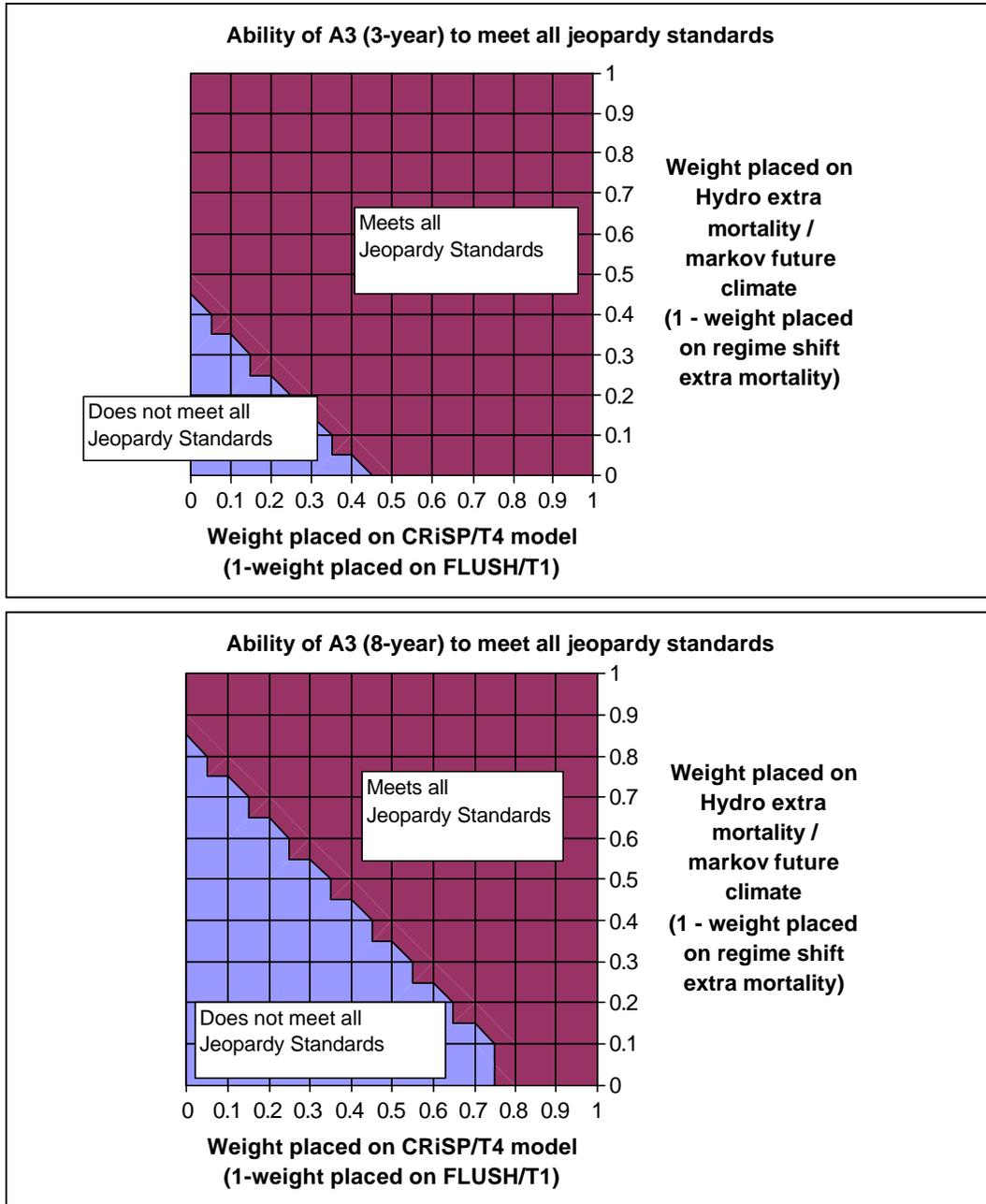


Figure 3-6: Ability of actions to meet all standards as a function of the weights placed on the passage/transportation model and the hydro vs. regime shift extra mortality hypothesis.

4.0 Evaluation of Key Hypotheses

4.1 Aggregate Hypotheses

4.1.1 Description

Analyses in Section 3 (and the Categorical Regression Trees in Appendix E) have identified subsets of hypotheses that have the most influence on: 1) the ability of actions to meet all of the jeopardy standards; and 2) the relative ranking of actions. Though these two subsets of hypotheses differ somewhat (Table 3-6), in general, the hypotheses with the most influence on results are:

1. Passage/transportation models
2. Extra mortality hypotheses
3. Life-cycle models
4. Drawdown assumptions, particularly the transition period
5. Hypotheses about predator removal effectiveness
6. Turbine/bypass survival

Combinations of these alternative hypotheses are called “aggregate hypotheses”. Many aggregate hypotheses are possible – 24 for A1 and A2, and 192 for A3². Although the weight of evidence process is focussed on evidence for and against individual hypotheses (e.g., the hydro extra mortality hypothesis vs. the regime shift hypothesis), there are a couple of reasons to also be interested in aggregate hypotheses.

The first reason to consider aggregate hypotheses is that there are some data sets that are more applicable to aggregate measures of stock performance (i.e., survival over many life cycles). The spawner-recruit data are one example. Given an estimate of historical mortality due to passage through or around the hydrosystem from the passage/transportation models, either version of the life-cycle model can be used to estimate post-Bonneville and overall spawner to recruit mortality. The delta version of the life-cycle does this by calculating the MLE of overall mortality from the historical spawner-recruit data, then calculating the post-Bonneville mortality as the difference between the MLE of overall mortality and the passage-related mortality from the passage/transportation models. The alpha version calculates post-Bonneville mortality directly, then calculates overall mortality as the sum of passage-related and post-Bonneville mortality. In the particular application of the alpha model used here, post-Bonneville mortality is computed from various climatic factors (Astoria flows, PAPA drift, and a STEP function related to oceanic regime shifts). Comparing the model-derived estimate of overall mortality to observed measures of stock performance (e.g., spawner to recruit survival rate, smolt-to-adult survival rate) indicates the degree to which each of the aggregate hypotheses fits the observed temporal and spatial patterns. Note that only hypotheses which affect retrospective simulations (i.e., passage models, life cycle models, turbine/bypass survival) can be examined against historical data.

The second reason for examining aggregate hypotheses is that it is the aggregate hypotheses which generate predictions about the future. Two aggregate hypotheses might equally well explain the past but

² A1 and A2: 2 passage/transportation models X 3 extra mortality hypotheses X 2 life-cycle models X 2 predator removal hypotheses.

A3: The same 24 as A1 and A2 X 2 pre-removal period hypotheses X 2 equilibrated juvenile survival rate hypotheses X 2 transition period hypotheses.

have very different implications for the future. Though it is difficult to assess the reasonableness of the projected impacts of actions under alternative hypotheses, particularly for actions which have not yet been experienced (e.g., A3), one can examine the projected impacts of actions that are similar to current operations (e.g., A1).

Although the primary focus of the weight of evidence process is on individual hypotheses, we present a limited evaluation of aggregate hypotheses in this section. Although this evaluation does not apply directly to evaluation of individual hypotheses, the reasonableness of aggregate hypotheses does have implications for the evaluation of their components.

4.1.2 Evaluation of Aggregate Hypotheses

Criterion 1) Clarity and 2) Mechanisms

We evaluate clarity and mechanisms on component hypotheses, rather than on aggregate hypotheses, for two reasons. First, the component hypotheses are independent of one another, so the clarity and mechanistic defensibility of the aggregates are really dependent only on the components. Second, it is impractical to evaluate 2400 aggregate hypotheses. Even the reduced set of aggregate hypotheses generated from the ‘key’ component hypotheses listed at the beginning of Section 4.1 add up to 144 combinations ($2 * 3 * 2 * 2 * 2 * 3$).

Criterion 3) Consistency with Empirical Evidence

Our evaluation is based on comparison of aggregate hypotheses to two data sets, and a qualitative assessment of the “reasonableness” of the trends in spawners that are projected under the various hypotheses. Data used to evaluate aggregate hypotheses are most useful if they are “out-of-sample”, that is they cannot have been used to develop or calibrate any component of the aggregate hypothesis. There are two empirical data sets that potentially provide an opportunity to evaluate aggregate hypotheses by comparing them to historical data. The first is the set of spawner-recruit data developed for Snake River index stocks (Beamesderfer et al. 1997). The second data set is the smolt-to-adult return rates (SARs) that describe survival of fish from the time they pass the uppermost dam as smolts to the time they return to that dam as adults.

Historical estimates of spawner-recruit survival and SARs are determined by passage / transportation models, versions of the life-cycle model, and assumptions about bypass / turbine mortality in some historical years (TURB assumptions; see Section A.2.3 in Preliminary Decision Analysis). Extra mortality hypotheses apply prospectively, not retrospectively, and therefore cannot be compared to empirical data directly. One can look, however, at the pattern of model-derived estimates of extra mortality to see if the trend over time is consistent with the hypothesized source of extra mortality (Section 4.2.2).

a) Fit of Aggregate Hypotheses to Spawner-Recruit Data

Retrospective hypotheses need to be able to explain the pattern of changes in historical spawning and recruitment. Spawning escapements showed steep declines in all 7 Snake River stocks since the early 1970’s, and no obvious patterns in the 6 lower Columbia River stocks (Figure 4-0a.). The recruits per spawner were generally similar for the two subregions before brood year 1970, and generally lower for Snake River stocks since the 1970 brood years, with the notable exception of brood years 1980 to 1983 (Figure 4-0b). Note that brood years are 3-5 years earlier than years of escapement estimates.

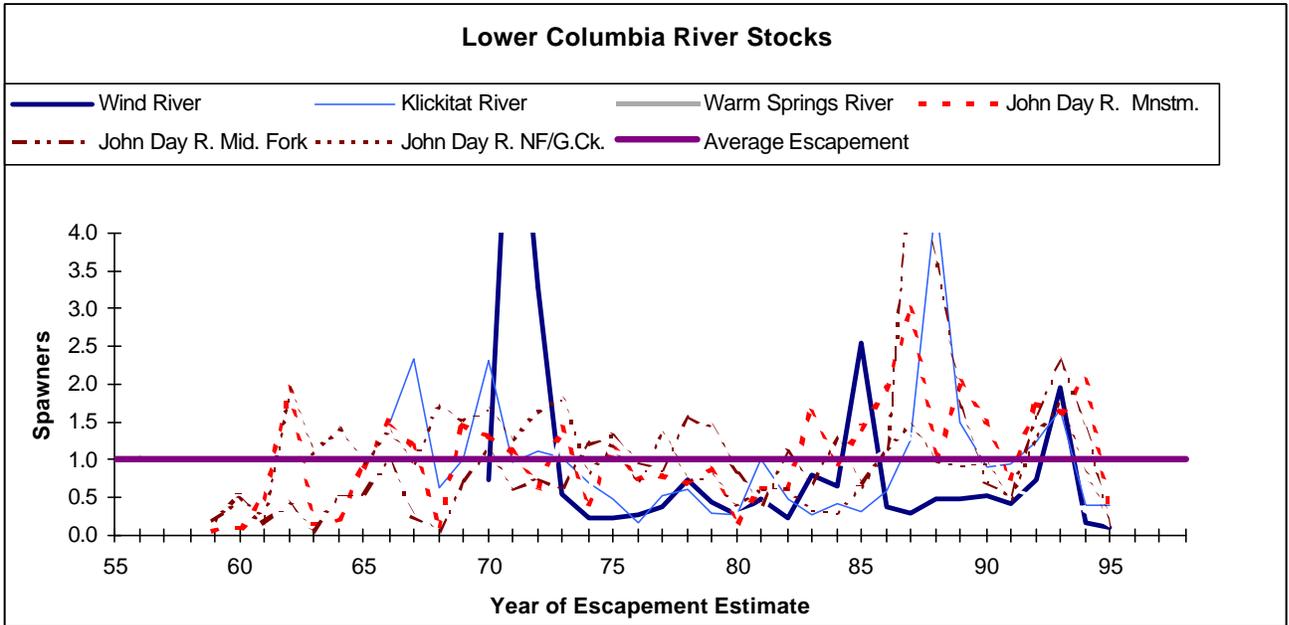
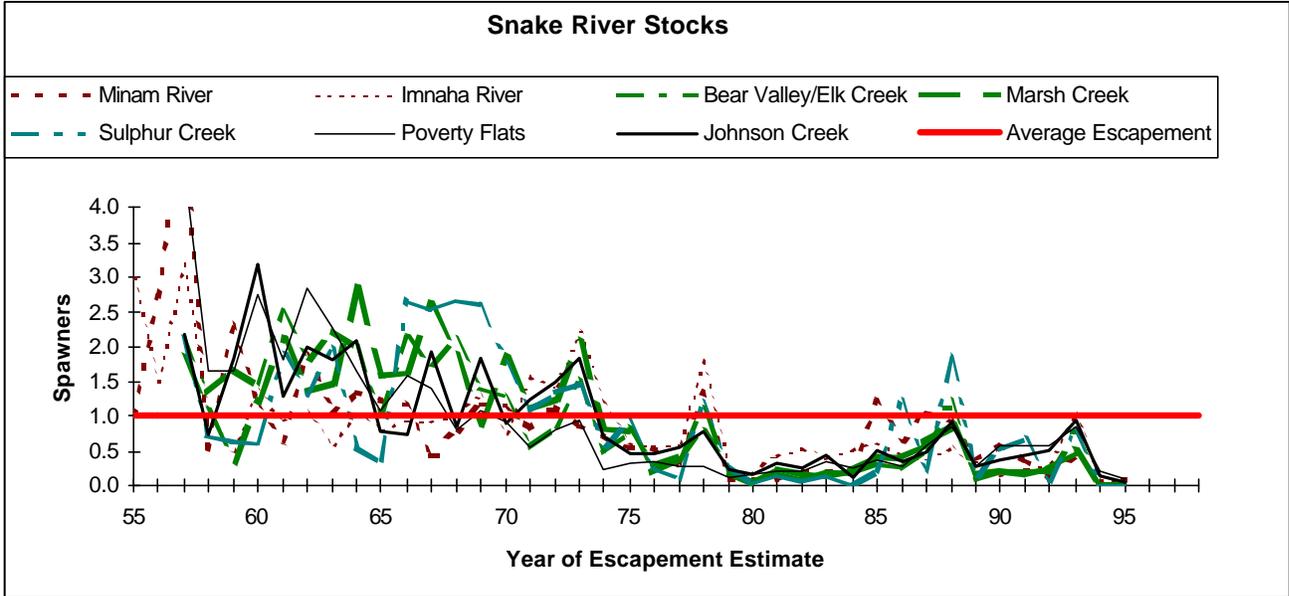


Figure 4-0a: Spawning escapements (normalized to average escapements) for seven Snake River spring/summer chinook index stocks (upper graph) and six lower Columbia River index stocks (lower graph). Source: Schaller et al. (1993).

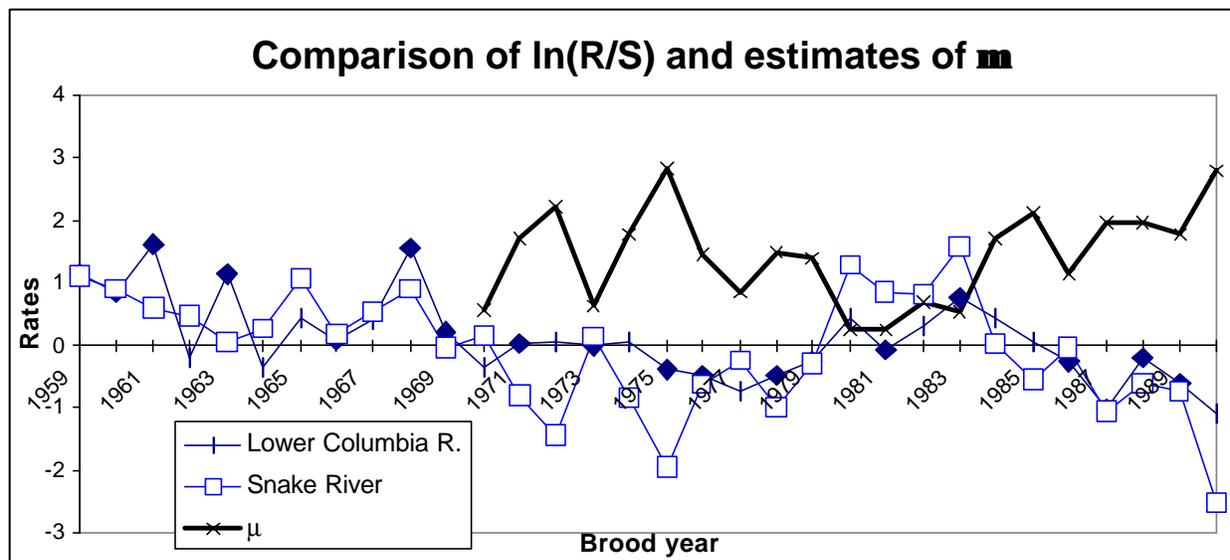


Figure 4-0b: Comparison of normalized $\ln(R/S)$ for Snake River and Lower Columbia River index stocks, and MLE estimate of m (incremental life-cycle mortality of Snake River stocks to John Day dam). The variable m is roughly proportional to the Lower Columbia River line minus the Snake River line, but also takes into account stock-specific productivities. Source: Deriso et al. (1996).

Spawner to recruit survival rates estimated by specific aggregate hypotheses are directly compared to empirical measurements using AIC and BIC model comparison criteria. These criteria consider both the fit of the aggregate hypotheses to the data and the number of parameters in the models. In general, both criteria will favor simpler models (fewer parameters) over complex ones, unless the complex models provide a very substantial improvement in the fit to the data. We use this criteria not to reject one model in favor of another, but as a measure of the relative fit of different aggregate hypotheses to observed data. The theory of AIC/BIC criteria is concerned largely with providing guidance about which models have superior forecasting ability. Our applications of AIC/BIC differ from that theory in that our forecasting models involve more parameters (and assumptions) than the parameters counted in our AIC/BIC test statistics. In particular, no parameters from the passage models are counted in the AIC/BIC statistics even though prospective passage model projections involve numerous assumptions about model parameters.

(i) Applicability: Score = 2-3

The spawner-recruit data is applicable to evaluating aggregate hypotheses about declines in Snake River spring/summer chinook stocks because it is specific to the index stocks of concern, and incorporates survival through all life cycles. Also, the time series extends prior to when the Snake River dams went in, providing greater contrast in the data.

However, its applicability is limited because it is not out-of-sample; both the delta model and the alpha models are calibrated to it. Also, the data has been circulating for some time. This means that hypotheses may not be completely independent of the data, in the sense that knowledge of the data and its patterns may have influenced formulation of hypotheses. A true out-of-sample dataset would be completely unknown to developers of hypotheses and would constitute a stronger test. Therefore, these data can provide only a qualitative comparison of the relative fit of historical estimates under different aggregate hypotheses to the empirical data. In addition, the spawner-recruit data is limited to comparisons between alpha and delta variants of aggregate hypotheses, and comparisons within alpha model variants. This dataset cannot be used to compare different aggregate hypotheses with the delta version of the life-cycle model because the

delta model estimates total mortality from the data, then back-calculates the post-Bonneville component of the overall mortality. The result is that all aggregate hypotheses that are modeled with the delta model have identical fits to the spawner-recruit data.

- (ii) Clarity: Score = 2 to 3

Justification for a score of 2 (greater clarity):

Spawner-recruit data provides clear evidence for assessing the ability of aggregate hypotheses to explain past trends in total mortality.

Justification for a score of 3 (lower clarity):

Comparison of aggregate hypotheses to the spawner-recruit data is confounded somewhat because different aggregate hypotheses may allocate mortality quite differently between sources and life stages but produce the same overall rate of mortality. Therefore, survival data in individual life stages (e.g., reach survival studies) should be considered when different aggregate hypotheses produce similar estimates of overall mortality. Also, some PATH participants have raised concerns about the uncertainty and assumptions used in the run reconstructions. Previous sensitivity analyses on some of these concerns have not shown significant effects on overall results (see Appendix G); further examination of these issues is planned for next year.

- (iii) Rigor: Score = 1 to 3

Justification for a score of 1:

The spawner-recruit data has been thoroughly reviewed by PATH participants and by the PATH Scientific Review Panel. Some PATH participants have commented on particular assumptions behind the analysis, but these assumptions do not introduce significant error or bias into the modeling results (see Appendix G).

Justification for a score of 3:

Concerns about uncertainties and assumptions in the data that have been raised by some PATH participants have not been adequately dealt with. PATH scientists intend to review the remaining concerns later this year.

Results

AIC and BIC scores are shown in Table 4-1. The table shows the # of parameters, the Sum of Squares (SSQ), $\ln(\text{likelihood})$, AIC, and BIC scores for each combination of passage / transportation model, TURB assumption, and version of the life-cycle model. Lower AIC and BIC scores are better. The 86 parameters listed for the Delta model include Ricker a and b parameters for 13 stocks (i.e., 26 parameters in total), estimates of m for 20 years (see Fig. 4.0b), estimated year effects d for 38 years, and two other parameters (Deriso et al. 1996). The 32 parameters for the Alpha model include the 26 Ricker parameters, and 3 regression parameters (STEP, PAPA, Astoria flow – explained in Section 4.2.2), each estimated for both upstream and downstream subregions (total of 6) (Anderson and Hinrichsen, 1997).

Retrospective aggregate hypotheses that include the Delta model provide a better fit to the spawner-recruit data (i.e., has a lower AIC and BIC score) than aggregate hypotheses that include the Alpha model. The differences are smaller but still significant using the BIC criterion, which places more emphasis on the number of parameters in the model. The Delta model has more parameters (86) than the Alpha model aggregate (32) but fits the data better, as indicated by the lower SSQ value. Note that all Delta model aggregate hypotheses have the same AIC and BIC scores.

Comparisons of AIC and BIC scores within the alpha model can also be used to evaluate different passage/transportation assumptions (i.e., FLUSH/T1 and CRiSP/T4). In general, FLUSH/T1 has lower

AIC/BIC scores than CRiSP/T4, and TURB5 has lower AIC/BIC scores than TURB4. These assumptions are discussed in more detail in Section 4.2.1.

Table 4-1: Comparison of aggregate hypotheses to S-R data: Goodness of fit statistics.

Aggregate Hypothesis	# parameters	SSQ	ln(L)	AIC	BIC
<i>Delta Model</i>					
FLUSH/T1 TURB1	86	112.35342	-315.19	802.38	1146.93
FLUSH/T1 TURB4	86	112.35342	-315.19	802.38	1146.93
FLUSH/T1 TURB5	86	112.35342	-315.19	802.38	1146.93
CRiSP/T4 TURB1	86	112.35342	-315.19	802.38	1146.93
CRiSP/T4 TURB4	86	112.35342	-315.19	802.38	1146.93
CRiSP/T4 TURB5	86	112.35342	-315.19	802.38	1146.93
<i>Alpha Model</i>					
FLUSH/T1 TURB1	32	268.33339	-491.92	1047.84	1176.04
FLUSH/T1 TURB4	32	271.91397	-494.61	1053.22	1181.42
FLUSH/T1 TURB5	32	264.59125	-489.07	1042.14	1170.34
CRiSP/T4 TURB1	32	286.38945	-505.14	1074.28	1202.48
CRiSP/T4 TURB4	32	302.56896	-516.30	1096.59	1224.79
CRiSP/T4 TURB5	32	291.30429	-508.59	1081.19	1209.39

b) Fit of Aggregate Hypotheses to Smolt to Adult Return Data

The historical pattern of Smolt to Adult Return data (which different retrospective hypotheses attempt to explain) is shown in Figures 4-0c and 4-1. SARs were generally high in Period 1 (averaging 4%), low in Period II (except for the 1975 outmigration year), somewhat higher in Period III, and uniformly low in Period IV.

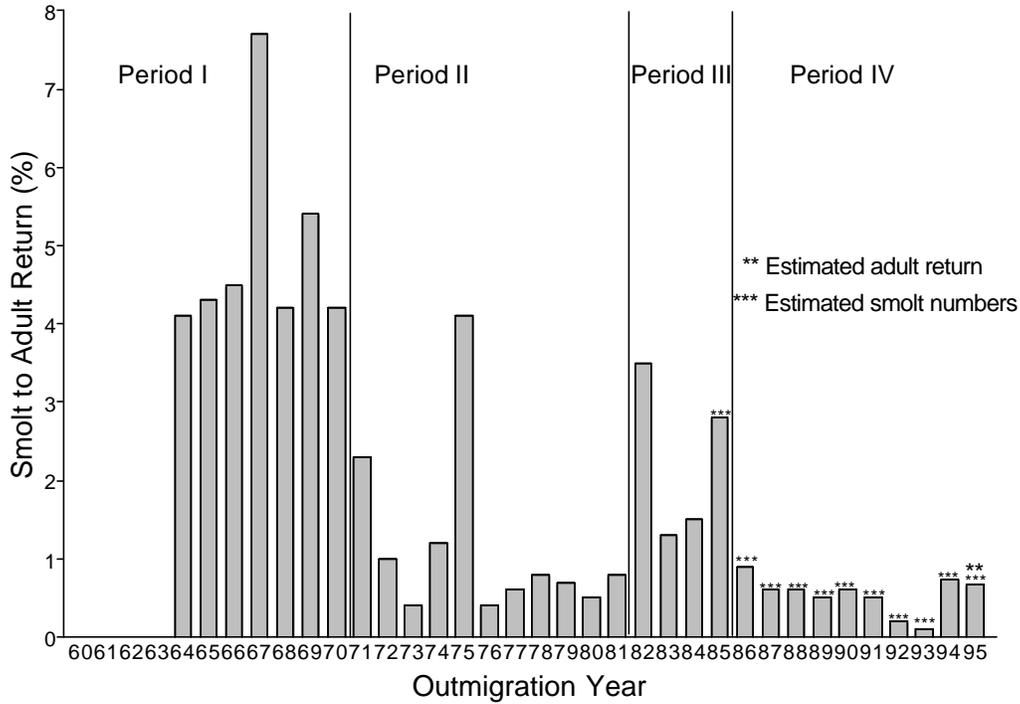


Figure 4-0c: Wild Snake River spring chinook salmon smolt to adult return (catch + escapement) rates to the upper dam on the Snake River broken out by periods. (Estimates of wild smolts from 1964 to 1984 from Raymond [1988]. Estimates of smolts from 1985 to 1995 based on an index of smolts compared to indexed redds from 1976 to 1984.). SOURCE: Williams et al. (Submission 9).

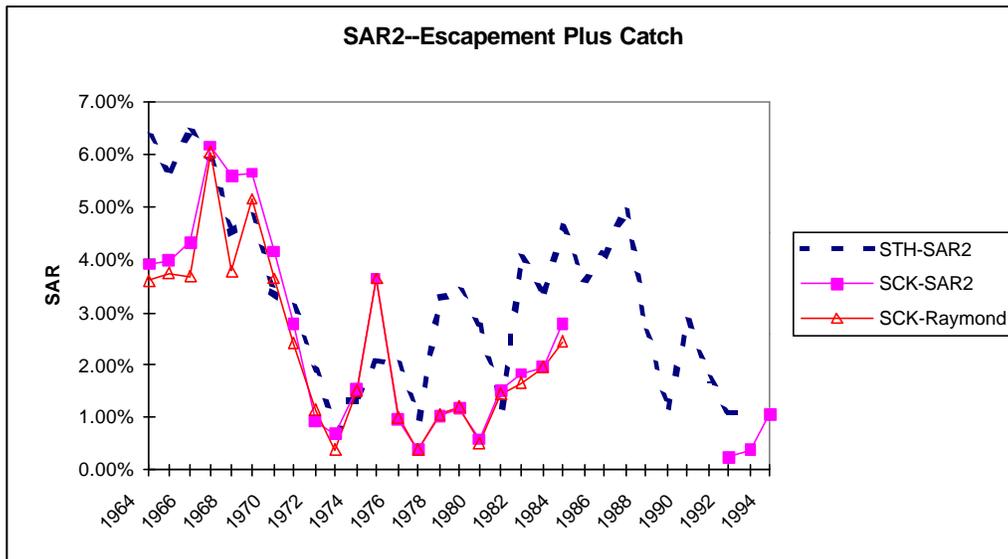


Figure 4-1: Wild Snake River spring and summer chinook salmon smolt to adult return (catch + escapement) rates to the upper dam on the Snake River (SCK-SAR2, two methods); and steelhead (STH). SOURCE: Petrosky and Schaller, Submission 10. See table "Fig_data".

SAR data is from Raymond (1988) for brood years 1962-1982, excluding 1972 (Toole 1997). Predictions of SARs (both historical and future) in the life-cycle model are based on the equation:

$$\text{prediction ln(adj. SAR)} = [\text{year-effect}] - [\text{small m}] + [\text{proportionality constant}]$$

where in the delta model,

$$\begin{aligned} [\text{year-effect}] &= \text{delta parameter} \\ [\text{small m}] &= \text{total passage + extra mortality} \\ [\text{proportionality constant}] &= \text{MLE estimate based on fit to SAR data} \end{aligned}$$

the alpha model has analogous parameters.

Both Raymond's SAR data and the predicted adjusted SAR do not include up-river conversion, first-pool mortality, and in-river harvest. To get a predicted SAR from the predicted adjusted SAR, therefore, one applies the appropriate up-river conversion, first-pool mortality, and in-river harvest to the predicted adjusted SAR's. Hence model-predicted SARs include errors in both the above predictive relationship, and errors in the predictors (i.e., year effect, m).

- (i) **Applicability: Score = 2**
 SAR data provide useful information about the relative merit of the alpha and delta models because they were not used in the fitting of the model or in the estimation of their parameters. Although the SAR data is not stock-specific, it does encompass periods before and after the Snake River dams were constructed.

Like the spawner-recruit data, however, SARs cannot be used to test aggregate hypotheses within the delta model because delta model predicted SARs do not vary with different aggregate hypotheses. Also like the spawner-recruit data, the SAR data has been widely cited and analyzed, and thus the potential exists for the data to indirectly influence formulation of hypotheses.

- (ii) **Clarity: Score = 2 to 4**
Justification for score of 4:
 Later authors have been unable to reproduce Raymond's SAR estimates. In addition, some PATH participants have raised concerns about the assumptions, methods, and documentation of Raymond's SAR estimates.

- (iii) **Rigor: Score = 2 to 3**
Justification for score of 2:
 Raymond's SAR estimates have been published in a peer-review journal, tend to corroborate the patterns seen in the spawner-recruit data, and are an independent measurement.

Justification for score of 3:
 Although Raymond's SAR estimates have been published in a peer-reviewed journal, later authors have raised doubts about some of the methods (Skalski 1998). In addition, some PATH participants have raised concerns about the assumptions, methods, and documentation of Raymond's SAR estimates. The relationship used to predict SARs has several weaknesses (Toole 1997): it does not account for intraseries correlation or ageing errors in the original SAR data; it does not include the 1972 data point (because it was clearly an outlier); and it has a low R^2 . Nevertheless, estimates of the SARs required for recovery using BSM and this relationship (2 to

7%) are similar to those developed by comparisons with historical data (Preliminary Decision Analysis Report, page 86).

Results

As a measure of the comparison of the predicted SAR to the historical data, we calculate:

$$\text{Variance of predicted SAR} = \text{sum} (\text{observed SAR} - \text{predicted SAR})^2 / (\# \text{ observations})$$

Variances of predicted SARs are considerably smaller for the Delta life cycle model than the Alpha life cycle model (Table 4-2). Within the Alpha life cycle model, variances are lower with FLUSH than CRiSP, for all TURB assumptions. TURB5 has similar variances to TURB4. These results are generally consistent with the comparison of the aggregate hypotheses to the spawner-recruit data.

Table 4-2: Comparison of aggregate hypotheses to observed SAR data. This comparison uses the same number of parameters as in Table 4-1.

Aggregate Hypothesis	Variance of Predicted SAR
Delta model hypotheses	0.081
Alpha, CRiSP TURB1	0.345
Alpha, CRiSP TURB4	0.381
Alpha, CRiSP TURB5	0.359
Alpha, FLUSH TURB1	0.207
Alpha, FLUSH TURB4	0.188
Alpha, FLUSH TURB5	0.190

Criterion 4) Validity of method of projection

The implementation of aggregate hypotheses involve assumptions about the future conditions relative to past conditions. In most cases, these assumptions cannot necessarily be tested with empirical data. We can test the aggregate effects of both the hypotheses and their implementation, however, by looking at the trends in stocks that are projected by each of the aggregate hypotheses under the A1 (current operations) option. Since A1 represents “status quo” conditions and operations, we can use projections under this action as an indicator of the basic behaviour of the aggregate hypotheses. In addition, by looking at these projections it should be possible to at least make a qualitative assessment of how reasonable projected trends are with no change in management in light of past and current levels of abundance.

In this section, we explore the range of projected trends in stocks that result from a “best-case” and a “worst-case” set of aggregate hypotheses under action A1. We chose best-case and worst-case aggregate hypotheses from the categorical regression tree for the 24-year survival standard (Fig. 3-2). Tracing the chain of hypotheses on the left-most branch of the tree in Figure 3-2 indicates that the lowest 24-year survival probability (0.46) is generated by the BKD extra mortality / Markov climate hypothesis, the FLUSH/T1 passage/transportation models, and the Alpha life cycle model. The highest 24-year survival probability (0.85) is produced by a combination of the Hydro extra mortality / Markov climate hypothesis, the CRiSP/T4 passage/transportation model, and the Delta life-cycle model. Both aggregate hypotheses assume the least favorable passage conditions (i.e., low FGE and Predator Removal effectiveness) and TURB1 historical turbine/bypass survival assumptions. (Note that best-worst is not simply a function of passage/transportation model: FLUSH/T1 aggregates also can generate relatively high 24-year survival probabilities (e.g., 0.82 with the Delta model, A3 and high predator removal

effectiveness), and CRiSP/T4 aggregates can generate intermediate 24-year survival probabilities (e.g., 0.65 with the BKD hypothesis, alpha model, and A1/A2) (Fig. 3-2.)

Projections of spawners using these aggregate hypotheses are shown in Figure 4-2. This figure shows historical estimates and future projections of spawners for Johnson Creek, one of the 7 Snake River spring/summer index stocks. Johnson Creek was used because it was the 6th best most often in terms of its 100-year survival probability. Future projections are **median** values in every 5th year of the 100-year simulation period, starting in 1996 (the use of medians explains the lack of year-to-year variability in these graphs). Survival and recovery thresholds are shown for comparison.

The best-case aggregate projects equilibrium spawners of around 450 spawners, which is considerably higher than the 1990-1995 average of around 150. The first projected point in 1996 for this aggregate is around 400 spawners, compared to recent empirical spawner estimates of around 20 (1995), 58 (1996), and 236 (1997). The worst-case aggregate hypothesis projects an equilibrium level of spawners at around the 1990-1995 average. With no assumed benefit of extended-length screens to FGEs and no assumed effect of the predator removal program, future passage conditions would seem to be no better than conditions in recent years. However, spawner abundances show an immediate doubling of spawners over the average number in recent years.

The relatively high projected escapements for 1996 (compared to actual data) demand some explanation. Predicted median spawner levels for 1996 are based on predictions of surviving recruits since 1991 **under the prospective hypotheses**. Some components of A1 have only been implemented since the 1995 Biological Opinion, but in these simulations these improvements are assumed to exist since 1991. In particular, as shown in Figure D-1, system survival is expected to improve dramatically under management action A1. In addition, prospective simulations sample from environmental conditions randomly from the long-term historical record, whereas recent years have generally had lower flow levels. Finally, the spawner-recruit functions maintain moderate recruitment levels even at low spawning levels. High spawner predictions are therefore generated by the combination of improved system and in-river survival, particularly under the hydro extra mortality hypothesis, average environmental conditions, and insensitive spawner-recruit functions.

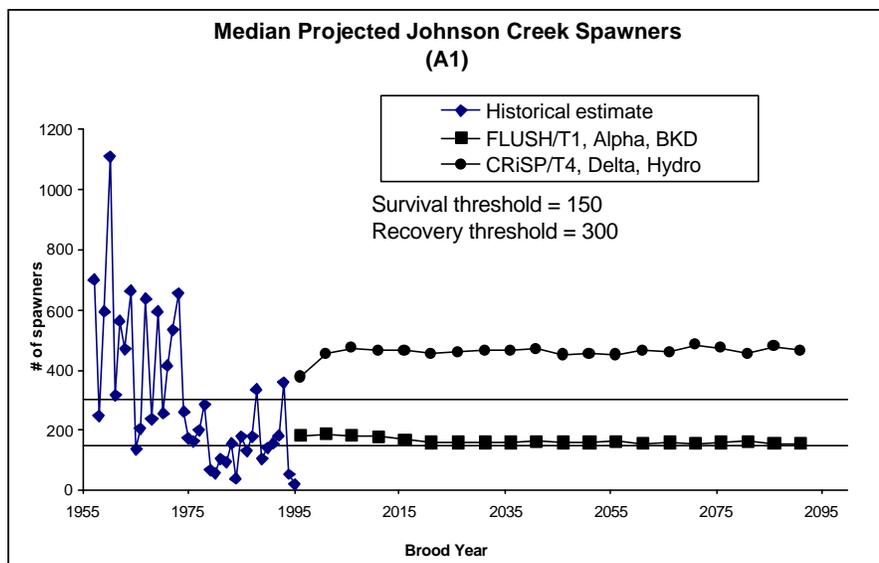


Figure 4-2: Projected spawners for Johnson Creek with best (CRiSP/T4, Delta, Hydro) and worst-case (FLUSH/T1, Alpha, BKD) aggregate hypotheses.

We can also examine other aggregate hypotheses besides the best/worst combinations. Table 4-3 shows the mean projected escapements (i.e., the mean of the median projected escapements for 1996-2090) for various aggregate hypotheses. These are provided to illustrate the range of projected escapements under different prospective aggregate hypotheses; to provide a basis for comparison, the median escapement for the entire historical period (1957-95) was 202, and for the recent period (1974-95) was 152.

Table 4-3: Mean of median projected escapements for 1996-2090 with management action A1, and under various aggregate hypotheses. These runs all assume lower FGEs (FGE2) and predator removal efficiencies (PREM1), and historically high turbine/bypass survival (TURB1 – which implies low reservoir survival in future projections under FLUSH).

Extra Mortality Hypothesis	Alpha Model		Delta Model	
	CRiSP/T4	FLUSH/T1	CRiSP/T4	FLUSH/T1
Hydrosystem	378	218	458	300
Regime Shift	458	176	483	212
BKD	279	162	312	214

4.2 Evaluation of Individual Hypotheses

4.2.1 Evaluation of Passage/Transportation Models

Section 3 showed that passage/transportation models are a critical factor in determining the modeled outcomes of actions. In general, FLUSH/T1 tended to strongly favor A3 over A1 and A2. CRiSP/T4 also tended to favor A3 over A1 and A2, but not quite as strongly. Unlike FLUSH/T1, there were a small number of situations where CRiSP/T4, in combination with the regime shift extra mortality hypothesis, the alpha model, and pessimistic drawdown assumptions, resulted in A1 or A2 having slightly higher jeopardy probabilities than A3.

Diagnostic analyses in Section 3 of this report suggest that the transportation component of the passage/transportation model is the most important cause of differences between outcomes of management actions in FLUSH/T1 and CRiSP/T4 passage/transportation models. That is, differences in the historical (retrospective) D values estimated by each modeling system explain most of the differences in their results. Additional diagnostic analyses showed that differences in the retrospective D values are largely due to differences in the in-river survival rates that are combined with Transport:Control ratios to estimate D (FLUSH “s”, CRiSP Vcl; Figure 3-23). FLUSH “s” values are generally higher in pre-1980 years than CRiSP Vcl values, while in post-1980 years CRiSP Vcl values are higher than FLUSH s values.

Differences in survivals of control fish are a function of a) differences in the structures of the passage models themselves, and b) different assumptions about historical turbine/bypass survival (TURB assumptions). Therefore, we deal with each of these separately. The Preliminary Decision Analysis Report (pgs. A-6 to A-9) reviewed the structural differences in the two passage models, and so we do not repeat this material here. The passage models themselves are evaluated in Section 4.2.1.1, while TURB assumptions are reviewed in Section 4.2.1.2. An evaluation of the assumptions underlying the different transportation models (i.e., TRANS1 associated with FLUSH, TRANS4 associated with CRiSP) is provided in Section 4.2.1.3.

4.2.1.1 Passage models

Criterion 1) Clarity

Both passage models are complexes of other hypotheses about flow-survival relationships, direct hydrosystem effects on biological and physical processes, allocation of mortality in the hydrosystem to different sources, etc. In addition, both models represent different approaches to modeling (i.e., detailed representation of biological mechanisms in CRiSP vs. general relationships based on empirical data in FLUSH) (see Section A.2.1.1 and Table A.2.1-1 in the Preliminary Decision Analysis Report). Based on diagnostics in the Preliminary Decision Analysis Report (Figure A.2.1-8), differences between the passage models are primarily in their reservoir survival estimates. FLUSH models reservoir survival based on empirical flow-survival relationships, while CRiSP models reservoir survival based on explicit gas and predation mortality functions. Issues related to these functional forms are presented in Tables 4-3 and 4-4.

In general, the original review of the passage models found strengths and weaknesses in both of the two passage models, and have not unequivocally endorsed one model over another (Barnhouse 1993). Some reviewers have commented that simpler models are generally better (e.g., original reviews of models, Review 1 p.18; Review 3 p. 3, while others have suggested that a more detailed model is appropriate (e.g., original reviews of model, Review 2 p.2, 30; Review 4 p. 10).

The interpretation of the causes of very low passage survivals in 1973 and 1977 distinguish the travel time-survival relationships in the models. The FLUSH model associates low survival rates in those years with long travel times, while the CRiSP model associates the low survival rates with poor dam passage conditions in the 1970's. The implied relationship between flow and reservoir survival in CRiSP is much weaker (flow has less effect on survival) than FLUSH. Note that there are many possible shapes of these relationships (flow-survival, gas mortality, and predation).

Past reviews of flow-survival relationships have been extensive (e.g., Cada 1994). The most recent of these, the Independent Scientific Groups's "Return to the River" report concluded that the assumption that "survival is positively related to the water velocity prevailing during the outmigration" has theoretical support with some evidence from experiments or observations (p. 55), but also concluded that "because it has not been possible to separate the influence of flow from that of other variables on survival, the relation between flow and survival remains obscured" (p. 55). They also point out that "the relationship between exposure time of emigrating smolts to mortality factors in the hydroelectric system and the overall survival of smolts is intuitively reasonable, but has not been demonstrated conclusively" (p. 54). Finally, they make the point that "it seems unlikely that an incremental quantitative relationship between these variables would apply equally to all species and life history types or necessarily be constant over time and space" (p. 55).

Table 4-3: Specific Issues and pieces of evidence relating to **Clarity** of hypotheses (counterpoints in *italics*).

Issues Relating to Clarity	Source (Submission – page)
FLUSH	
The FLUSH model is based on the assumption that the longer the fish spend in the hydro system, the higher their mortality rate will be.	22-15
CRISP	
Functionally, the average rate of mortality is constant with time giving an exponential-like survival function.	14-3
The consumption of stream type chinook by several predators is well documented.	22-16
Mortality due to GBT is not documented but straight forward from the theoretical point of view.	22-16

*Criterion 2) Mechanism***Table 4-4:** Specific issues and pieces of evidence relating to **Mechanisms** of hypotheses (counterpoints in *italics*; rebuttals to counterpoints in curly brackets { }).

Mechanism	Source (Submission – page)
FLUSH	
Researchers have reported the substantial increases in travel times associated with impounded water (Raymond 1968), and the poor survivals associated with poor migration conditions such as low flows, low spill rates and lack of night time flows (Raymond et al. 1973; Ebel et al. 1973; Park et al. 1978). Increased travel times have the following effects: <ul style="list-style-type: none"> • Longer travel times alter seawater entry timing which is poorly synchronized with the physiological state of the smolts (CBFWA 1991). • Places greater demands on energy reserves of actively migrating yearling chinook. • Greater exposure to predation. Predation by northern squawfish is well documented (Poe and Rieman 1988). 	22-15
<i>Adjustment of yearling chinook to seawater is not related to gradual exposure to salt water.</i>	<i>Oral comments by John W., Hoar 1976.</i>
<i>The FLUSH flow-survival relationship is largely based on the 1970-1980 Sims and Ossiander reach survival data, after removing dam passage mortality. Other model forms using different assumptions on the level of dam passage mortality have not been explored. Reviews of the Simms and Ossiander data by Steward 1994, Mathews and Williams 1985, and Skalski 1998 raise concerns about the scientific rigor of these estimates and their applicability to the existing and future hydro system conditions.</i>	14-3

Mechanism	Source (Submission – page)
<i>The positive relationship between mortality rate and travel time is dependent on the 1973 and 1977 data points. Using these two years is problematic because they are the only two significant outliers in the survival estimates dating back to 1966. The relationship based on these long travel time data is used to project outcomes for A3, where travel times are short.</i>	14-6; oral comments by Jim A.
<p>{In a statistical sense, outliers are data points that have independent variable values which overlap with those of other data points, but have anomalous dependent variable values. The fish travel times for 1973 and 1977 study reaches are the two longest, by a substantial amount (PATH Decision Analysis Report, Figure A.2.1-18a). There are no other data points at those fish travel times to contradict the dependent variable values. They are therefore not outliers but “influential points”, and in fact provide much of the contrast in the available data. }</p> <p>{Even under TURB4, which assumes very high dam mortality in 1973 and 1977, there is still a shallow relationship between surv.rate/day and travel time }</p>	22-6 Prelim. Dec. Analysis Rep., Fig. A.2.1-9.
<i>The upside-down logistic curve used to model travel time-survival relationship in FLUSH is unique to ecology. No similar relationships or biological evidence has been cited. This relationship implies that fish traveling together will have different rates of mortality in FLUSH depending on release date and location.</i>	14-4
<p>{The FLUSH model assumes that dam and reservoir passage has both instantaneous and cumulative effects on survival, resulting in a “Type 1” survival vs. time relationship. The complement of the survival vs. FTT relationship is a sigmoidal cumulative mortality vs. time curve, which is exactly the form one would expect from sub-lethal impacts on the fish from each project that was passed. Lethal exposure times (i.e., LE₁₀₀) to different stressors are ubiquitous in the biological literature (e.g., Pearl, 1928). }</p>	22-5
<i>The Sims and Ossiander data show that survival from the first project to Ice Harbor dam was lower than survival from Ice Harbor dam to the last project. The FLUSH hypothesis is inconsistent with this pattern.</i>	14-7, 14-8
<p>{FLUSH is not inconsistent with this pattern – mortality is increasing per time, but mortality per project is not increasing because fish are moving faster relative to the water in the lower reaches, and water itself moves faster in Columbia R. than Snake R. }</p> <p>{1997 PIT-tag data and passage model estimates indicate that survival rates are lower in the lower reaches. }</p> <p><i>Uncertainty in the 1997 data is too large to conclude any difference in upper and lower reach survival.</i></p>	PATH FY96 Retrospective Report, Ch. 6 App. 5 Prelim. Dec. Analysis Report Fig. A.2.1-14; oral comments by Paul W. <i>Oral comment by Steve S.</i>
<i>FLUSH assumes that time in the river does not affect survival above Lower Granite pool, which is inconsistent with the assumption that survival below Lower Granite pool is dependent on travel time.</i>	14-4

Mechanism	Source (Submission – page)
CRiSP	
In CRiSP, the average rate of mortality is constant giving an exponential-like survival function. Variations in the rate of mortality depend on water temperature and gas levels but these factors do not alter the basic relationship.	14-4
<i>CRiSP has many more parameters than FLUSH to determine the relationship of reservoir survival to FTT – it is impossible to tell how much impact total dissolved gas has on this relationship.</i>	22, pp. 2 to 4
{The CRiSP group has done this by setting GBT to zero and observing effects – gas portion has minimal effect on survival (results not yet distributed).}	Oral comments by Jim A.
<p><i>Information on signs of Gas Bubble Trauma (GBT) in juvenile salmonids at less than 130% saturation is limited.</i></p> <p><i>It is unclear how big of a problem GBT is – fish may be able to avoid GBT by going deep. ISAB is currently looking at this.</i></p>	<p><i>Mesa et al. 1997 and 1998; Dawley et al. 1974 and 1975;</i></p> <p><i>Oral comments by Chris P.</i></p>

Criterion 3) Consistency with empirical evidence

FLUSH draws on three data sources: reach survival estimates, dam mortality estimates, and estimates of fish travel time. FLUSH is calibrated to a series of annual reach survival estimates. The dam survival for each reach, each year, is backed out and the remainder is reservoir survival which is allocated among the reservoirs in the reach based on the fish travel time.

CRiSP also uses dam mortality estimates for the dam component of project mortality, but uses data on consumption of juvenile chinook to estimate reservoir survival.

a) Dam survival estimates

- (i) **Applicability:** Score = 1 to 2 (placeholder)
Dam mortality estimates are from research conducted on spring/summer chinook in the Columbia River hydropower system.
- (ii) **Clarity :** Score = 1 to 4 (placeholder)
There is agreement within PATH on the data sets concerning dam survival, reach survival and fish travel time in recent years, but considerable disagreement over some of the studies in the 1970's.
- (iii) **Rigor:** Score = 2 to 4 (placeholder)
Dam survival estimates (Turbine survival, spill survival, bypass survival, FGEs) have been reviewed within PATH but these estimates are less certain.

b) Consumption data

- (i) **Applicability:** Score = 3 to 4 (placeholder)
Consumption rates for spring/summer chinook were estimated in John Day Reservoir from 1983 through 1986. Data for earlier years are not available. Predation indices for this species were developed for other reservoirs in more recent years.
- (ii) **Clarity:** Score = 3 to 4 (placeholder)
Consumption based models are established in fisheries literature, although they are not generally used to predict predation. However, you have to assume predator-prey distributions to predict predation (Petersen et al. 1993), but we don't know anything about them.

There are no estimates of prey densities for many years which adds considerable uncertainty to estimated functional response. There are large confidence intervals around predator estimates and indices. CRiSP model is sensitive to functional response – it uses a Type I response, but Vigg (1988) suggested a Type II response.
- (iii) **Rigor:** Score = 2 to 4 (placeholder)
Consumption rates were based on accepted fisheries methods and studies were peer reviewed. Predation indices are conducted regularly but are considered “spot checks” by researchers.

c) Reach survival estimates and fish travel times

Consistency with survival estimates over entire reach

The PATH hydro group agreed on a set of reach survival estimates to use for model calibration and comparison (Preliminary Decision Analysis Report, Table A.2.1-2). There are 19 of these estimates, from 1966 to 1996. Reaches covered by the estimates varied from year to year.

- (i) **Applicability:** Score = 2-3
Reach survival estimates and passage model estimates of in-river survival of control fish in transportation studies are not directly comparable because the survival studies do not encompass all eight of the projects that Snake River fish currently pass through, and because the dam operations that were in place during the older studies are different from current operations (Toole et al. 1996).
- (ii) **Clarity:** Score = 1
Reach survival data are relatively clear and unconfounded. The set of reach survivals used for calibration by FLUSH (and possibly by CRiSP – see Submission 22 p.1-2) and comparison of models has been agreed upon by the PATH hydro work group.
- (iii) **Rigor:** Score = 1-2
The NMFS reach survival estimates are widely cited and used throughout the basin. Details on the methodology for recent PIT-tag studies (since 1993) have been peer-reviewed and published in Skalski (1998) and in Skalski et al. (in press, CJFAS). Older reach survival studies have also been published in peer-reviewed journals (Raymond 1977, 1978), although some aspects of their mark/recapture and statistical methods have since been questioned (Matthews and Williams 1985; Toole et al. 1996; Skalski 1998, Steward 1994).

Chris Toole’s analysis of this in the Preliminary Decision Analysis report showed that both models give roughly equal fits to 1966-1980 and 1993-1996 survival data over the entire reach (Table 4-5). Differences between model and empirical estimates of reach survivals are generally much smaller than the differences between the models’ estimates of in-river survival of control fish, particularly in 1971, 1972, 1975 and 1976 (Figure 4-3).

Table 4-5: Summary of linear regression fits (r^2) to empirical reach survival estimates for CRiSP and FLUSH model estimates. FLUSH retrospective mode implemented observed fish travel times (FTT), so comparisons are not applicable. CRiSP comparisons are for 1966-1980 and 1993-1996 (n=19); FLUSH comparisons are for 1970-1980 and 1993-1996 (n=15). From Table A.2.1-3 in Preliminary Decision Analysis Report.

	FLUSH	CRiSP
Retrospective Model		
FTT	N/A	N/A
TURB1	0.85	0.83
TURB4	0.77	0.87
TURB5	0.86	0.84

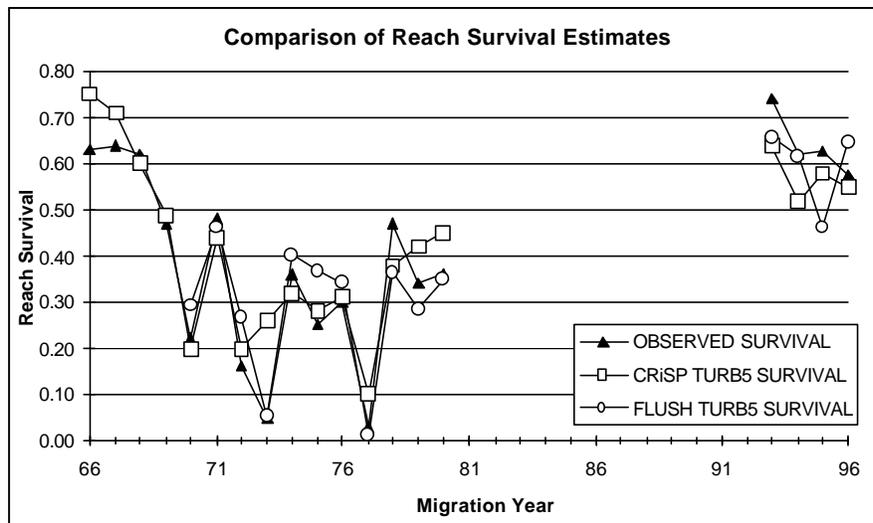


Figure 4-3: Comparison of observed and predicted reach survival estimates.

Consistency with 1989 to 1992 proportion of PIT-tagged fish detected at Lower Granite and McNary

NMFS recently completed an analysis of 1989-1992 detection probabilities from above Lower Granite Dam to McNary Dam and compared these data to estimates provided by the passage models (Submission 16, starting on p. 29).

(i) Applicability: Score = 2 to 4 (placeholder)

Justification for score of 2 (greater applicability):

This data sets was not used in the calibration of the models so it provides an out-of-sample test. The data are useful to investigate the degree to which passage model predictions agree with the proportions of fish detected at McNary dam (submission 16, p. 30). NMFS made no estimates of survival and did not compare model outputs of survival estimates to "observed" estimates (Submission 16)

Justification for score of 4 (less applicability):

Estimates are approximations, not statistical observations, and were never intended for use as survival estimates. Detection probabilities were higher in 1992 than 1991, but survival was lower (Submission 22 p. 9-13). Data were from a mixture of wild and hatchery fish.

(ii) Clarity: Score = 2-4 (placeholder)

Justification for score of 2:

The data are quite clear: release fish above, count how many show up below, compare to what the models predict. Several model parameters other than survival might contribute to the discrepancy, particularly FGE at Lower Granite, Little Goose, and McNary Dams. In fact, data on detections at Lower Granite and Little Goose Dams suggest that FGE values assumed in PATH are in error. However, no reasonable combination of assumptions for FGE at the dams can explain the entire difference between observed and predicted percentages detected at McNary Dam (Submission 16, p.15).

Justification for score of 4:

The NMFS estimates are very sensitive to assumptions about FGE, spill, flow, and method of calculations. These assumptions are not well documented. PIT-tag releases were not made throughout the smolt migration season, and trap catches were flow-dependent (Submission 22 p. 9-13).

(iii) Rigor: Score = 2-3

Justification for score of 3:

There appears to be within-year temporal patterns in detection probabilities, but fish in the 1992 study were only released early in the season (prior to May 20). Also, few tags were released or recovered (Submission 22 p. 12,13)

Results of Comparison to 1989 to 1992 PIT-tag data

Details results are provided in Submission 16 (starting on p. 29); we summarize their main conclusions here.

- In every year from 1989 to 1992, the CRiSP model predictions were closer to observed data than the FLUSH model or a spreadsheet model developed using survival estimates from the 1970's (see Submission 16, p.39 Figure 2).
- CRiSP predictions were considerably lower than observed data in 3 out of 4 years; the lower flow years.
- The best predictions were made by a spreadsheet model using survival estimates from the 1990's.
- In the 3 lowest flow years, FLUSH predictions were closer to the spreadsheet model using the 1970's survival estimates than they were to the observed data.

- Models used to project survival in the future must reflect that relatively high survival is possible (and has occurred in recent years) under low flow conditions in the Snake River.

Comments on results {rebuttals are from Submission 16 p. 15-17}

FLUSH detection and survival estimates matched more closely in 1989, which was year with the largest number of releases.

{This was also the year with highest flow }

NMFS survival estimates exceeded 1.0 in 2 of 4 years (Submission 22, Table 1).

{The purpose of the exercise was not to develop reach-by-reach survival estimates for 1989-1992. Survivals can be reduced to below 1 by adjusting FGEs.}

Although FGE assumptions do not affect detection probabilities, they do affect the passage model predictions (Submission 22, p.9,10).

{Increasing FGE at LGR would reduce the passage models' predicted detection, increasing the difference between predicted and observed}

Consistency with 1997 and 1998 PIT-tag survival estimates

NMFS also estimates survival rates from the tailrace of Lower Granite Dam to the tailrace of Bonneville Dam for the years 1997 and 1998, and compared these estimates to CRiSP and FLUSH passage model output (Submission 16, pp. 41-45; FLUSH outputs were not available for 1998).

- (i) Applicability: Score = 1 to 4 (placeholder)

Justification for score of 2 (greater applicability):

The data are for wild yearling chinook salmon from the Snake River drainage, migrating through nearly the entire hydrosystem. These estimates were not used for calibration and were not provided to the modeling groups (Submission 16, p. 41). Applicability of 2 might be considered only because the data are not stock specific.

Justification for score of 4 (less applicability):

Using coho detection probabilities as a surrogate for chinook detections ignores differences between the two species. Errors around survival estimates for those years precludes strong inference (Submission 22, p.13).

- (ii, iii) Clarity and Rigor: Score = 2 to 3 (placeholder)

Justification for score of 2:

Although the data are weaker than the data agreed on by the PATH hydro group, they still provide some relevant information.

Justification for score of 3:

Estimates are inconsistent with methods established by the PATH hydro group. The group agreed that the relevant reach for the 1996 survival study was from Lower Granite to McNary tailrace, because the number of detections at the lower projects were too few to allow estimation of survival rates over that reach. Similar concerns about the lack of detections at lower projects were expressed by NMFS about estimation of survival to John Day dam tailrace in 1997. Methodology for estimating survival to Bonneville is unclear. (Submission 22, p.13).

Results of Comparison to 1997 to 1998 PIT-tag data

Detailed results are provided in submission 16, p. 42

- Based on empirical data, estimated survival from Lower Granite Dam tailrace to Bonneville Dam tailrace for wild yearling chinook salmon was 45.5% in 1997 and 57.8% in 1998.
- CRiSP predictions matched the estimates better in the higher flow and spill years of 1997 (CRiSP: 43.7%) and 1998 (51.5%) than in the lower flow years 1989-92.
- the FLUSH prediction for 1997 (adjusted to remove Lower Granite project survival) was 43.6% (averaged over 4 TURB assumptions).

Consistency with Per-Project Survival Estimates

NMFS compared the survival rates of control fish estimated by FLUSH and CRiSP to historical survival rate estimates from Raymond (1979). Since the reaches covered in the Raymond estimates differed from the reaches covered by the transport studies, comparisons were made on a per-project basis. Details on methods and results are provided in Submission 16 (starting on p. 48). Raymond reach survival data is evaluated above.

In general, the analyses found that FLUSH average per-project transport control survival estimates are at least 10% higher than Raymond run-at-large estimates in most years for which comparisons are available (exceptions 1978-80). CRiSP per-project estimates are within about 2% of Raymond's estimates in all years except 1976 and 1979, when they are lower than Raymond's estimates. Model comparisons assumed TURB4 assumptions. The discrepancy is particularly large in 1971, when FLUSH per-project survival of control fish was 0.94 per project and the Raymond estimate was 0.65.

Response to these observations (submission 22, p. 13-14):

- *The 1971 estimate of control fish survival was overestimated because it does not account for gas-related mortality, which was probably high in that year because of high flows and spills, and no mitigation structures in place at the dams.*
- *The 1971 survival estimate from FLUSH does not influence the T:C vs. Vc relationship or D values used in retrospective or prospective modeling.*
- *Figure D-16 in the Appendix documenting reason for differences in passage models suggests that even if the 1971 estimate were adjusted downward to account for gas problems, it would result in lower retrospective and prospective D values.*
- *The FLUSH group has questioned the credibility of the TURB4 hypothesis. Assuming TURB4 results in a poorer fit to the empirical reach survival data (Table 4-5) and to the spawner-recruit data under the Alpha model (Table 4-1). Comparing FLUSH TURB4 survival estimates to the Raymond estimates is not a proper test of the STFA hypothesis.*

Comparison to 1994-1996 PIT-tag data from Lower Granite tailrace to Lower Monumental tailrace

The relationship between survival and travel time from the 1994-1996 PIT-tag data from Lower Granite tailrace to Lower Monumental tailrace is weak. The weak relationship between survival and travel time predicted by CRiSP is closer to the estimated relationship than the strong relationship predicted by FLUSH (Submission 14 p.5,6).

The year with the lowest flows, 1994, was also the year with the lowest survival (oral comment by Paul W.).

d) Spawner-recruit data

Retrospective recruit/spawner survival estimates can be generated from passage model input with the life-cycle models and compared to historical spawner-recruit data. S-R data is evaluated in Section 4.1. These comparisons are limited by the constraints discussed in Section 4.1 above (e.g., they are only possible with the alpha model because all delta model aggregate hypotheses fit the S-R data equally well), but they do allow some qualitative observations. In general, FLUSH estimates have lower AIC and BIC scores than CRiSP outputs (Table 4-1).

e) Smolt to Adult Return data

Similar to the spawner-recruit data, retrospective Smolt to Adult Return (SAR) rate estimates can be generated from passage model input with the life-cycle models and compared to historical SAR data. SAR data is evaluated in Section 4.1. In general, SARs predicted with the FLUSH model fit the SAR data better than CRiSP outputs (Table 4-2).

f) Other Potential Data Sets

Paulsen et al. (Submission 5) suggest a number of possible tests of the passage models. These include:

- Recovery proportions of PIT-tagged fish, 1989-97, from various Snake and Columbia projects as far down the system as data permits;
- TCRs for transport study groups, passage year 1995 (should have complete recoveries of 3-ocean fish in 1998);
- Fish travel times for PIT-tagged fish, starting in 1987 (from Snake River traps);
- Reach survivals of PIT-tagged fish for 1996-98, as low in the system as data permits;
- Potentially, recoveries of fish below Bonneville in experimental trawls and by the Rice Island tern colony. Given the small number of recoveries to date, these would be pilot studies.

These tests would potentially provide useful information because they are generally not used in the calibration of the models. Some of them have already been completed (e.g., NMFS analyses of recent PIT-tag studies). However, because of time constraints it will not be possible to do the remaining analyses until after October.

4.2.1.2 TURB assumptions

Assumptions about historical levels of turbine/bypass survival (TURB assumptions) affect D estimates through their influence on estimated survival rates of control fish on which D's are based. The main issue is to what extent descaling of fish as they pass through a turbine or bypass system causes mortality, and to

what extent such mortality continues to occur. TURB4 assumes that the rate of mortality through the turbines and bypass is related to the amount of descaling, and that the cause of decaling in the past (excess debris in the forebays) has been corrected. TURB5 assumes that bypass mortality is equal to the descaling rate, and turbine mortality is equal to one-half of the descaling rate. As more mortality is ascribed to the dams, less is allocated to the reservoirs. Thus reservoir survival is higher under TURB4 in prospective simulations. Both TURB assumptions are described in more detail in Section A.2.3 of the PATH Preliminary Decision Analysis Report.

A sensitivity analysis of the effects of using TURB4 D values vs. TURB5 values suggests that the model results are not particularly sensitive to these assumptions (Appendix F). Nevertheless, some evidence pertaining to the effects of descaling on turbine/bypass survival is summarized here; more details are included in Submission 22, pp. 7-9, 20-21.

Criterion 1) Clarity

There is generally strong agreement that descaling could result in mortality but there is uncertainty as to the source of descaling. While debris in trash racks could be one cause of descaling (Williams and Matthews 1995) it appears that it is not the only one and may not be a prominent one (Submission 22).

Criterion 2) Mechanism

Descaling is a plausible source of mortality and may be a good indicator of fish condition. Debris, however, does not appear to be the only factor affecting descaling. Wilson et al. suggest that TURB4 overestimates turbine/bypass mortality because it ascribes bypass mortality to bypassed fish that were transported at Little Goose after 1977 – turbine/bypass mortality is on average 270% of the descaled fish prior to 1981 (Submission 22-8). *Anderson (oral comment) counters that descaling is used only as a correlative indicator of fish condition.*

Criterion 3) Consistency with empirical evidence

As noted above, Wilson et al. (Submission 22) argue against the theory that debris in trash racks was the primary source of descaling and mortality in the 1970's. They suggest this mortality was confounded by other factors; that other sources of mortality were noted in the literature at the time the research was conducted; that this literature discounted the debris problem; and that high levels of descaling continue to occur (pg. 22-8; and Preliminary Decision Analysis Report, Table A.2.3-3). These recent descaling levels contradict the TURB4 assumption that the condition of fish improves in 1980 and no additional mortality (beyond standard rates) occurs.

Oral comment by Jim A.: TURB4 may overestimate project mortality, but TURB5 may underestimate project mortality and TURB1 clearly underestimates historical project mortality because it assumes no change in hydrosystem operations. Therefore, the TURB hypotheses at least bracket the range of historical turbine/bypass mortalities. Also, descaling levels have generally declined substantially from the mid-1970's (pg. 22-8), especially at Lower Granite.

4.2.1.3 Transportation Models

The efficacy of transportation is expressed as D, the ratio of post-Bonneville survival of transported fish : post-Bonneville survival of non-transported fish. In general, the TRANS4 transportation hypothesis assumes that D values in the “current” (1980-present) period have increased substantially since the pre-1980 period. Post-Bonneville survival of transported fish is related to the degree of descaling and stress experienced during collection for transport (Table 4-6, 4-7). Since the primary cause of descaling (excess

trash) is assumed to have been corrected, post-Bonneville survival of transported fish (represented by D) is higher than in the pre-1980 period.

The TRANS1 model assumes that current D values are lower than TRANS4 values, and have not improved since 1980. Details on the calculation of D values is provided in Appendix D. In general, it was concluded that the primary difference in the D's estimated by the two transportation model (TRANS1, associated with FLUSH and TRANS4, associated with CRiSP) was due to differences in the survival of control fish estimated by the two passage models.

Criterion 1) Clarity

Table 4-6: Specific issues and pieces of evidence relating to **Clarity** of hypotheses. Counterpoints are in *italics*. Rebuttals to counterpoints are in curly brackets {}.

Clarity	Source (Submission- page)
FLUSH/TRANS1	
Closely aligned with research results, although TCRs may be biased high.	22-17
<i>Model used to generate FLUSH estimates of control fish survival (s) is undocumented.</i>	<i>Oral comment, Jim A.</i>
{FLUSH s values are based on simple expansions of control survival estimates from NMFS transport studies in 1970's, and from the FLUSH passage model in 1986 and 1989. The transport model is intended to be as independent of the passage model as possible.}	23-7; Oral comment, Howard S.
<i>It appears that the FLUSH model underestimates survivals in many years, particularly those with lower flows, under present hydropower conditions. This has the effect of decreasing D.</i>	16-9
<i>FLUSH/TRANS1 D values ignore improvements in fish handling (lower rate of descaling) in the recent (post-1980) period relative to the pre-1980 period.</i>	14-10
{SAR and stock-recruit data do not support hypothesis of improvement in migration conditions, including transportation survival.}	22-7; 23-10
CRiSP/TRANS4	
Prospective CRiSP D values are based on transportation data from the 1980's and 1990's which are representative of current transportation conditions. The consideration of a pre-1980 and post-1980 period incorporates improvements in transport conditions since 1980. Estimates of D based on descaling comport with estimates based on T/C's and control survivals.	14-10,11
<i>Survival of control fish always less than survival of in-river fish</i>	23-7

<p><i>There are several potential problems with the current implementation of the TRANS4 hypothesis, including:</i></p> <p><i>Inconsistent use of TURB4 and TURB5 assumptions</i></p> <p><i>The use of the median as an estimate of the central tendency of D values</i></p> <p><i>Arbitrary selection of 1980 as the break point for the “historical” and “current” periods</i></p> <p>{Sensitivity analyses in Appendix F suggest that these issues have minor effects on the results.}</p>	23
--	----

Criterion 2) Mechanism

Table 4-7: Specific issues and pieces of evidence relating to **Mechanisms** of hypotheses. Counterpoints are in *italics*. Rebuttals to counterpoints are in curly brackets {}.

Mechanism	Source (Submission – page)
FLUSH/TRANS1	
Assumes some advantage to transported fish, but assumes some delayed mortality of transported fish. Mechanisms for delayed mortality of transported fish include stress and injury during bypass and collection and disease transmission.	22-19
<i>According to results shown in Table D-3 (top graph), system survival decreases with decreasing survival of non-transported fish (directly related to travel time.) This model does not explain a mechanism of why transported fish mortality decreases as survival of non-transported fish decreases.</i>	16-9
<i>FLUSH/TRANS1 transportation assumptions relates post-Bonneville survival of transported fish only to the experience of the in-river fish. Effects of collection conditions on transported fish are ignored.</i>	14-11
{D is calculated from the T/C data and an estimate of control (in-river survival) in both models, and thus is dependent on the survival rate of non-transported fish. T/C must vary with in-river survival, since avoidance of the direct mortality of the hydrosystem below the collection point is presumed to be the primary benefit of transportation. Since in FLUSH model T/C increases with decreasing Vc, underestimates of Vc will be mitigated to some extent by the increase in T/C, since $D = T/C * Vc/Vt$. There is no such effect to mitigate underestimates (or overestimates) of Vc in CRiSP method, since T/C does not vary with Vc.}	22-6
CRiSP/TRANS4	
The CRiSP/TRANS4 transportation model assumes that D depends on the condition of the transported fish, which has improved since 1980.	14-11
<i>Although descaling likely had large impacts on fish during the many of the years in the 1970s (see Williams and Matthews 1995), it is not clear how descaling would have differentially impacted transported fish compared to downstream migrants.</i>	16-9

Mechanism	Source (Submission – page)
{Survival of control fish (Vc) is adjusted for descaling.}	Oral comment, Jim A.
<i>There is no apparent biological reason to make a break between pre-1980 and post-1980 for historical and current periods. 1969, 1970, 1975, 1982-85 were years with high SARs between 1969 and 1985 (a nearly, to fully completed hydropower system). In many of these years, descaling was high but little to no debris existed at the uppermost powerhouse on the Snake River.</i>	16-8,9; 9-Fig. 5
<i>Increases in the effectiveness of the transportation program manifested in the CRiSP D's are all related to assumption that pre-1980 fish condition was poor due to debris in the trash rack that were going uncleaned, and since then the descaling related problems are nonexistent. However, high descaling rates have continued into the present. No plausible reason has been given why descaling was so deadly prior to 1981 but innocuous afterward.</i>	22-19; 16-9
<i>Descaled fish were systematically removed from transport experiments.</i>	23-8,9

Criterion 3) Consistency with empirical evidence

Transportation models are based on data on relative returns of transported and non-transported fish (controls) in transportation experiments. They are also based on passage model estimates of the in-river survival of control fish (evidence pertaining to passage model estimates is reviewed in the previous section).

a) Transport:Control Ratios

(i) Applicability: Score = 2 to 3 (placeholder)

Justification for score of 3:

It is not justifiable to use T/C values from transport studies where fish were trucked because nearly all fish are now barged. In 1978 trucked fish from Lower Granite Dam returned at 0.013% rate; those barged from Lower Granite Dam returned at 0.116% rate (9 X greater than that of trucked fish). The ISAB Report 98-2, dated 27 February 1998, evaluated the efficacy of transportation and concluded “Trucks should not be used in the transportation program due to a lack of information needed to advise management, due to an absence of current research programs to collect such information, and because historical indications on truck transport are negative.” (Submission 16, p. 8).

Justification for score of 2:

All transport data pertains to transported spring/summer chinook and their controls. Relative returns of trucked and barged fish depend on whether fish trucked in saltwater are included. FLUSH/T1 and CRiSP/T4 T/C estimates exclude these groups because fish were never transported in saltwater. For 1978, the ratio of SAR of barged fish to that of fish trucked in freshwater only is 1.55, as opposed to 9 as estimated above (Submission 22, p.17). Also, the highest T/C estimate ever observed was in 1973, when all fish were trucked (oral comment by Paul W.).

(ii) Clarity: Score = 3 to 4 (placeholder)

There is considerable disagreement concerning the interpretation of the transportation data and which data series to use. Disagreements arise out of the methodologies used to estimate D_s because the choice of individual T/C data points, as well as factors unrelated to T/C ratios such as in-river survival (discussed in Section 4.2.1.1 above) and TURB assumptions (discussed in Section 4.2.1.2 above), can all significantly affect D estimates.

(iii) Rigor: Score = 2 to 3 (placeholder)

PATH has adopted a series of transport studies but has not thoroughly investigated the sensitivity of model results to different assumptions about controls. Transport studies have been extensively peer reviewed. Prominent among reviews is Mundy et al. (1994). These reviewers noted that transport experiments lacked true controls and that fact may lead to over estimates of TCR's. The possibility that TCRs are close to 1:1 or less is reinforced by comparisons of the SARs of transported fish (see following section) and those that were never detected (Kiefer, IDFG, pers. comm.). The undetected fish returned at higher rate than transported fish. Thus TCRs may be overestimated.

The claim that transport TCR are possibly close to 1:1 is a selective use of data. Table 1 (Submission 16, p. 11) provides complete TCRs from 1993 to 1995 and those derived from the 2-ocean returns from the 1996 outmigration. There is a wide range in TCRs of transported fish compared to those never detected in the system. Further, if a belief exists that old transport studies have little value because of handling of controls that were transported (as detailed by Mundy et al. 1994), then only the 1995 transport study has any validity in the PATH process. (Williams, oral comment).

b) Spawner-Recruit Data

Wilson et al. (Submission 22-19) suggest that the FLUSH/TRANS4 hypothesis of low D values is consistent with estimates of m (Deriso et al. 1996). Passage models which assumed low delayed mortality of transported smolts had the poorest fit in the MLE analysis (e.g., Fig. 5-5 of Deriso et al. 1996). Total mortality of Snake River spring/summer chinook (μ) tended to be highest in low flow, low spill years which had higher proportions of smolts transported. Spawner-recruit data is evaluated in Section 4.1 above.

Williams et al. argue this point in Submission 16, p.10: The basic hypothesis of low D values is NOT IN ALL CASES consistent with m over the range of data. The FLUSH model predicts low D_s in all years that survival of downstream migrants was low with assumed levels of TCRs that are in the 2:1 range. The CRiSP model assigns low D_s to periods when descaling was considered high and higher D_s when descaling was considered low. In the CRiSP model, levels of descaling are assumed low between 1980 and 1992 and high prior to 1980. In the FLUSH model, survivals of downstream migrants are consistently low from 1972 to the late 1980s, while TCRs over the period were generally ranged near 2:1. However, m was low in 1982 to 1985 and high from 1986 to 1992. Further, neither model predicted the low m that occurred in 1975 when the TCRs were low, downstream migrant survivals low, and descaling high. Finally, although Water Travel Time (WTT) is the main component in the FLUSH model to predict downstream migrant survival, WTT was NOT a good predictor of m (Deriso et al. 1996, page 5-17).

{There is considerable variation in the graph of m vs Water Travel Time but it did have a 0.58 correlation with ($p < 0.01$) despite the outlier data points of 1974 and 1976 with high

gas mortality (i.e., low WTT but high μ). Furthermore, a simple model using WTT outperformed models using either passage model (Deriso et al. 1996).}

c) Smolt-to-Adult Return Rates

The FLUSH/TRANS1 hypothesis of low D values is consistent with direct measures of transport survival (SARs), which indicate that transported spring/summer chinook are surviving at rates far below the PATH established goal of two to six percent and that no increasing trend since 1980 is apparent (Submission 22-18; Figure 4-5). SAR data is evaluated in Section 4.1 above.

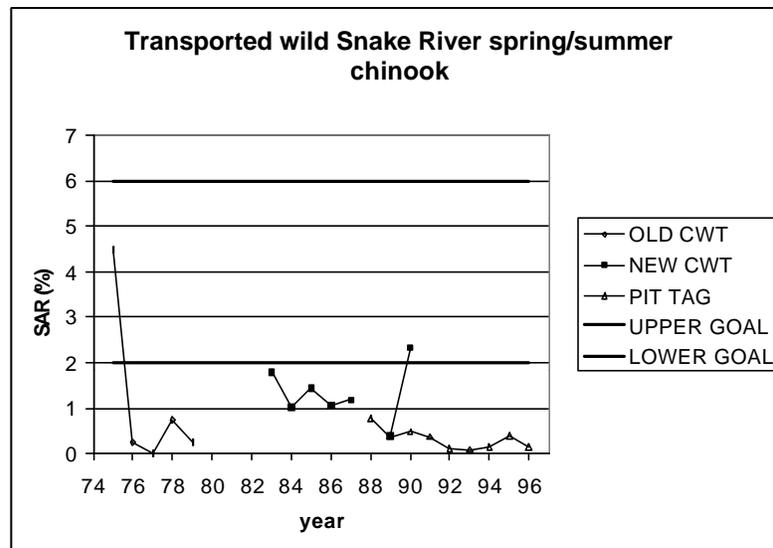


Figure 4-5: SAR's of wild Snake River spring/summer transported fish. DATA are from NMFS.

From Submission 16, p.10-11: It is a specious argument to use low SARs of transported fish in recent years to suggest that transportation is not working. Even with a nearly doubled increase in survival of fish that passed downstream through the hydropower system in recent years compared to 1969, 1970, and 1975, there was nearly a 10-fold decrease in SARs. This data overwhelmingly indicates that other factors, in addition to the hydropower system, are responsible for low SARs. It is NOT an issue of transportation.

Criterion 4) Validity of method of projection

It is useful to look at the trends in stocks that are projected by each of the passage/transportation models under the A1 (status quo) option. This is similar to the exercise conducted above for aggregate hypotheses (Figure 4-2). While it is true that stock projections are dependent on many hypotheses and assumptions, the sensitivity analyses in Section 3 suggest that passage / transportation models have the greatest influence on the relative ranking of actions. We used the BKD extra mortality / Markov climate hypothesis for these comparisons because future values of extra mortality and climate under these hypotheses are closest to their historical estimates. That is, the BKD hypothesis assumes that future extra mortality remains at the same level as the retrospective estimates. Markov climate samples randomly from retrospective climate effects, with autoregressive properties.

Results for Johnson Creek are shown in Figure 4-4 for combinations of passage/transportation model and life-cycle model with the BKD extra mortality / Markov climate hypothesis. These results assume

TURB1, low FGE, and low predator removal effects. We show these combinations not to suggest that all combinations are viable aggregate hypotheses, but simply to isolate the effects of passage/transportation models on projected spawners.

CRiSP/T4 projections are higher than FLUSH/T1 projections regardless of which version of the life-cycle model was used (Figure 4-4). CRiSP/T4 projects spawner abundances that are about double the 1990-1995 average of 150 spawners (actual escapement for Johnson Creek was 58 in 1996 and 236 in 1997). FLUSH projections are close to the recent year average, particularly in conjunction with the alpha model. Based on the diagnostic analyses in Appendix D, we conclude that the improvement is greater in CRiSP/T1 because of the large jump in retrospective D from the pre-1980 to the post-1980 period (Figure D-9). FLUSH/T1 D values are also higher prospectively than retrospectively, but the difference is smaller over all water years, which results in a smaller relative improvement in projected spawners.

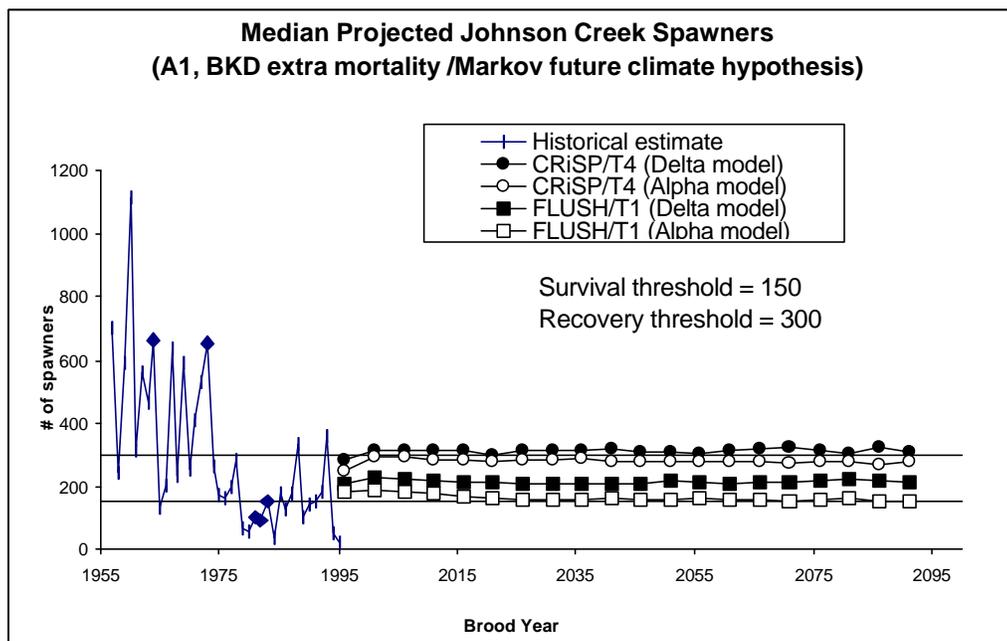


Figure 4-4: Projected spawners for Johnson Creek with various combinations of passage/transportation models, life-cycle models, and the BKD extra mortality / Markov climate hypothesis.

4.2.2 Life Cycle Models

4.2.2.1 Description of Alternative Hypotheses

Recruitment estimates are projected through the Bayesian Simulation Model (BSM), described in Deriso 1997, FY97 Report). Two alternative life cycle models are incorporated into BSM: the delta model, and the alpha model (see Fig. 3-1).

Brief Summary

Delta Model: The delta version of the life-cycle first calculates the Maximum Likelihood Estimate of total mortality, and a common year effect, from the historical spawner-recruit data. It then calculates post-Bonneville mortality as the difference between the MLE of overall mortality and the passage-related mortality from the passage/transportation models. It is assumed that: 1) year to year variations which

affect all stocks will be picked up the year effect parameter; 2) random variations in ocean survival will be assigned to a residual error term; and 3) variations among stocks in ocean survival are not systematic differences between Snake River and mid-lower stocks, as regional stock groups.

Alpha Model: The alpha version of the life-cycle model calculates post-Bonneville mortality directly from assumed relationships between mortality and various climatic factors (Astoria flows, PAPA drift, and a STEP function related to oceanic regime shifts), then calculates overall mortality as the sum of passage-related and post-Bonneville mortality. A different value for STEP is estimated for Snake River and mid-lower stocks, allowing for different responses to ocean survival (and other factors) in different regions.

Mathematical Representation

Delta Model:

The formal structure of the delta model is described in detail in Wilson et al. (1997; FY97 Report) and the Preliminary Decision Analysis (Appendix A pg. 87-91). Briefly, it can be described as follows:

$$\ln(R_{t,i}) = (1 + p) \ln(S_{t,i}) + a_i - b_i S_{t,i} - M_{t,i} - \Delta m_{t,i} + \delta_t + \epsilon_{t,i} \quad [\text{Eq. 4.2.2-1}]$$

for which,

- $R_{t,i}$ = Columbia River “observed” returns (recruitment) originating from spawning in year t and river sub-basin i
- $S_{t,i}$ = “observed” spawning in year t and river sub-basin i
- a_i = Ricker *a* parameter, which depends on sub-basin i
- b_i = Ricker *b* parameter, which depends on sub-basin i
- p* = depensation parameter
- $M_{t,i}$ = direct instantaneous passage mortality, which depends on year and region (computed from any passage model as $-\ln(N_b/N_0)$, defined below)
- $\Delta m_{t,i}$ = extra mortality rate, expressed as an instantaneous rate, which depends on year and region, and is calculated as the difference between total mortality m_t (described below) and passage mortality $M_{t,i}$;
- δ_t = year – effect parameter for year t (common year effects affecting both upstream and downstream stocks in year t)
- $\epsilon_{t,i}$ = normally distributed mixed process error and recruitment measurement, which depends on year t and sub-basin i.
- N_0 = total number of smolts at top of first reservoir in a season
- N_b = total number of smolts alive at Bonneville during a season

The total mortality, m_t , which includes both passage and extra mortalities, is estimated from stock recruitment data by using the maximum likelihood estimates for two parameters, \mathbf{m} and *X*:

$$m_t = \mathbf{m} + n_t X \quad [\text{Eq. 4.2.2-2}]$$

for which:

- m_t = Incremental total mortality between the Snake River Basin and the John Day project in year t ;
- $n_{t,i}$ = Number of first level dams (John Day, Dalles and/or Bonneville, called X-dams) stock i must pass in year t
- X = estimated total mortality per first level dam (does not vary with year).

As implemented in BSM, the total mortality (passage plus extra), m_t , constrains the value of extra mortality to:

$$Dm_{t,i} = m_t - M_{t,i} \quad [\text{Eq. 4.2.2-3}]$$

Thus assumptions which generate a greater amount of passage mortality $M_{t,i}$ have lower extra mortality $Dm_{t,i}$ in a given year. Other expressions for $Dm_{t,i}$ could be used (e.g., as a function of climate variables), generating a structure similar to the alpha model except for the common year effect δ_t .

Alpha Model:

The formal structure of the alpha model is described in Anderson and Hinrichsen (1997; Prospective analysis for the alpha model, in FY97 report), the Preliminary Decision Analysis Report (Appendix A, pg. 91-92) and Hinrichsen and Paulsen (Submission 3). The following summary draws from each of these sources.

Like the delta model, the alpha model uses a Ricker spawner-recruit relationship. However, it does not rely upon comparisons of upriver and downriver stock mortality estimates to derive a differential mortality (m) estimate that is assumed to bound the total (passage + extra) mortality (m_t). Instead, it calculates total mortality from passage mortality plus estimates of extra mortality. The alpha model is of the form:

$$\ln(R_{t,i}) = (1 + p) \ln(S_{t,i}) + a_i - b_i S_{t,i} - M_{t,i} - \mathbf{a}_{t,j} + \mathbf{e}_{t,i} \quad [\text{Eq. 4.2.2-4}]$$

The alpha term can be written as

$$\mathbf{a}_{t,j} = \mathbf{a}_n - \bar{\mathbf{a}}_n - \ln(DP_t + 1 - P) + \overline{\ln(DP_t + 1 - P)} \quad [\text{Eq. 4.2.2-5}]$$

where:

$R_{t,i}$, $S_{t,i}$, a_i , b_i , $M_{t,i}$, and $\epsilon_{t,i}$ are as defined above for equation [4.2.2-1]

$\mathbf{a}_{t,j}$ = extra mortality in year t for subregion j , here all upriver (Snake) stocks. A different alpha value applies to different subregions (i.e., Snake River, Lower Columbia River), hence the subscript j . The two averaged terms ensure that $\mathbf{a}_{t,j}$ sums to zero over 1952-1990. These terms are best associated with the Ricker a parameter, as outlined on page A-93 of the Preliminary Decision Analysis.

D_t = Ratio of post-Bonneville transport survival to post-Bonneville in-river survival for year t .

P_t = Proportion of fish arriving below Bonneville that were transported for year t .

In the version of the alpha model used to date, the series of extra mortalities is described by a linear relationship with further explanatory variables. The retrospective alpha series is modeled in the BSM as follows:

$$\mathbf{a}_n = c_1 / F_t + c_2 E_t / F_t + STEP_j \quad [\text{Eq. 4.2.2-6}]$$

where:

c_1, c_2 = estimated coefficients

F_t = Average Flow (in KCFS) at Astoria for year t during April-June.

E_t = Climate index variable (PAPA drift). Represents the latitude of a drifting object after three months drift starting at station PAPA.

$STEP_j$ = $STEP_j$ takes the value zero prior to 1975, and the estimated value $STEP_j$ afterwards. It is formulated to model the effect of a 1975 (brood year) climate regime shift, which has different effects in different subregions.

General Comments on the Two Model Structures

- If \mathbf{a}_t were set equal to $(Dm_{t,i} - \mathbf{d}_t)$ in equation [4.2.2-4], then the alpha model structure would be equivalent to the delta model structure in equation [4.2.2-1]. Therefore in years with relatively poor conditions (i.e., negative common year effects, $\mathbf{d}_t < 0$), the extra mortality in the alpha model, \mathbf{a}_t , will tend to be greater than the extra mortality in the delta model, $Dm_{t,i}$. In the Delta model, the common year effects \mathbf{d}_t and the total mortality m bound the magnitude of estimated extra mortality.
- Hinrichsen (Submission 25) points out that whereas prior to 1970, the year effect reflects changes in both lower river and Snake River stocks, after 1970 it only reflects changes in lower river stocks. *Comment: A common year effect is an explicit assumption as stated in Deriso (1996), and in the description at the start of Section 4.2.2.1 (point # 3).*
- Maximum potential productivity levels (Ricker 'a' levels) are generally greater in the delta model than the alpha model, though not after the alpha 'a' is corrected for the non-zero averages of extra mortality terms (i.e., average value of $\ln(DP+1-P)$, and average [STEP]) (Preliminary Decision Analysis, pgs. 38 and 93).
- The estimated values for STEP for Snake River and lower Columbia River stocks (equation 4.2.2.7) greatly depend on which passage and transportation models are used (Table 4-8). CRiSP/T4-TURB4 shows a 100-fold greater STEP value for Snake River stocks than lower Columbia River stocks (i.e., strong regional differences are required to fit the spawner-recruit data). This changes to a 50-fold difference with CRiSP/T4-TURB5. The CRiSP results in the alpha model are consistent with the hypothesis of regional differences in year effects.
- In Table 4-8, FLUSH/T1 shows low STEP values for both subregions, with little difference between them. This is primarily because FLUSH/T1 has lower D values for recent years than CRiSP/T4 (i.e., poorer post-Bonneville survival of transported fish), which absorbs enough of the decreases in recruitment that there is no need for the Snake River STEP parameter to have a value much above zero. The FLUSH results in the alpha model are consistent with hypothesis of common year effects.

- With both models, there is a higher value for Snake River STEP with TURB4 than under TURB5, because the higher reservoir survivals under TURB4 have to be counterbalanced with a stronger STEP to generate enough of a recruitment decline.

Table 4-8: Values of STEP estimated for Snake River and lower Columbia River stocks, for three different assumptions about turbine/bypass survival. Higher values of STEP result in greater decreases in recruitment from brood years 1975 to 1991. All runs assumed optimistic passage conditions (high fish guidance efficiencies (FGE1) and predator removal benefits (PREM3)).

Passage Assumptions	<i>CriSP/T4</i>		<i>FLUSH/TI</i>	
	Snake River	Lower Columbia	Snake River	Lower Columbia
TURB4	1.077	0.010	0.027	0.006
TURB5	0.855	0.017	0.002	0.009
TURB1	0.960	0.010	0.002	0.004

4.2.2.2 Evaluation

Criterion 1) Clarity

Peters and Marmorek (1997, A Generalized Prospective Modelling Framework, FY97 report) set out four principles for the prospective model:

1. It should be relatively simple to implement alternative hypotheses.
2. The model should allow for a common method of expressing alternative hypotheses.
3. The implications of alternative hypotheses on the decision to be made should be clear.
4. The model should be easy to explain in non-technical terms to decision makers.

Equations 4.2.2-1 and 4.2.2-4 are equally clear representations of changes in recruitment over time. The delta model assumes that common year effects affect both upstream and downstream stocks, whereas the alpha model does not. The validity of these assumptions can be questioned (they are examined below), but they are both clearly stated.

Both life cycle models would need to be rethought for simulating actions which had major changes in flow (e.g., major changes in storage reservoirs), but are adequate for simulating the range of actions currently considered in PATH. Equation 4.2.2-5 of the alpha model uses flow at Astoria (F) as an environmental variable, but F is unlikely to change significantly under any of the hydrosystem actions being considered (Hinrichsen, Submission 25). The year effect δ in the Delta model is uncorrelated with water travel time, and therefore there is no confounding in simulating changes in this water travel time from A1/A2 to A3 (Deriso, 1996). However, δ is weakly correlated with flow (Hinrichsen, Submission 25), so the Delta model might also need to be reconsidered for a different set of hydrosystem actions.

Criterion 2) Mechanism

A key distinction between the two life cycle models is the question of common year effects. Both models accommodate distinctive fluctuations in individual stocks through the $\varepsilon_{t,i}$ term. The delta model uses the δ_t term to extract that level of common variation in recruitment between lower Columbia River and Snake River stocks, presumably due to broad regional effects of climate and ocean conditions.

The delta model assumes that estuary and early ocean conditions have a similar effect on survival for stream-type chinook across regions of the interior Columbia River basin (i.e., Lower Columbia and Snake River). This assumption is based on a number of arguments summarized in Table 4-9, with associated counterpoints and rebuttals that were generated in either written or oral comments on an earlier draft of this report.

The alpha model allows different levels of extra mortality to be assigned to different regions by estimating region-specific values of STEP from the spawner-recruit data. With CRiSP/T4 input, significant differences in STEP are estimated for Snake River and lower river stocks (Table 4-8). Paulsen and Fisher (1997, FY97 Report) proposed four mechanisms which might generate such regional differences, which are summarized in Table 4-10, together with counterpoints and rebuttals. Note that some of these arguments are also relevant to the discussion of the regime shift extra mortality hypothesis in Section 4.2.3.

Table 4-9: Mechanistic arguments in support of the **delta** model representation (*with counterpoints in italics, and {rebuttals to these counterpoints in brackets}*). Main references listed in Schaller et al. 1996, (page 3-18 of FY96 report) and the Preliminary Decision Analysis (1998, Appendix A, pg. 98).

<p>Snake River, John Day River and Warm Springs River spring chinook smolts all have similar timing of arrival in the estuary, primarily in late April and May (Lindsay et al. 1986, 1989; Raymond 1979; Hymer et al. 1992).</p> <p><i>Coincident timing may no longer occur due to transportation (Anderson).</i> {This is a hydrosystem effect, accounted for by D (Wilson)}. {Without any dams, travel time is about a week; with transportation about 9 days; and with eight dams (in-river fish) about 3 weeks}.</p>
<p>The first few months at sea are generally considered critical in determining ocean survival of Pacific salmon (Pearcy 1988, 1992; Lichatowich 1993);</p> <p><i>This neither supports nor refutes either of the life cycle models. In either, the mortality could be set in the first year of life in the ocean. (Hinrichsen).</i></p>
<p>The importance of the first year in the ocean is supported by the close correlation of adult escapement estimates with jack estimates (Fryer and Swartzberg 1993), and the fact that (unlike fall chinook) ocean fisheries have little impact on spring/summer chinook (near absence of recovery of coded wire tag (CWT) captures in ocean fisheries (Berkson 1991, PSC 1994));</p> <p><i>The ratios of jacks to adults change in brood years with different age structure, so it is not only early life history that determines recruitment (Paulsen, comment based on Ch. 4, FY96.).</i></p>
<p>Although ocean recoveries of CWT spring/summer chinook are infrequent (Berkson 1991), the average annual proportion of CWT recoveries from ocean fisheries north and south of the Columbia appear to be similar between Snake and lower Columbia River hatcheries;</p> <p><i>The CWT evidence has very low power and should be discarded (Peterman) (evidence discussed below);</i></p>

A variety of evidence supports the idea that climatic changes have common effects on many stocks over a wide geographic range (Beamish and Bouillon 1993, Anderson 1996, Deriso 1996).

Mid-Columbia stocks show different extra mortality pattern from Lower Columbia stocks in alpha model (Anderson).

{These Mid-Columbia analyses have not been scrutinized, and there are many problems with the recruitment estimates, especially age structure estimates (Petrosky)}

Beamish and Bouillon (1993) and Anderson (1996) both use catch data, which are inappropriate for examining common year effects (Peterman; expanded comment below)

This evidence does not actually address the stronger assumptions actually employed in the Delta model, namely that the lower stocks act as a control for Snake River stocks (Hinrichsen, Submission 25)

Table 4-10: Mechanistic arguments in support of the **alpha** model representation and regional differences in extra mortality (*with counterpoints in italics, and {rebuttals to these counterpoints in brackets}*). Source of main references: Paulsen and Fisher (1997, FY97 Report).

The two stock groups are genetically distinct from one another, and are considered by NMFS to be in two distinct Evolutionarily Significant Units (Matthews and Waples 1991, West Coast Chinook Biological Review Team 1996). Also, the longer migration distances mean that they are in different physiological states when they enter the ocean.

Genetically distinct stocks often show co-variation in survival (residuals from $\ln(R/S)$ vs. S), but far distant stocks do not. For example, Bristol Bay sockeye stocks show within-region co-variation in survival, as do Fraser River sockeye, but there is no correlation in survival between these two groups/regions (Peterman; Adkison et al. 1996; Peterman et al. 1996).

Prior to 1970, Snake River stocks showed survivals much more similar to downstream stocks, which undermines the genetic distinctness argument. One has to imagine a genome that cues Snake River fish to seek out the bad ocean spots when poor conditions prevail, which doesn't seem likely to have evolved. (Petrosky et al., Submission 21)

Physiological state of the Snake River stocks is also a consequence of the hydrosystem (Weber).

Snake River stocks are more likely to return at age 5 than age 4 (Beamesderfer et al. 1996), giving ocean mortality additional time to affect Snake River fish;

Inter-annual variability is what is important, not average total survival rate. Sockeye data show much lower correlations in survival rate by brood year, than by year of ocean entry (Peterman). (Ed note: The SRP has suggested examining whether (R/S) varies with age of return for stocks which otherwise have similar exposure to stressors, but this analysis has not been completed.)

Ocean distributions of hatchery stream-type chinook collected and reared in the two areas are “quite possibly different” (based on the analysis of CWT recoveries in Paulsen and Fisher (1997));

The CWT evidence has very low power and should be discarded (Peterman) (evidence discussed below);

Even if the distributions were different, this does not necessarily mean that survival would be different (Cooney).

Recent ocean conditions since the mid-1970's have been more favourable in the Gulf of Alaska than the California current (Anderson (1996), Chapter 12, FY96 Report; SGI 1996; Hare et al., in press), and if Snake River stocks are more likely to be found in the California current than lower Columbia River stocks (based on Wahle et al. 1981), then Snake River stocks may have experienced poorer ocean survival.

Anderson (1996) and Hare et al. (in press) both rely on catch data to make inferences on survival, which is inappropriate (see comment below; Peterman)

The CWT evidence has very low power and should be discarded (Peterman) (evidence discussed below);

I know of no statistical test of inverse covariability between Gulf of Alaska salmon stocks and California Current salmon stocks. Such tests were not in Mantua, et al. (1997), Francis and Sibley (1991) nor Anderson (1996). Since then the paper by Hare, et al. (in press) has appeared. That paper compares catch data of the five species in the two regions by performing a spatial principle components analysis. Thus there is no specific statistical test of covariability between stocks or species in the two regions. The implications of results for each species are indicated by correlations between catch of that species and the first principle component for each region. The correlations for chinook salmon are not significant in any of the regions. So the statement is still true that there is no statistical test of inverse covariability between salmon stocks in the Gulf of Alaska and the California Current. (Botsford)

Two other points were mentioned at a meeting (July 30/31) to review an earlier draft of this report. First, Randall Peterman stressed the point that catch data alone provide very weak evidence to make inferences about similarities or differences in spatial/temporal patterns of survival (see Adkison et al. 1996, Peterman et al. 1996). Such inferences should be drawn from examinations of recruitment data (catch plus escapement) using residuals from spawner-recruit relationships (i.e., $\ln(R/S)$ vs. S). Examinations of data sets where both catch and recruitment residuals are available has demonstrated that catch data explain roughly only 20% to 44% of the variance in productivity of stocks, as measured by those residuals (Peterman et al. 1998, unpubl. data). Thus, catch data alone are inappropriate for drawing inferences on survival. These comments apply both here and to the discussion of extra mortality (Section 4.2.3). Second, Tom Cooney pointed out that the key issue is what causes upstream-downstream differences in survival, not the alpha vs. delta formulation. The alpha vs. delta differences are less significant in the results than are the effects of passage/transportation model and extra mortality assumptions.

The Return to the River report (ISG, 1996) provides some insights on mechanisms affecting synchrony in populations:

“Adjacent local populations are more likely to respond synchronously to environmental factors, whereas local populations that are more distant are more likely to experience asynchronous dynamics (Harrison and Quinn, 1989; Hanski, 1991; Rieman and McIntyre, 1993). However, the dynamics of geographically diverse populations can become correlated if, at some stage in their life history, individuals from diverse populations share a common environment, such as the ocean or a common migratory pathway (Rieman and McIntyre, 1993).

Salmon likely experience some degree of synchrony in dynamics due to the effects of natural environmental factors acting on regional scales in the ocean and freshwater...Synchrony can also be induced in common migratory pathways and the ocean as a result of mortality due to excessive harvest, construction of dams, degradation of mainstem habitats. Synchrony may be more likely if migration timing of diverse populations is seasonally restricted.”

Criterion 3) Consistency with Empirical Evidence

Some of the evidence related to mechanisms is cited in the previous section. Here we discuss four types of evidence relevant to this question: spawner-recruit data for wild spring-summer chinook in the Snake River and lower Columbia River; SAR data for Snake River spring-summer chinook; CWT data for Snake River and lower Columbia hatcheries; and escapement information for stream-type chinook from the Fraser River.

Spawner-Recruit and SAR Data

The spawner-recruit and SAR data are discussed above in Section 4.1.2. The delta model fits the spawner-recruit and SAR data significantly better than the alpha model, regardless of which passage model or TURB hypothesis is implemented (Tables 4-1 and 4-2). Some comments on these results from PATH participants (*with counterpoints in italics*) are listed in Table 4-11.

Table 4-11: Comments on tests of the alpha and delta model to spawner-recruit and SAR data. *Counterpoints in italics.*

<p>Table 4-1 understates the differences between the alpha and delta models because the alpha model's PAPA drift 'index' is not a physical measurement; it is actually a model prediction from the OSCURS ocean circulation model. Thus the number of parameters in the alpha model is actually greater than stated, and their AIC and BIC scores should be higher (Petrosky et al., Submission 21, and oral comments)</p> <p><i>OSCURS not used to fit spawner-recruit data and therefore its parameters should not be included in AIC/BIC tests (Hinrichsen).</i></p>
<p>The delta model is based on fits to both the lower river and Snake River data. The alpha model was only specified for the Snake River stocks, and not for the lower river stocks. Therefore, until an alpha model is developed for the lower river stocks, the AIC/BIC scores (which assess the fit to <u>both</u> lower river and Snake River data) unfairly penalise the alpha model (Hinrichsen, Submission 25, page 25-1). The fit to lower river stocks does not affect prospective predictions (Hinrichsen).</p> <p><i>The AIC/BIC statistics give the fits to the historical data for the models as they have been specified and utilized in PATH (Deriso).</i></p>
<p>The alpha model utilizes explanatory variables that have some scientific or physical meaning (e.g., flow), whereas the delta model uses factor variables with no physical or biological meaning. These results in a poorer fit for the alpha model, but a greater potential for scientific explanation (Hinrichsen, Submission 25, pg. 25-1).</p> <p><i>There have been many incarnations of physical variables used to explain extra mortality. The existence of historical correlations (after many tries with different explanatory variables) gives no more scientific validity than empirically derived estimates of common year effects. (Deriso (pers. comm.), Petrosky et al., Submission 21-7). The list of past candidate environmental variables used in PATH includes: 21 possible upwelling indices (Paulsen 1995); snow-pack, precipitation and spring flows (Paulsen 1996; Chapter 4); two sea surface temperatures and spring/fall transition dates (Paulsen 1997; revised Chapter 4) the North Pacific Index and one ocean upwelling index; E/F, E²/F, 1/F and a drought index term (Anderson and Hinrichsen 1997).</i></p> <p>{Many of these indices are correlated; the alpha model just chose a representative indicator. (Hinrichsen)}</p>

Several other retrospective analyses have been completed using the spawner-recruit data which have relevance for an assessment of these alternative life cycle models (i.e., Chapters 2, 3, 4, 5, and 12 of the FY96 Retrospective Analyses; revised Chapter 2 in FY97 Analyses). Some key points from these studies:

- Covariability in survival (residuals from fits to Ricker stock-recruitment relationships) is strongest within each of three sub-basins (Snake, Mid Columbia, and John Day), though moderate covariability was found between the Snake River and Mid-Columbia sub-basins, and between the Snake and Klickitat Rivers (Botsford and Paulsen 1997; revised Chapter 2);

- Stocks differ in their degree of statistical association between survival [$\ln(R/S)$] and selected indicators of ocean conditions and terrestrial climate, but there are no consistent differences in response between upstream and downstream stocks (Chapter 4, FY96 Retrospective Analyses).
- Common year effects (δ) derived from the spawner-recruit data (Figure 4-6); Deriso et al. 1996) are consistent with the hypothesis that the climate generally became less favourable for both Snake River and Lower Columbia stocks after brood year 1974, and are weakly correlated with the Pacific Decadal Oscillation index (Submission 24). The magnitude and variability of common year effects is a significant influence on estimated historical recruitment in the Delta model.

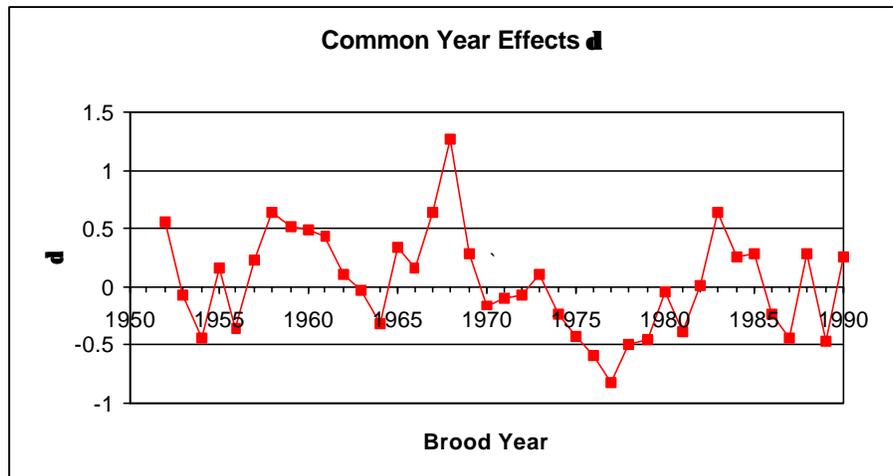


Figure 4-6: Common year effects derived from the Lower Columbia and Snake River spawner-recruit data (Deriso et al. 1996).

Coded Wire Tag Recoveries

Coded wire tag data from Snake River and lower Columbia River hatcheries were analyzed by Paulsen and Fisher (1997, in FY97 report), and Weber et al. (1997). We apply the evidence criteria used earlier.

(i) Applicability: Score = 4

Paulsen and Fisher (1997) have examined the 1983-1990 CWT data from PSMFC, which allow comparisons of upstream-downstream differences in ocean distributions. These data are from several lower Columbia River (Bonneville-McNary region) and Snake River hatcheries. Though 2.1 to 2.8 million tagged fish were released in this period, only 62 to 95 tags were recovered (ranges depend on whether or not test and first generation releases are included). Only a few brood years were included. Extrapolating from such a meager number of tag recoveries of *hatchery fish* to ocean distribution patterns of *wild fish* seems tenuous. In comments on an earlier draft of this report, Budy noted:

Libero (1986) demonstrates the sharp increase in the coefficient of variation at low numbers of observed recoveries. Assuming a 20% sampling rate (the coast wide standard), an average of 35 estimated adult equivalent recoveries across brood years corresponds to a coefficient of variation of approximately 30% for multiple tag codes released from a single location. At less than 24 observed recoveries (averaged across years), the coefficient of variation rises exponentially.

To date, all other PATH analyses have focused on wild fish in recognition of the differences between wild and hatchery stocks. Paulsen and Fisher found no differences in the regional distributions of wild and hatchery fish for a separate 1978-1982 data set for the Bonneville-McNary region, and concluded: “using hatchery fish as surrogate for naturally produced fish is not unreasonable when looking at broad patterns of ocean recoveries”. However, the power of these data to detect such differences is very low; there were only 13 wild fish recoveries in this data set, distributed over five regions (Table 2 in Paulsen and Fisher). Furthermore, there were no data to complete wild-hatchery comparisons for the Snake River subregion. These data are clearly insufficient to test the hypothesis that hatchery and wild fish have the same ocean distributions. As a counterpoint, Paulsen (Submission 25) notes that the run reconstructions assume the equivalence of hatchery and wild fish for estimating harvest.

The data from Wahle et al. (1981) describing recoveries of marked yearling chinook from brood years 1970-71 seem to have larger sample sizes, though it appears that Table 8 in Paulsen and Fisher (1997) only shows expanded (not observed) recoveries. The data do have the same limitation that they are from hatchery fish only, and cover only two brood years, which yielded contradictory results (see below).

(ii) Clarity: Score = 4

PSMFC data (4): The low numbers of CWT recoveries in the 1983-1990 data result in an average of only 6 to 9 tags per cell (2 subregions times 5 recovery areas), *after* lumping observations from different years together. Year effects thus add confounding to the results, and there are not enough data to analyze a fully interacted model (Paulsen and Fisher, 1997).

Wahle et al. (1981) data (4?): only expanded data are shown so it is difficult to assess the quality of these measurements.

(iii) Rigor: Score = 4

The spring/summer chinook CWT recovery data from PSMFC has not been published, nor has it been reviewed internally by PATH. A thorough review would include the details of the individual release groups, as well as recovery methods. Wahle et al. data: we assume that because these data were published in Marine Fisheries Review they have been peer-reviewed, though noting the limitations of such reviews (Starr 1998).

What do these admittedly weak CWT data say about ocean distributions? Are Snake River stocks more likely than Lower Columbia stocks to end up in the California Currents (i.e., south of the Columbia River), which may or may not have been worse for fish (see comment a4-3 in Table 4-10 [above], and discussion of regime shift in Section 4.2.3). Paulsen and Fisher found differences in recovery locations between Lower Columbia and Snake River stocks, but the subregion * recovery area interaction was not significant when bootstrapped. The 1983-90 data do not appear to show any significant interaction between hatchery location and ocean recovery (Preliminary Decision Analysis Report, page A-99) though the power to detect such differences is again very low.

The Wahle et al. data (1981) show very different results for the two brood years examined. In brood year 1970, fish from lower Columbia River hatcheries were less likely to be recovered in California fisheries than fish from Snake River hatcheries (8.1% and 17.0%, respectively, of the total recoveries from each subregion). The following year showed the reverse pattern, however, with 9.6% of the lower Columbia River recoveries in California, and only 3.5% of the Snake River recoveries (Table 9, Paulsen and Fisher 1997). Furthermore, there was no systematic recovery point

for these tags. Therefore, the data do not permit any firm conclusions regarding upstream-downstream differences in ocean distributions.

Fraser River Chinook Escapements (Bradford 1994)

Data from chinook populations in the Fraser River are relevant to the issue of common year effects in the Columbia River. Bradford (1994) examined spawning escapement of 12 Fraser River chinook stocks (1974-1991 data). These twelve stocks span a 600 km stretch of the Fraser River, not including the distances to tributary spawning areas. Thus these data are relevant to the question of whether other widely-spaced chinook stocks show common patterns of marine survival, or are markedly different.

- (i) Applicability = 3-4. These Fraser River stocks are all stream type chinook, as are the spring-summer chinook we are examining in the Columbia River. These are wild fish, not hatchery fish. However, they are from a different river system. The data are for escapement, not estimates of survival such as residuals from stock-recruitment curves. Variation in harvest rates also can affect the observed escapement patterns. Paulsen (Submission 25) points out that substantial changes in harvest did occur during this period, which likely increases the amount of correlation among stocks.
- (ii) Clarity = 2. The escapement data were collected by consistent methods over the period 1974-1991.
- (iii) Rigor = 1. This work has been published in Can. J. Fish. Aquat. Sci., and extensively reviewed both before publication and subsequently (as part of the hearings on the Kemano Completion Project for the Nechako river).

Bradford (1994) found that common year to year trends (the first factor in a factor analysis) explained 73% of the total variance in spawning escapement of 12 unregulated Upper Fraser River chinook stocks (1974-1991 data). The broad geographical distributions of these stocks indicate that chinook from widely separated streams can show very strong common year effects. The strength of the year effects is striking. When one of the outlier stocks was dropped from the data set, the year effects explained 86% of the variance in the escapement data of 11 stocks. (The Nechako River, which has only about 30% of its original flow due to flow diversions since 1952, showed a distinctly different temporal trend from the other 12 stocks (Figure 3 in Bradford 1994)). Though changes in ocean harvest rates undoubtedly affected these stocks over this period (Paulsen, Submission 25), the fact that such strong year effects were present suggests that overall marine survival (including climatic and harvest impacts) did not vary substantially for widely spaced stocks.

Criterion 4) Validity of Method of Projecting Hypotheses into the Future

The alpha and delta life cycle models use different methods of projecting future changes in salmon stocks, and also implement extra mortality hypotheses differently as a consequence of their different structures. We therefore discuss the form in which extra mortality hypotheses are implemented in this section, though the evidence for (and against) these hypotheses are presented in the next section.

Delta Model

The delta model projects future recruitment by:

$$\ln(R_y) = (1 + p)\ln(S_y) + a - bS_y - m_r + \ln(\mathbf{w}_y / \mathbf{w}_r) + \ln(\mathbf{I}_{n,y} / \mathbf{I}_{n,r}) + \mathbf{d}_y + \mathbf{e}_y \quad [\text{Eq. 4.2.2-7}]$$

where:

w_y / w_r = Ratio of system survivals estimated from a passage model runs for prospective year y and retrospective year r. These are matched by water year (e.g., whenever water year 1980 is selected in the future and used to compute a future system survival w_y , the system survival from 1980 is used for w_r).

$I_{n,y} / I_{n,r}$ = Ratio of post-Bonneville survival factors for non-transported smolts in prospective year y to retrospective year r.

Thus the delta model responds to the ratio of system survivals, and the ratio of post-Bonneville survival factors for a given water year. The value of λ_n for a particular retrospective year ($\lambda_{n,r}$) can be estimated from the total passage plus extra mortality (m_r), and the system survival estimate (ω_r), as described on pg A-101 of the Preliminary Report.

In the delta model, each extra mortality hypothesis must also generate a post-Bonneville survival factor $I_{n,y}$, so that the ratio term in equation [4.2.2-8] can be calculated:

1. For the BKD hypothesis (extra mortality is here to stay), it is simply assumed that $I_{n,y} = I_{n,r}$. Selection of system survival values is the same as for the regime shift hypothesis (see #3 below).
2. For the hydrosystem extra mortality hypothesis, it is assumed that for a given water year, the post Bonneville extra mortality factor for non-transported smolts ($1 - I_{n,r}$) is proportional to their in-river mortality ($1 - V_{n,r}$). For each of the sixteen retrospective years, BSM computes a factor (say $slope_r$) equal to $(1 - I_{n,r}) / (1 - V_{n,r})$, meant to capture the relationship between the post-Bonneville mortality factor and in-river mortality **in that year**. Then, for a future year y, the future post-Bonneville mortality factor ($1 - I_{n,y}$) is computed as $(slope_r)(1 - V_{n,y})$, where the retrospective year r and the prospective year y are again matched by water year (see Section 4.2.3 for more details and discussion). System survival (which is affected by D values) is provided by water year; this is different from the other two extra mortality hypotheses.
3. For the regime shift hypothesis, equation 4.2.2-7 is applied as stated, except that to get an estimate of $I_{n,y}$ the prospective water year y is matched with a retrospective year r that is in the same phase of the cycle. For example, until brood year 2005 (relatively poor climate) the coupled retrospective years are chosen from brood years 1975-1990, then from brood year 2006 for the next 30 years the coupled retrospective years are those chosen from brood years 1952-1974 (relatively good climate). The system survival in prospective years, ω_y , is calculated based on input M and P values, but the D values in the prospective years are chosen randomly from the 1980 to present water year estimates, which are thought to be representative of current D conditions. Note that with the regime shift hypothesis year effects **d** are selected according to a cyclical / autoregressive process **independent** of the 60-year regime cycle (Preliminary Decision Analysis Report, pg. A-118).

Alpha Model

In the alpha model, future recruits are projected by:

$$\ln(R_y) = (1 + p) \ln(S_y) + a - bS_y - M_y - \mathbf{a}_y + \mathbf{e}_y \quad [\text{Eq. 4.2.2 - 8}]$$

with a_y estimated from equation [4.2.2-5], with an appropriate value for $STEP$. In the alpha model the equivalent term to the delta model's $I_{n,r}$ is $\exp(1-STEP_r)$; (see Preliminary Report, pg. A-93 for equivalence of terms between the alpha and delta models, and equation [4.2.2-6] for the definition of the regime factor $STEP$). The three extra mortality hypotheses are then implemented as follows in the alpha model:

1. For the BKD hypothesis (extra mortality is here to stay), it is simply assumed that $STEP_y = STEP_r$. (Preliminary Report, pg. A-107).
2. For the hydrosystem extra mortality hypothesis, the $STEP_r$ term is assumed to be proportional to $(1-V_{n,r})$. The procedure is thus similar to the delta model approach except that only one *slope* factor is computed based on the average value of $V_{n,r}$ over 1975-1990, that is:

$$slope_r = (1 - \exp(-STEP_r)) / (1 - \overline{V_{n,r}}) \quad [Eq. 4.2.2 - 9]$$

Then the future $STEP_y$ is then calculated as $-\ln[1 - slope_r (1-V_{n,y})]$. Thus the two approaches are similar, except that the alpha model assumes all future years have the same relationship between passage mortality of non-transported fish and the post-Bonneville mortality factor. Putting all these terms together one gets:

$$STEP_y = -\ln[1 - (1 - \exp(-STEP_r))(1 - V_{n,y}) / (1 - \overline{V_{n,r}})] \quad [Eq. 4.2.2-10]$$

Details are described in the Preliminary Report (page A-104). The prospective V_n , D , and P are specified in the input file. In prospective years, there are new F, E variables according to the particular climate hypothesis.

It is instructive to work through an example application of equation 4.2.2-10. With CRISP/T4, the estimated value of $STEP_r$ is close to 1 (Table 4-8). If for a given management scenario there were a 50% reduction in passage mortality (i.e., $(1 - V_{n,y}) / (1 - V_{n,r}) = 0.5$), then $STEP_y$ would be 0.37, and the extra mortality term a_y would be significantly reduced in equation [4.2.2-5]. However with FLUSH/T1 $STEP_r$ is close to zero (Table 4-8) and so both $slope_r$ and $STEP_y$ will also be close to zero. Thus when the alpha model implements the hydrosystem extra mortality hypothesis with FLUSH/T1, the only possible change in a_y in equation [4.2.2-5] will be due to any changes in D and/or P in the $\ln(DP+1-P)$ term. (In FLUSH/T1, D changes with the predicted (T/C).)

3. For the regime shift extra mortality hypothesis, the $STEP_y$ value chosen for a given prospective year is the one which occurred during the same phase of the cycle. For example, until brood year 2005 $STEP_y$ is the one applicable to brood years 1975-1990 (i.e., $STEP_r$), then from 2006 for the next 30 years $STEP_y = 0$, which is the one applicable to brood years 1952-1974. With FLUSH/T1, this should make little difference since $STEP_r$ is close to zero. Thus, we would not expect the alpha model to show much of a regime shift with FLUSH/T1 input. In prospective years, there are new F, E variables. The value for F_y (Astoria flow in future year y) is chosen according to its negative correlation with unregulated water transit time (i.e., higher flow is associated with shorter unregulated WTT). The alpha value in prospective years, a_y , is then calculated in equation [4.2.2-5] based on $STEP_y$, E and F , an input P value, and a D value chosen randomly from the post-1980 period.

Differences in Projected Escapements

Delta model projections consistently show higher projected escapements than alpha model projections, for both passage models (see Figs. 4-4 in Section 4.2.1 and Fig. 4-13 in Section 4.2.3) in the next section of this report). This general result is related to differences in stock productivity, or estimated Ricker ‘a’ values (Fig. 4-7). We can see no independent means of verifying which Ricker ‘a’ values are more likely to be correct. In our analysis of Johnson Creek projections for A1 under the BKD extra mortality hypothesis (Section 4.2.1 above, Fig. 4-4), the passage model exerted much more influence than the choice of alpha or delta life cycle model. Hinrichsen (Submission 25) comments that Ricker *a* values in the Delta model are sensitive to a few observations in the Middle Fork of the John Day River, and that when these points are removed the alpha and delta models have similar values. Other sensitivity analyses, however, show that removal of the Middle Fork data has little effect on the jeopardy standards (Appendix G). Also, we have not yet looked at influential points in the Alpha model.

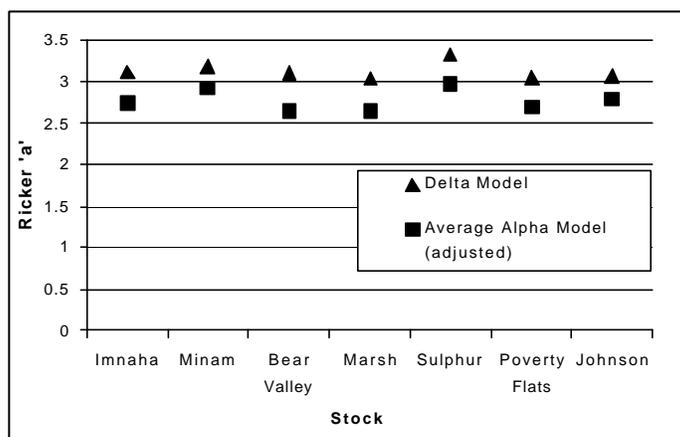


Figure 4-7: Comparison of stock productivities estimated by the delta and alpha models.

Implementation of Extra Mortality Hypotheses

Here we examine the ability of the alpha and delta models to implement extra mortality hypotheses; the strengths and weaknesses of these hypotheses are discussed in Section 4.2.3. It appears that both life cycle models reasonably implement the extra mortality hypotheses, though their behaviour is complex. The differences between delta and alpha model projections vary with both the passage model and the extra mortality hypothesis, as shown in the average projected escapements for Johnson Creek under A1 (Table 4-12). Relative to CRiSP, FLUSH showed a greater percentage change in escapement when the delta model is used in place of the alpha model (last two columns of Table 4-12). With both passage models, the percentage increase under the delta model was lowest for the regime shift extra mortality hypothesis, and highest for the hydrosystem-related hypothesis (last two columns).

Since the BKD hypothesis is expected to generate the lowest escapements, we use the ratio of mean escapements (i.e., Hydrosystem:BKD and Regime Shift:BKD in last two rows) as an indicator of responsiveness of each life cycle model to these extra mortality hypotheses. As expected from the low value for *STEP_r*, under FLUSH/T1, the alpha model is less responsive to the regime shift hypothesis with FLUSH/T1 input than with CRiSP/T4 input (Regime:BKD ratios of 1.09 and 1.64, respectively). That is, FLUSH retrospective estimates imply no regime shift, so none is implemented prospectively. The delta model also showed a greater response to the regime shift hypotheses with CRiSP/T4 input (Regime:BKD ratio of 1.55) than with FLUSH/T1 input (Regime:BKD ratio of 0.99). The pattern of escapements (Fig.

4-13 in Section 4.2.3) shows the contrasting response of CRiSP and FLUSH during the good regime of 2005-2035. This contrast is primarily due to differences in the retrospective system survival estimates, which are in turn the result of differences in estimated D values (Deriso, Submission 24).

The alpha model showed a similar response with the hydrosystem hypothesis under both CRiSP/T4 and FLUSH/T1 (i.e., same Hydro:BKD ratio of 1.35). This occurred despite the near zero value for STEP under FLUSH/T1, because the changes in D values with FLUSH/T1 are sufficient to shift α_t in equation [4.2.2-5].

Table 4-12: Effect of alpha and delta approaches on projected escapements for Johnson Creek under scenario A1. Table shows the mean of projected median escapements (5-year intervals) for 2000-2095, with two different passage model and transport assumptions, and three different extra mortality hypotheses. The escapement trends are graphed in Figures 4-4, Fig. 4-13 (Section 4.2.3). These runs all assume lower FGEs (FGE2) and predator removal efficiencies (PREM1), and historically high turbine/bypass survival (TURB1 – which implies low reservoir survival in future projections under FLUSH).

Extra Mortality Hypothesis	Alpha Model		Delta Model		Delta-Alpha		Delta-Alpha % Diff.	
	CRiSP/T4	FLUSH/T1	CRiSP/T4	FLUSH/T1	CRiSP/T4	FLUSH/T1	CRiSP/T4	FLUSH/T1
Hydrosystem	378	218	458	300	80	82	17%	27%
Regime Shift	458	176	483	212	25	36	5%	17%
BKD	279	162	312	214	33	52	11%	24%
Means	372	185	418	242	46	57	11%	23%
Hydro:BKD	1.35	1.35	1.47	1.40				
Regime:BKD	1.64	1.09	1.55	0.99				

4.2.3 Extra Mortality Hypotheses

4.2.3.1 Definition of Extra Mortality

Extra mortality is any mortality occurring outside of the juvenile migration corridor that is not accounted for by either: 1) productivity parameters in spawner-recruit relationships; 2) estimates of direct mortality within the migration corridor (from passage models); or 3) for the delta model only, common year effects affecting both Snake River and Lower Columbia River stocks. The magnitude and pattern of extra mortality therefore depend on which prospective modeling framework (alpha or delta) is implemented, and which passage model / transportation assumptions are used (CRiSP/T4 or FLUSH/T1). (See Section 4.1 for discussion of aggregate hypotheses.) Extra mortality can in theory occur either before or after the hydropower migration corridor. In the life cycle modeling, extra mortality hypotheses are applied to λ_n , the post-Bonneville survival of in-river fish. The extra mortality of transported fish is calculated from the extra mortality of non-transported fish by the 'D' parameter (the relative post-Bonneville survival of transported and non-transported fish; i.e., $\lambda_t = D \times \lambda_n$).

4.2.3.2 The Historical Pattern of Extra Mortality

The extra mortality hypotheses deal with the future responses of non-passage parts of the life cycle to particular management actions. Historical patterns of estimated extra mortality and their correlates is one form of evidence for assessing these hypotheses, though clearly less valuable than the results of actual

management experiments. The inferred pattern of extra mortality in non-transported fish depends on which life cycle and passage models are used. In transported fish, extra mortality inferences depend on the estimated D values, which in turn depend on passage model estimates of the survival of control fish (see Section 3.4.5). For the delta model, both CRiSP and FLUSH generate fairly similar patterns of extra mortality (Figure 4-8).

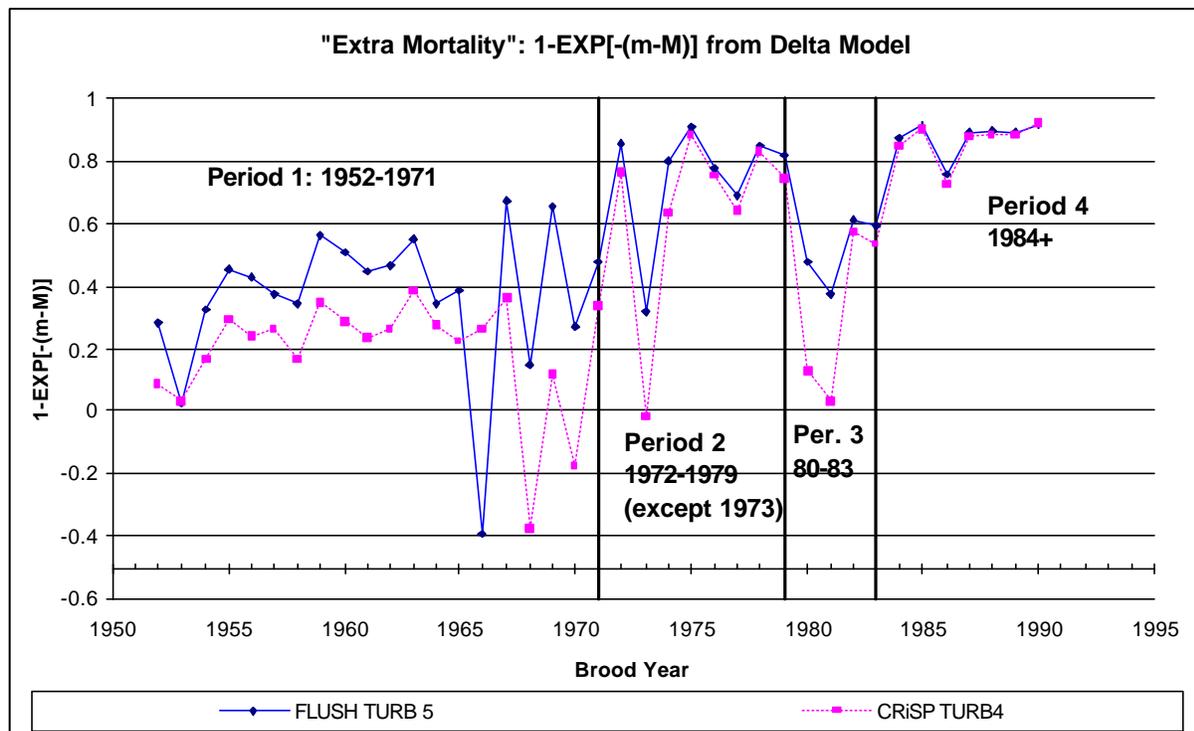


Figure 4-8: Extra mortality estimated from the delta model ($m-M$), where m is the total passage plus extra mortality rate, and M is the passage instantaneous mortality rate for all wild smolts (both transported and non-transported) from Lower Granite pool to below Bonneville Dam (see Section 4.2.2 for equations and definitions of terms). Both CRiSP and FLUSH estimates of M were used to compute ($m-M$).

The alpha model estimates a very different pattern in historical extra mortality. We computed the extra mortality term $a_{t,j}$ from equation [4.2.2.5] for the Snake River region, excluding the two averaging terms in this equation, which are better considered as part of the Ricker a parameter (as per pg. 93 of the Preliminary Report). Since we are only dealing with the Snake River region, we simplify the nomenclature to a_t . Note that $a_t = a_n - \ln(DP + 1-P)$ (equation [4.2.2-5]). To clarify the relative influence of different terms, we graph $\{1-\exp(-a_n)\}$ in Figure 4-9, and $\{1-\exp(-a_t)\}$ in Figure 4-10 (the latter is comparable to $\{1-\exp[-(m-M)]\}$ shown in Figure 4-8). The $\ln(DP + 1-P)$ term is responsible for all the higher frequency fluctuations in a_t with both CRiSP/T4 and FLUSH/T1 (compare Figs. 4-9 and 4-10). Equation [4.2.2-6] provides a STEP response in $\{1-\exp(-a_n)\}$ under CRiSP/T4 after brood year 1975 (but not under FLUSH/T1) (Fig. 4-9). Under both passage models the first two climate terms in equation [4.2.2-6] (i.e., $(1/F)$, inverse of Astoria flow and (E/F) , PAPA index divided by flow) make only negligible contributions to a_n add no observable fluctuations to $\{1-\exp(-a_n)\}$.

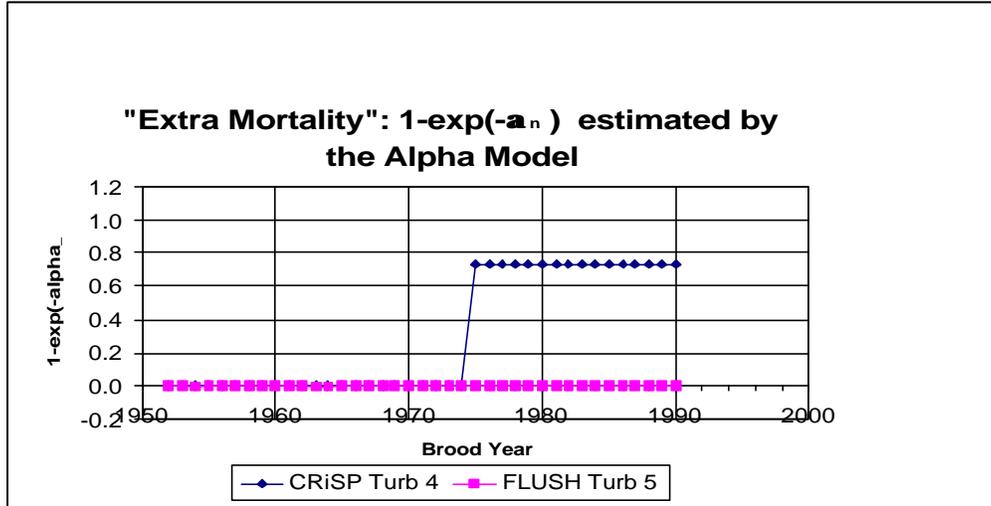


Figure 4-9: Pattern of extra mortality $\{1-\exp(-a_n)\}$, as estimated by the alpha model, for the same passage model assumptions as in Fig. 4-8. See equation [4.2.2-6].

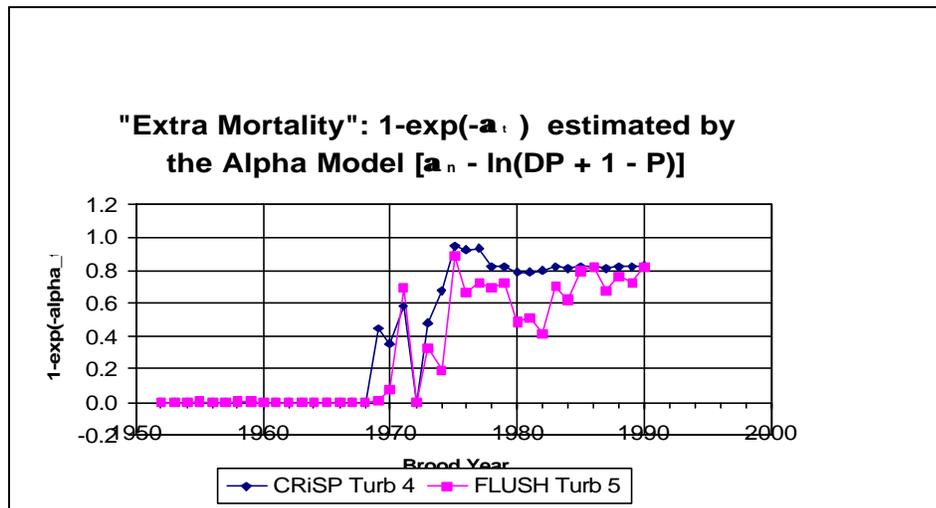


Figure 4-10: Pattern of extra mortality $\{1-\exp(-a_t)\}$, as estimated by the alpha model, for the same passage model assumptions as in Fig. 4-8. Here $a_t = a_n - \ln(DP + 1 - P)$; a_n is graphed in Fig. 4-9 (see equation [4.2.2-5]).

With FLUSH/T1, most of the post-1975 extra mortality is taken up by the extra mortality of transported fish via the D term (instead of the regime shift factor STEP). The lower D values with FLUSH/T1 generate a larger magnitude for the $-\ln(DP+1-P)$ term in equation [4.2.2-5] after brood year 1975, so that the estimated value of STEP is close to zero³. Under the alpha model, FLUSH/T1 does show a reduction in extra mortality during the 1980-83 brood years (due to an increase in estimated D values), but CRiSP/T4 does not, as it uses constant D values for this period.

³ As an example, values for D and P typical of CRiSP/T4 (0.63 and 0.9) generate a value for $-\ln(DP+1-P)$ of 0.4; keeping the same P value but changing the D value to 0.3 (a value typical of FLUSH/T1 retrospective runs) generates a value of 1.0 for $-\ln(DP+1-P)$.

In general, the extra mortality pattern estimated by the delta model is close to the inverse of the pattern of changes in SARs (Figs. 4-0b and 4-0c). The alpha model under FLUSH/T1 captures some of the patterns, but under CRiSP/T4 does not capture most of the recent variation in SARs. These graphs illustrate why the AIC and BIC scores for fits to the SAR data are better for the delta model than the alpha model, and better for FLUSH/T1 than CRiSP/T4 under the alpha model (Table 4-2).

Williams et al. (Submission 1) divide the pattern in Fig. 4-8 into four periods, and compare them to SAR patterns (Fig. 4-0a):

“(1) Period 1 (1952 through approximately the 1971 BY): “Extra mortality” was experienced by Snake River stocks at what might be considered a “background level.” When expressed using the $1-e^{-(m-M)}$ metric, the range of estimates is generally between 0-0.6. Estimates based on both passage models are synchronous through about BY 1965, and there is considerable uncertainty regarding the pattern in the 1966-1970 brood years. However, even with this uncertainty, the range of estimates is generally consistent with, or even lower than, the 1952-1965 BY range, so we have included these years in the first period.

(2) Period 2 (Approximately 1972-1979 BY): The range of “extra mortality” estimates was higher than the range during the first period for all years except BY 1973. The range using the described metric was approximately 0.6-0.9. The BY 1973 estimate, however, was as low as those during Period 1.

(3) Period 3 (1980-81 BY and Possibly 1982-83 BY): “Extra mortality” during the 1980-81 brood years dropped to levels experienced during Period 1. The “extra mortality” experienced by the 1982-83 BY was intermediate between the levels in Period 1 and Period 2 (approximately 0.6), so may be considered in either category.

(4) Period 4 (1984-90 BY): “Extra mortality” has been very high relative to that in other periods (generally >0.8).

We note that the general pattern of Snake River stock-recruitment residuals corresponds closely to the pattern of Snake River smolt-to-adult return rates (SAR), estimated using two different historical age structure assumptions (see Submissions 9 and 10). Although the years corresponding to each of the four phases of the temporal pattern vary slightly among these approaches, high SARs tended to occur during the first and third periods, while low SARs tended to occur in the second and fourth periods. As with the “extra mortality” estimates, SAR during the 1973 brood year (1977 migration year) was considerably higher than other SARs during the second period.”

4.2.3.3 Description of Extra Mortality Hypotheses Modeled to Date

Three alternative extra mortality hypotheses have been modeled so far: hydrosystem, BKD (or stock viability), and regime shift. The following provides a brief verbal description of each hypothesis; the mathematical representation of the hypothesis and specific evidence is examined in more detail below. Additional extra mortality hypotheses that have been recently proposed (hatcheries, multi-factor) are discussed in Section 4.2.3.4.

It is important to recognize that each of the three modeled extra mortality hypotheses are linked to other sets of assumptions, providing hydrosystem, climate and hatchery components to the aggregate hypotheses that are simulated. In other words, when applied to the whole population, these three hypotheses are not ‘single factor’ hypotheses, despite their names. Independent of the extra mortality hypotheses, the passage models estimate direct effects of the hydrosystem; alpha and delta life cycle models, together with future climate hypotheses, provide mechanisms for incorporating climate effects; and the ‘D’ parameter absorbs any impacts of hatchery fish on the post-Bonneville survival of transported

fish. Where these hypotheses differ is in the level of importance of these factors in affecting the post-Bonneville survival of **in-river** fish (λ_n).

There are several difficulties in assessing these extra mortality hypotheses (discussed in more detail later): 1) all three hypotheses have similar expectations about past patterns of λ_n ; 2) estimates of λ_n are derived from other parameters and model estimates, including m , M , D , and P (see Figure 3-1), all of which are subject to errors; and 3) there are many alternative indices by which to assess shifts in climate regimes.

EM1. Hydrosystem-related extra mortality (Marmorek and Peters, 1998; pg. A-95 to A-104)

Description: The completion of the Federal Columbia River Power System in the late 1960's through the mid-1970's and subsequent operation, has increased the direct and delayed mortality of juvenile migrants, which resulted in considerably sharper declines in survival rates of Snake River spring and summer chinook stocks (over the same 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).

Mechanisms: Proposed mechanisms causing delayed mortality of smolts migrating through the hydrosystem include: altered saltwater entry timing poorly synchronized with the physiological state of the smolts; stress from crowding and injury (including descaling) during bypass, collection and holding; increased vulnerability to disease outbreak (e.g., BKD and fungal infection) due to stress and injury; and increased vulnerability to other stressors in the environment or to predation, particularly by northern squawfish.

Management Implications: If post-Bonneville extra mortality of non-transported smolts is related to the above hydrosystem effects, then significant reductions in hydrosystem related passage mortality (e.g., A3) should also cause significant reductions in extra mortality of non-transported fish. The effects are likely to vary from year to year with the amount of flow.

EM2. BKD or Stock Viability Hypothesis (Marmorek and Peters, 1998; pg. 105-106)

Description: This hypothesis proposes that the viability of Snake River stocks declined as a direct or indirect result of the hydrosystem construction in the 1970s. It also proposes that the current extra mortality is not related to either the hydropower system or climate conditions, and is here to stay, even if hydrosystem direct mortality is reduced and / or the climate improves. This hypothesis was originally proposed as a 'worst case' for preliminary modeling. Since this hypothesis overlaps strongly with the recently proposed hatchery extra mortality hypothesis (EM4), it is generally accepted in PATH that EM4 should replace EM2 in future modeling.

Mechanisms: One mechanism proposed 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 proposed stock viability mechanism 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.

Management Implications: This hypothesis represents a worst case scenario for management, in that neither hydrosystem improvements nor regime shifts would reduce extra mortality.

EM3. Regime shift (Marmorek and Peters, 1998; pg. A-107 to A-113; Ch. 12 FY96 report)

Description: 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. 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.

Mechanisms: 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.

Management Implications: Mantua (1997) points out that some fish management goals may not be attainable when environmental conditions are unfavorable, and conversely, in a period of favorable conditions, managers shouldn't claim credit for a situation that may be beyond their control. The implication for the Columbia River is that the climatic regime could undermine or enhance the survival improvements achieved by various management actions. Hare et al. (in press) stress however that this does not imply holding back on efforts to restore watersheds during unfavorable periods, since:

“Pacific salmon have evolved their metapopulation structures over millennia to deal with variations in ocean conditions..., [and] the physical template provided by naturally functioning watersheds is the ultimate source of ‘climate insurance’ necessary for wild salmon populations to persist.” Hare et al. (in press)

4.2.3.4 *New Extra Mortality Hypotheses (not yet modeled)*

Two additional extra mortality hypotheses have recently been proposed involving hatcheries (EM4a and EM4b below; Submissions 1 and 2). Note that EM4a discusses total hatchery and steelhead production, while EM4b looks only at hatchery releases of yearling chinook. A “Multi-factor” hypothesis (EM5) has also been proposed (Submission 12). There has not been much time to thoroughly evaluate these new hypotheses. The hatchery hypothesis is critically reviewed in Submission 19.

EM4a. Hatchery Hypothesis (Williams et al.; Submission 1):

Description: Snake River hatchery smolt production, particularly that of steelhead, increased greatly during the time period experienced by Snake River spring/summer chinook salmon smolts of the 1968-1990 brood years. This increase is generally coincident with an increase in “extra mortality” experienced by most of those brood years. A negative effect of hatchery production on wild spring/summer chinook survival is hypothesized, particularly for brood years 1984-90.

Mechanisms: The proposed mechanisms are reduced growth rate, and increased stress, predation, and disease transmission. Effects are likely greater for Snake River stocks than for lower Columbia River stocks because: 1) Snake River wild smolts are exposed to hatchery smolts for a longer distance; 2) Snake River wild smolts are the only ones concentrated with hatchery fish in barges; 3) all seven Snake River

index stocks encounter hatchery fish in the migration corridor above the first dam, whereas only three of six lower river index stocks do (Wind, Klickitat and Warm Springs); and 4) natal streams of Snake River stocks may be more nutrient depleted than those of lower river stocks, affecting fish condition and exacerbating the effects of hatchery interactions above the first dam.

Management Implications: Williams et al. have the following discussion on the potential impacts of hatchery activities under management actions A2 and A3:

A2 -Maximize transport: If the effect of hatchery fish is significant and is primarily the result of stress before or during bypass capture and barging, the prospective effect would most likely remain unchanged from recent levels or be increased as a result of exposure of an increased proportion of the run. If the effect of stress is primarily the result of interactions during the in-river migration below the collector projects, the maximum transportation option may result in a decrease in effect due to reductions in the proportion of the run remaining in-river below the collector projects. If the effect is due to interactions on entry to the estuary or ocean, maximizing transport will increase mortality as a result of a higher proportion of the hatchery and wild smolts reaching the estuary and ocean.

A3 -Snake River Drawdown In general, interactions with hatchery fish would be reduced under this option. Migration time through the area currently impounded above Lower Granite Dam would be reduced. Collection and handling at the Snake River projects would be eliminated resulting in less crowding. The Snake River collector projects currently remove more than 50% of the steelhead juveniles prior to passage below Little Goose Dam. Drawdown would eliminate Snake River transportation, resulting in significant increases in the number of smolts migrating in the mainstem below Little Goose Dam. To the extent extra mortality is increased by competition for food or by exposure of wild spring/summer smolts to hatchery smolts, detrimental impacts could be increased. This may be particularly true at McNary Dam, where a larger number of smolts will arrive at the project with the elimination of Snake River transportation than under current conditions. Alternatively, negative interactions in the corridor below the Snake River collector projects could be reduced under this option, the amount of change being dependent upon the relative importance of cumulative stress prior to the reach and to the nature of the bypass systems remaining in place.

Reductions in the number of hatchery yearlings released during the wild spring/summer chinook migration would reduce extra mortality under all of the combinations of possible interactions and system options described above. Plotting estimates of Snake River wild spring/summer chinook extra mortality against hatchery steelhead release levels indicates a possible non-linear relationship (Figures 6,7 in Submission 1). If that relationship holds, reductions in hatchery releases to or below those levels prevalent in the mid to late 1970's would be required to significantly reduce extra mortality.

EM4b. Paulsen and Hinrichsen Hatchery Hypothesis (Submission 2):

Description: “Variation in releases of Snake River hatchery spring/summer chinook is associated with variation in extra mortality of naturally produced Snake River spring/summer chinook. Reducing hatchery releases would reduce extra mortality, independent of hydrosystem actions.”

Mechanisms: Paulsen and Hinrichsen do not discuss mechanisms in much detail, other than to say:

“Biological mechanisms that could explain these results might include horizontal transmission of disease, competition for food or other resources, or hatchery fish may help maintain larger predator populations than would exist otherwise. In addition, stress during transportation may play a role.” (pg. 5, Submission 2)

Management Implications: Paulsen and Hinrichsen state:

“If the correlations between hatchery releases and mortality reflect an underlying causal relationship, then reducing hatchery output should reduce extra mortality. The correlation analysis indicates the direction of the hypothesized relationship, but does not say anything about the size of that relationship.” (pg. 5, Submission 2)

EM5. Multiple Factor Hypothesis (Submission 11; Submission 16, pgs. 8-10)

Description: Extra mortality of non-transported smolts is related to multiple factors. These include changes in: 1) hydrosystem flow related to the construction of storage reservoirs, 2) ocean and climate conditions, 3) estuary predator populations, 4) hydrosystem passage stress; and 5) hatchery production. As stated in the introduction to Section 4.2.3.3, other extra mortality hypotheses also recognize a mix of hydrosystem, hatchery and ocean/climate effects (when combined into an aggregate hypothesis). The multiple factor hypothesis would link these factors directly to the post-Bonneville survival of **in-river** fish, by estimating coefficients for these five factors to generate an a_n (see equation [4.2.2-5]) which best fits historical spawner-recruitment information. The justification for selecting these five factors rest on two criteria: 1) they have experienced significant changes coincident with the construction of the hydrosystem, and 2) there are ecological bases for how they may affect the extra mortality of the non-transported fish.

Mechanisms: From Anderson (Attachment 11):

1. *River discharge as it affect the dynamics of the estuary and the dynamics of the river plume in the ocean have been shown to correlate with early-ocean survival (Pearcy 1992).*
2. *Ocean and climate regimes shifts are correlated with the survival and catch of a large number of salmonid species (Anderson 1996 and in press). The drift parameter is an index of climate regime shifts (Ebbesmeyer et al. 1998).*
3. *Avian predators in the Columbia River estuary have been shown to consume more smolts than in-river fish predators (Roby et al. 1998). Submission 16 (page 9-10) elaborates in more detail on recent increases in bird predation.*
4. *The explanation for the impacts of the hydrosystem on extra mortality was developed in the PATH Preliminary Decision Analysis Report (1998).*
5. *The effect of hatchery smolt interactions with wild smolts could occur though several mechanisms, including competition for food and transfer of hatchery-borne disease into wild populations (as developed by NMFS).*

Management Implications: These are not discussed by Anderson in Submission 11, but clearly, if extra mortality is distributed among five different factors, changes in any one factor (such as the hydrosystem) will have less benefit.

Criterion 1) Clarity

Are each of these extra mortality hypotheses clearly implemented in the models? Some of the issues associated these hypotheses are listed in Table 4-13 (not a comprehensive list):

Table 4-13: Potential problems with the clarity of how extra mortality hypotheses are currently implemented in the models (EM1 to EM3), or how they might be (EM4 and EM5). *Counterpoints listed in italics.*

Hypothesis	Issues
EM1. Hydrosystem	Analyses of I_n vs. V_n (defined in Section 4.2.2) indicate inconsistency in the relationship between extra mortality of non-transported smolts and their in-river mortality (see discussion under Evidence below, and Hinrichsen and Paulsen – Submission 3). <i>Alternative implementation of this hypothesis (Appendix H) does not have this problem, and produces similar results.</i>
EM2. BKD	Most straightforward to implement in models because it replicates past. Problems with this hypothesis are in its proposed mechanisms and supporting evidence (see below).
EM3. Regime Shift	PAPA index is a model, not a measurement. Model has not been reviewed by PATH. (Submission 21). Other issues discussed under mechanisms and evidence. <i>PAPA index is present in all EM hypotheses implemented through the Alpha life cycle model.</i>
EM4. Hatcheries	Similar confounding problems to EM5, but not as many variables. Confusion in relative magnitude of impacts of hydrosystem and hatcheries. Many components of the hatchery hypotheses are inseparable from the other competing hypotheses for extra mortality (or percent of fish transported), so this hypothesis has poor ‘clarity’ <i>All extra mortality hypotheses have confounding effects.</i> Hatcheries were built as mitigation for fish and wildlife losses. Correlations of extra mortality with hatchery smolts may be confusing cause and effect. If the dams were removed, hatchery impacts would be less, and hatchery production as mitigation for hydrosystem effects would end.
EM5. Multiple Factors	MLE model would likely be over-parameterized due to colinearity of independent variables, and posterior distributions for parameters could be so wide as to render almost any prediction possible. <i>All extra mortality hypotheses have confounding effects.</i> Points raised in Submission 16 (pg. 10): Bird populations increased in later years of time series of most extra mortality estimates (hence little contrast); not clear how model would appropriately allocate predation effects to various survival and productivity terms. Only one of the five factors listed (i.e., hydrosystem) can explain upstream-downstream differences in survival.

Criterion 2) Mechanism

In going through past PATH reports, it is difficult to separate out evidence related to a mechanism of mortality (criterion 2), from evidence showing the consistency of the proposed stressor with various survival measures, and that mortality *has* occurred to the magnitude specified in the hypothesis (criterion 3). Table 4-14 is a proposed framework for organizing and evaluating existing evidence regarding mechanisms. It was synthesized from material in the Preliminary Decision Analysis Report and attached Submissions for this report. There has not been enough time to include all of the mechanisms and associated evidence, nor to evaluate it according to the stated criteria. Individual mechanisms have been numbered for ease of reference.

In Table 4-14 some sources of mortality (e.g., crowding of hatchery and wild fish at forebays) are labeled a “hydrosystem effect” in one hypothesis, and a “hatchery effect” in another, when they are likely an interaction of both stresses. We have not had time to separate out each of the independent mechanisms (e.g., impacts above Lower Granite pool could not be hydrosystem effects) from those which are confounded, and where possible weigh the evidence for each impact pathway. Clear separation of these effects is only possible through strong manipulations of different combinations of factors through adaptive management experiments, ideally with contrasts in both space and time. The benefits and risks of such experiments have not yet been evaluated by PATH.

Table 4-14 Key mechanisms proposed for each extra mortality hypothesis, and source of more detailed descriptions.

Mechanism	Sources
Extra Mortality Hypothesis: EM1. Hydrosystem (Prelim. Dec. Anal. Report, pg. 95)	
<p>Altered timing of saltwater entry by hydropower system may affect later survival of fish, since physiological changes (e.g., osmoregulation) in preparation for saltwater residence are time-dependent.</p> <p><i>Spring/summer chinook smolts can adapt to saltwater over a wide range of timings of ocean entry (Hoar 1976).</i></p>	CBFWA 1991; Fagurland et al. 1995
<p>Stress from crowding and injury during bypass, collection and holding leads to increased vulnerability to disease outbreak (e.g., BKD and fungal infection). Disease causes delayed mortality in saltwater stage. See 4.2 (hatchery).</p>	Basham and Garrett 1996; Williams and Matthews 1995; Mundy et al. 1994; Raymond 1988; Williams 1989. See EM4 (hatchery) refs and Submissions 21 and 19.
<p>The hydrosystem has altered the system such that conditions are now optimal for predation on salmonid smolts. The altered system, in addition to the stress and injury associated with the hydrosystem (above), causes increased vulnerability to predation below Bonneville, particularly by northern squawfish.</p> <p><i>Predation within the migration corridor is already accounted for by passage models.</i></p> <p><i>I_n should decline with the predator control program in prospective simulations.</i></p> <p>{This is accounted for, because V_n improves with predator control, which in turn reduces λ_n under EM1.}</p>	Submission 21
Extra Mortality Hypothesis: EM2. BKD / Stock Viability (Prelim. Dec. Anal. Report, pg. 105, Submission 1 in App. A)	
<p>BKD in Snake River hatchery chinook transmitted to wild fish causes mortality in freshwater and/or ocean that will exist with or without dams.</p>	Submission 1

<p><i>BKD is no more prevalent in Snake River systems than lower river, and infectious diseases would not constrain recovery of these stocks if migration conditions in the mainstem are corrected.</i></p> <p><i>Since potential for transmission is greatest under transportation, (and we have no certainty that transmission of BKD is completed), impacts of BKD would be reduced under drawdown.</i></p> <p>{The hypothesis pertains to populations already infected, not new disease transmission.}</p> <p>{Snake River fish have longer exposure to various types of pollutants which could express itself in higher rates of disease related mortality, or susceptibility to predation. This pollution would remain after drawdown. (J. Geiselman)}</p> <p><i>As modeled, the BKD hypothesis is independent of biological mechanisms (i.e., extra mortality simply here to stay), and is untestable.</i></p>	Submission 20
<p>Loss of stream nutrients from salmon carcasses in low productivity, high elevation streams give Snake River fish lower energy reserves, reducing their survival during passage and ocean entry.</p>	Submission 1
<p><i>Such a mechanism should show up as compensatory decline in (R/S) but there is could find no evidence of this. Ricker b parameter accounts for this effect already. This mechanism would in theory apply to all EM hypotheses, and is not an irreversible situation.</i></p> <p><i>Ln(smolts/spawner) has not decreased significantly over time.</i></p> <p>{There are several data uncertainty and consistency issues with smolt estimates in this analysis (J. Geiselman)}</p> <p><i>Downward trends for Lemhi River (higher nutrients) same as other stocks.</i></p> <p>{Sedimentation, dewatering problems are also factors in the Lemhi River (J. Geiselman)}</p>	<p><i>Deriso (1997; FY97 report); Section 5.7 of this report</i></p> <p><i>Petrosky and Schaller 1996</i></p>
<p>Extra Mortality Hypothesis: EM3. Regime Shift (Submission 17 (this report); Prelim. Dec. Anal. Report, pg. 107-113; Ch. 12, FY96 report)</p>	
<p>Pacific Decadal oscillation occurs every 30 years, changing air and ocean temperatures, as well as stream flow. This alters zooplankton production, causing major changes in Columbia River fish stocks' production, concurrent with inverse changes in Alaskan stocks.</p>	<p>Minobe 1997; Hare et al. 1997; Mantua et al. 1997; Pearcy 1992; Francis et al. 1997; Brodeur and Ware 1992; Sugimotoa and Takokoro 1997; Roemmich and McGowan 1995; Olovina et al. 1995; Ebbesmeyer et al. 1998.</p>
<p><i>These papers do not provide evidence of differential climate effects on Snake River and lower Columbia yearling chinook survival unrelated to the hydrosystem, but only evidence of common year effects across broad geographic areas for a variety of species (pink, coho, sockeye)</i></p>	<p><i>Schaller et al. in Prelim. Report (pg. A-107).</i></p>

<p><i>Many of the cited papers do not have rigorous evidence of spatial / temporal correlations. Peterman (July 30th meeting) noted that there was a 3-fold increase in Ricker alpha parameter for nine sockeye stocks in Bristol Bay after 1977, but no evidence for changes in Ricker alpha parameter for Fraser River sockeye. There is considerable evidence that some changes have occurred, but not for the strength of EM3 as formulated.</i></p>	<p><i>Botsford comments (A4-3 in Table 4-10, section 4.2.2 of this report; pg. A-113 of Preliminary Decision Analysis).</i></p>
<p><i>Several of the cited papers use catch data (weak evidence) as a surrogate for survival; should only use residuals from spawner-recruit curves.</i></p>	<p><i>Peterman’s comments (Table 4-10 and following).</i></p>
<p>Changes in river discharge affect the dynamics of the estuary and early ocean survival.</p>	<p>Pearcy 1992.</p>
<p><i>This is an effect of the hydrosystem as well, not just a climate effect, and does not explain differential decline in Snake River stocks.</i></p> <p>{River discharge is not affected by the actions we are considering. (Hinrichsen)}</p>	<p><i>Petrosky et al. (Submission 21)</i></p>
<p>Extra Mortality Hypothesis: EM4. Hatcheries (Submission 1)</p>	
<p>The large number of hatchery fish and their larger biomass reduces carrying capacity of the migration corridor, limiting growth and energy reserves of wild spring/summer chinook. This makes these fish more susceptible to predation and less able to make seawater transition.</p> <p><i>If 20 million hatchery smolts (with poorer survival than wild smolts) strain carrying capacity, how did the corridor handle juvenile production from the estimated 7.5 million salmonid adults and 150 million smolts in pre-development times? Smolt survival generally density-independent, and smolts don’t eat much as they move through cold, fast, turbid water. No observed declines in smolts/spawner in Snake River aggregate. Recent reviews (CBFWA) haven’t demonstrated hypothesized effects. Carrying capacity declines would need to be only for Snake River stocks to explain extra mortality.</i></p> <p><i>Lack of strong evidence for effects of hatchery fish on wild chinook in the Snake and Columbia rivers. Survival of hatchery fish is less than that of wild fish, suggesting that wild fish may out compete hatchery fish.</i></p> <p>{Ratio of steelhead:chinook smolts greater now than before, and carrying capacity has probably declined considerably since pre-development times due to much less estuarine habitat, non-point source pollution, etc. ISAB has moratorium on future hatchery construction.}</p>	<p>Submission 1, pg. 7-8</p> <p>Submission 19</p>

Hatchery fish increase stress levels on Snake R. fish in places where fish congregate (forebays), increasing susceptibility to: a) predation; b) disease; and reduction of energy reserves. <i>This mechanism is not separable from the effects of the hydrosystem. Stress (measured by saltwater challenge tests) greatest at dam sites such as raceways. See EM1 above (2nd point). Hatchery impacts show up under A1/A2 in D values due to crowding with 8 dams; these effects removed under A3.</i> {The fact that actions are not separable does not negate the hatchery hypothesis (Cooney)}	Submission 1, pg. 8-9 Submission 19
Hatchery fish affect Snake River spring/summer chinook more strongly due to longer distance of co-migration; greater concentration mechanisms; greater likelihood of encountering hatchery fish; more nutrient depletion (see EM2 above); and longer exposure to water pollution.	Submission 1, pg. 10-11
<i>General criticisms pertaining to all of the above mechanisms: Effects of the hypothesized mechanisms likely to vary dramatically with environment, life stages of the species present, relative and total densities, history of co-evolution, etc. Lack of data and experiments aimed at explicitly understanding hatchery-wild interactions. The direction and magnitude of the effects are unknown and could actually be opposite of those hypothesized. Correlation is not causation: I_n is as well correlated with percent of fish barged as with numbers of hatchery smolts.</i>	Submission 19
Extra Mortality Hypothesis: EM5. Multiple Factor Hypothesis (Submission 11, Submission 16; pg. 8-10)	
EM1; EM3; EM4; plus:	
Trends in SARs over time imply that different factors have been important at different times.	Submission 16 (pg. 8-9); Submission 9
<i>Avian predators in the Columbia River estuary consume more smolts than in-river predators. Caspian terns on Rice Island have increased from 1000 breeding pairs in 1984 to 8000 pairs by 1997.</i> <i>Of the five hypothesized factors, only hydrosystem effects can explain upstream-downstream differences.</i> <i>Not clear what stock composition is of prey consumed by birds (i.e., if transported or hatchery chinook, or if coho / steelhead, then cannot explain I_n). Transported fish have more shallow distribution than in-river fish (Shreck, unpub.), and are probably more vulnerable to bird predation.</i>	Roby et al. 1998. Submission 16, pg. 10.

Criterion 3) Consistency with Empirical Evidence

There is no direct measurement of extra mortality of either transported or non-transported fish. Ideally one would have marine survival estimates from Bonneville to Bonneville, but there are no such estimates, and any estimates would likely have huge variances with such small population sizes. In T/C studies, survival estimates of control fish have been based on releases of 8,000 to 100,000 tagged fish at Lower Granite Dam, from which only about 1 to 100 return.

The historical extra mortality of in-river fish is inferred from the differences between stock/recruitment data and estimates of passage survival (e.g., $m-M$, \mathbf{a} ; see Section 4.2.2). These patterns in extra mortality can be examined to see how consistent they are with temporal patterns in the hypothesized stressors under each extra mortality hypothesis. This section summarises several such analyses, which are detailed in attached submissions (i.e., Hinrichsen (Submissions 3 and 17), Williams et al. (Submission 1), Budy et al. (Submission 19), Petrosky (Submission 20), and Petrosky et al. (Submission 21)). Before diving into these analyses, however, it is worth noting the cautions made by Deriso with respect to testing extra mortality hypotheses (Submission 24):

1. All three hypotheses pertain to changes in I_n (post-Bonneville survival of non-transported fish) which have occurred during the last 20 years or so. The last Snake River dam (Lower Granite) went into operation in 1975 (brood year 1973); the Regime Shift cycle is assumed to have changed phases beginning in 1977 (brood year 1975); the BKD hypothesis could have begun during those years, earlier, or later. Therefore, all three hypotheses predict a “one-way trip” in the observed spawner-recruit data with I_n generally decreasing in recent years as compared to earlier years. Furthermore the hypotheses all predict nearly the same starting date of a big change in I_n , particularly if the last Snake River dam is disproportionately responsible for decreases in I_n . While analyses involving the specific mechanics about how the above 3 hypotheses have been implemented in BSM are useful to encourage us to examine the hypotheses in more fundamental detail, they dodge the basic problem that we’re dealing with: three alternative hypotheses that are fundamentally similar in their expectations about past patterns of I_n .
2. The delta model contains 86 parameters and many of the parameters have large covariances, which is particularly true for the F parameters that were transformed to obtain estimates of $[1-I_n]$. There is no statistical basis for conducting hypothesis testing by application of “t-tests” on groups of those Maximum Likelihood transformed I_n estimates, nor do I think it is particularly useful given the one-way trip that the alternative models all describe.
3. The I_n estimates are obtained by factoring out several passage model inputs and thus the I_n reflect all errors in those inputs. The errors are potentially quite large, especially during the 1970’s, due to lack of adequate data to support all the passage model inputs – particularly the “D” values. CRiSP and FLUSH produce similar Maximum Likelihood estimates of $[1-I_n]$ when they use the same D values, which indicates that D’s are at least as important as any other passage model input in determining the pattern of $[1-I_n]$. (D is the ratio of post-Bonneville survivals of transported and in-river fish. D affects I_n because I_n is computed from e^{-m} / \mathbf{w} and \mathbf{w} in turn depends on V_n , V_t , P , and D . Essentially m places a limit on the total mortality, and D shifts it around between transported and in-river fish.)
4. There are alternative indices for assessing the regime shift hypothesis, and the lack of an exact environmental feature is one of the difficulties in making a rigorous examination of the hypothesis.

EM1. Hydrosystem Hypothesis

Empirical evidence for and against this hypothesis are presented in Table 4-15. Though the focus of this chapter is on factors affecting the post-Bonneville survival of **in-river fish**, some of the evidence involves estimates of overall patterns of recruitment, which includes both transported and in-river fish.

Table 4-15: EM1. Empirical evidence and arguments in support of the hydrosystem EM hypothesis (a), and evidence/arguments against (b). *Counterpoints are presented in italics.*

a. Evidence/Arguments in support of EM1 <i>(and Counterpoints)</i>	Source
<p>Analyses of covariance show that productivity and survival rate of the upriver spring chinook aggregate remained fairly stable from early hydropower development (1939) until the era of major hydropower development (1970), when major declines began. In contrast, indices of climate change used by Beamish et al. (1997) varied widely from 1939 to 1970.</p> <p><i>Status reviews show major declines in the stocks prior to the 1950's (Mathews and Waples 1991) and multiple factors in addition to the hydro system that have caused declines (Chapman et. al. 1991, Snake River Recovery Team 1995, Litchatowich et.al. 1996, ISAB 1997)</i></p>	Schaller et al. (1996)
<p>Productivity (measured as the Ricker <i>a</i> parameter in the ln(Recruits/Spawner) versus Spawner relationship) declined more from the pre-1970 period to the post-1974 period in seven upstream stocks than in six downstream stocks. These changes are consistent in space and time with development of the hydrosystem.</p>	Schaller et al. (1996)
<p>A survival rate index, measured as ln[(observed R/S) / (predicted R/S)], also declined significantly more from the pre-1970 period to the post-1974 period in upstream than in downstream stocks. These changes are consistent in space and time with development of the hydrosystem.</p>	Schaller et al. (1996)
<p>MLE estimates of <i>F</i>, which include direct and delayed passage mortality components, were correlated with water travel times experienced during the smolt outmigration and were independent of year effects (<i>d</i>).</p> <p><i>Despite the fact that m was correlated with water travel times (WTT) in Deriso et al. (1996), the post-Bonneville mortality factor (1-I_n) shows no relationship with water travel time over BY75-90.</i></p>	Deriso et al. 1996
b. Evidence/Arguments against EM1 <i>(and Counterpoints)</i>	Source
<p>As formulated in the Delta model, EM1 is untestable because a different slope relationship is allowed for each year (BY1975-90). The hypothesis is formulated in both life cycle models so that a general positive relationship exists between <i>I_n</i> and <i>V_n</i> regardless of their true retrospective association (see note 1 below).</p>	Hinrichsen and Paulsen (Submission 3; Figure 2 and associated text)
<p>The relationship between <i>I_n</i> and <i>V_n</i> was examined for both passage models and life cycle models (i.e., 4 combinations). There was no statistically significant relationship during either of the two periods 1952-1990 and 1975-1990 except when the Alpha life-cycle model was used in conjunction with the FLUSH passage model (see note 2 below).</p>	Hinrichsen and Paulsen (Submission 3; Figures 3-8 and associated text)
<p><i>(1-I_n) is only a small component of the hydrosystem extra mortality hypothesis; direct mortality of in-river fish and delayed mortality of transported fish comprise larger components of the hypothesis.</i></p>	Petrosky et al. (Submission 21, pg. 1-2)
<p><i>The lack of correlation between I_n and V_n is not evidence that hydrosystem extra mortality does not exist. Rather, it reflects the fact that estimates of (1-I_n) are confounded by the very high proportion of fish surviving to Bonneville that are transported, so errors in transport model assumptions end up affecting I_n (see note 2 below).</i></p>	Petrosky et al. (Submission 21, pg. 1-2)

c. Evidence/Arguments against EM1 (and Counterpoints)	Source
<p><i>The results of the decision analysis for A3 do not depend on the slope assumptions in the current formulation of the hydrosystem extra mortality hypothesis. A sensitivity analysis to examine an alternative hydro hypothesis for A3, where I_n was selected by water year from the pre-1970 values, produces equal or higher probabilities of survival and recovery under both CRiSP/T4 and FLUSH/T1 in both Delta and Alpha models. For A1/A2, results would be similar to the BKD hypothesis (i.e., more pessimistic).</i></p> <p>{D-values are selected differently under alternative hydro hypothesis, as compared to original hypothesis. (Paulsen)}</p>	<p>Deriso (Appendix H)</p>

Notes on Evidence

- As described in Section 4.2.2, the existing hydrosystem extra mortality hypothesis assumes that for a given water year, the post Bonneville extra mortality factor for non-transported smolts ($1-I_{n,r}$) is proportional to their in-river mortality ($1-V_{n,r}$). For each of the sixteen retrospective years, BSM computes a factor (say $slope_r$) equal to $(1-I_{n,r})/(1-V_{n,r})$, meant to capture the relationship between the post-Bonneville mortality factor and in-river mortality **in that year**. Then, for a future year y , the future post-Bonneville mortality factor ($1-I_{n,y}$) is computed as $(slope_r)(1-V_{n,y})$, where the retrospective year r and the prospective year y are again matched by water year. The value of $slope$ varies considerably over time (Fig. 4-11): during the pre-1970's period of low in-river mortality, there is a higher value for $slope$, but it declines during the 1980's and 1990's as in-river mortality increases.

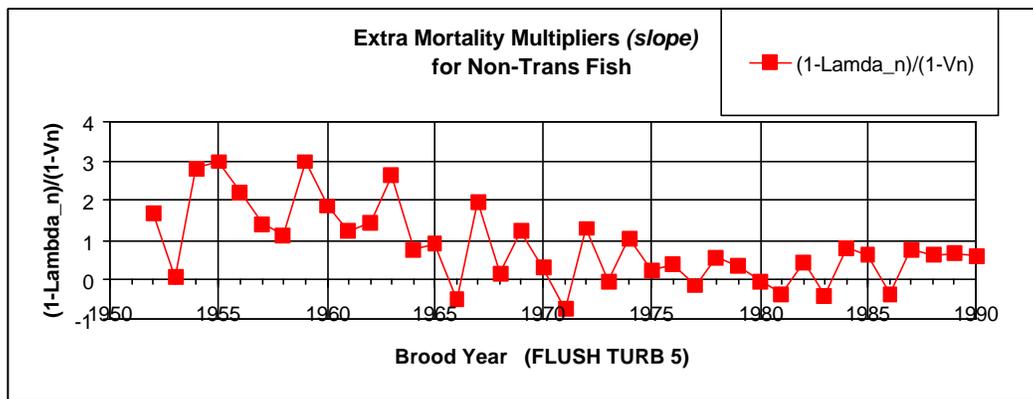


Figure 4-11: Value of $slope = (1-I_{n,r})/(1-V_{n,r})$ over time from the FLUSH model. Only the values from 1975-1990 are used in future simulations.

- Patterns of change in $(1-I_n)$ are shown in Figure 4-12. Both models' estimates of $(1-I_n)$ show general stability prior to the late 1960's, with increased variation subsequently. This variation is the main reason why no relationship is apparent between I_n and V_n , especially for post-1970 years with $V_n < 0.4$ (Hinrichsen, Figures 3 and 5 in Submission 3). I_n is a survival "factor" which can be > 1 (Wilson et al. 1997). However, values of extra mortality $(1-I_n)$ significantly less than zero (i.e., $I_n \gg 1$) could be due to errors in estimating m , V_n , V_t , P , or D , since $I_n = e^{-m} / \omega$, and system survival ω is computed from the last four terms.

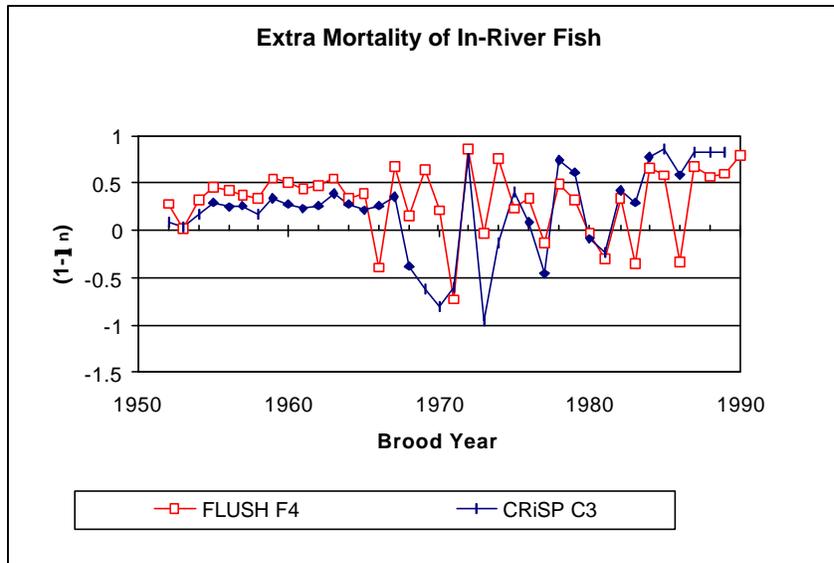


Figure 4-12: Variation in post-Bonneville mortality factor ($I-I_n$) over time under the Delta life cycle model, for FLUSH/T1-TURB5 and CRiSP/T4-TURB4 passage model inputs.

EM2. BKD Hypothesis

The main arguments for and against this hypothesis are summarized in Table 4-14, and in Submission 20. As discussed previously, it makes sense to replace this hypothesis with the much better elaborated hatchery hypothesis EM4. However, since EM4 has not been modeled, the runs completed for EM1 and EM2 provide a starting point from which to assess the probability of survival under EM4 (e.g., a 50:50 allocation of extra mortality of in-river fish to hydro and hatchery causes can be roughly assessed by a 50:50 weighting of EM1 and EM2).

EM3. Regime Shift Hypothesis

There are three components to the regime shift hypothesis:

- 1) Changes in climate and ocean regime between pre-1975 and post-1975 period have affected spring-summer chinook populations in the Columbia River Basin;
- 2) Snake River spring/summer chinook have been affected more severely by regime shift than lower Columbia River populations, and this is primarily responsible for extra mortality associated with post-1975 declines in Snake River spring/summer chinook;
- 3) The ocean regime will improve again in 2005, and remain favourable for another 30 years.

There is little argument about statement 1, though there are disputes about strength of regime shift (see Table 4-14). Statement 1 is consistent with the FY96 Retrospective Analyses (Chapters 3 (Schaller et al. 1996), 5 (Deriso et al. 1996), and 12 (Anderson 1996), and mechanisms are discussed above. We therefore focus attention on evidence for and against components 2 and 3 of the hypothesis.

Table 4-16: EM3. Empirical evidence and arguments in support of the **regime shift** EM hypothesis (a), and evidence/arguments against (b). *Counterpoints are presented in italics.*

a. Empirical Evidence/Arguments in support of EM3 <i>(and Counterpoints)</i>	Source
Under the alpha model, a strong regime shift (STEP close to 1) is estimated with CRiSP, generating a large and significant shift in I_n , coincident with the 1977 regime shift (BY75) recorded by the PDO index.	Hinrichsen (Submission 17)
<i>Not strong evidence. A weak regime shift (near zero STEP) is estimated with FLUSH, with no significant shift in I_n under either the Delta or Alpha model. FLUSH results were found to be more consistent with the spawner-recruit and SAR data in the alpha life cycle model (Tables 4-1 and 4-2).</i>	<i>Hinrichsen (Submission 17); Section 4.1 of this report</i>
<i>Climate correlations with the PDO are driven by D estimates. Both CRiSP and FLUSH passage models produce higher correlations (0.31-0.40) between the ML estimate of $[1-I_n]$ and the PDO index when they use the CRiSP set of D values than the FLUSH set (0.02 to 0.20).</i>	Deriso (Submission 24)
<p><i>The recent (1987-1989) brood years show high ML estimates of extra mortality ($1-I_n$) in all models in contrast to what is indicated to be a favorable climate period by the PDO index. The noisy delta model ML estimates of year-effect track PDO trends from 1969-79, but differ from PDO in BY87-89, and indicate unfavorable climate in two of the three years. No environmental factor shows STEP without noise, and no environmental factor correlates with I_n.</i></p> <p>{Astoria flow has a high frequency variation; STEP is for low frequency change.}</p>	Deriso (Submission 24)
<p>Tree ring records show that longest historical run is 30-35 years. The last change was in 1975. Therefore change expected in next 10 years.</p> <p><i>Some of these references have sampled trees only in S. Oregon.</i></p>	Anderson (Ch. 12); Minobe 1997; Ebbesmeyer (1998); Ware (1995)
b. Evidence/Arguments against EM3 <i>(and Counterpoints)</i>	Source
Component 2) on the previous page requires empirical evidence which shows that Snake River spring/summer chinook stocks are particularly vulnerable to climate variability, unrelated to their hydrosystem experience. No such evidence has been presented to date.	Petrosky et al. (Submission 21)
Recent work on fall chinook has shown no evidence of a regime shift in the Hanford or Lewis stocks (as estimated by residuals from a graph of $\ln(R/S)$ vs. S).	(Petrosky et al., Submission 21).
Sixty year cycle is based on evidence from catch data alone. However, these catch data are of limited value in estimating the duration of such cycles, let alone estimating the level of productivity to which Snake R. chinook will change to in the future. The reason for this limitation is that catch data (for sockeye salmon) are not well correlated with recruitment anomalies (residuals in $\ln(R/S)$).	Peterman
Indices of climate change used by Beamish et al. (1997) varied widely from 1939 to 1970, while stocks were reasonably stable (point a1 from Table 4-15).	

EM4. Hatchery Hypothesis

Empirical evidence is evaluated in Submissions 1, 2 and 19. A sensitivity analysis is included in Section 5.2.

EM5. Multi-factor Hypothesis

Empirical evidence is

Criterion 4) Validity of Method of Projecting Hypothesis into the Future

Issues related to the method of projection were already discussed under Criterion 1) Clarity. Here we compare projections to recent escapement data.

Comparison to Escapement Data

We used the median projected escapements for Johnson Creek under A1 (Current Operations), together with historical escapement estimates, to assess how reasonable the projections are for these two extra mortality hypotheses. A1 was chosen because recent escapements reflect operations very similar to this scenario, therefore providing some grounding in reality. We ran both FLUSH/T1 and CRiSP/T4 with each of the two extra mortality hypotheses, and each of the two life cycle models, so as to determine the relative influence of each factor. These results assume TURB1, low FGE, and low predator removal effects. The historical data and the projections (in five-year intervals) are shown in Figure 4-13. There is less variance in the projections than in the historical pattern since the projections are median values of several thousand runs.

Regime Shift Hypothesis (bottom of Figure 4-13)

With CRiSP/T4, the 1996 median projected escapements for Johnson Creek under the Regime Shift hypothesis are just under 300 spawners, while with FLUSH/T1 they are about 200 spawners (bottom of Figure 4-13). This compares to an average of 150 spawners from 1990 to 1995, and below average levels in 3 of the last 4 years (50, 20, 58 and 236 in 1994 to 1997). Thus both sets of 1996 projections are high relative to observed escapements, but particularly those of CRiSP/T4. For the A1 management scenario, the Regime Shift hypothesis under CRiSP/T4 projects median escapements of 600 spawners by 2010, which continue for the next 25 years. These **median** escapements are as high as any **single year's** escapement from the 1950's, except for 1957 and 1960. Thus CRiSP/T4 projections suggest that during the next positive climate regime, current operations with eight dams (A1) will yield escapements considerably higher than observed in the last positive regime of the late 1950's and early 1960's when there were only 3-4 dams. One possible explanation for this apparently optimistic projection is that mainstem harvest rates were 40-50% in the 1950's and 1960's, and only reach a median value of 15% over the period from 2010 to 2035. FLUSH/T1 shows little response to the regime shift hypothesis, and appears similar to the BKD hypothesis (Fig. 4-4) for reasons previously discussed (Section 4.2.2).

The large increases in escapement under CRiSP/T4 and the regime shift hypothesis (Fig. 4-13) explains why these assumptions increase the chances of A1/A2 having a higher probability of meeting the 24-year survival standard than A3 (Section 3.3.1). Though A1/A2 was only preferred over A3 in 106 out of 960 runs of CRiSP/T4, a disproportionate fraction of these runs had the regime shift/cyclical extra mortality hypothesis (Table 3-3). The 24-year standard covers the period from brood years 1996 to 2019, so the large jumps after 2005 exert a major influence on the probability of remaining above the survival level of 150 spawners.

Hydrosystem Hypothesis (top of Fig. 4-13)

Both models have higher 1996 projected median escapements with the hydrosystem hypothesis than with the regime shift hypothesis. Thus the above described concerns about overestimates of initial conditions apply even more strongly here. CRiSP projections are again higher than those of FLUSH. The median projected escapements for CRiSP are higher than any observed escapements since BY 1973, while those of FLUSH are higher than most of these escapements but still within the historical range.

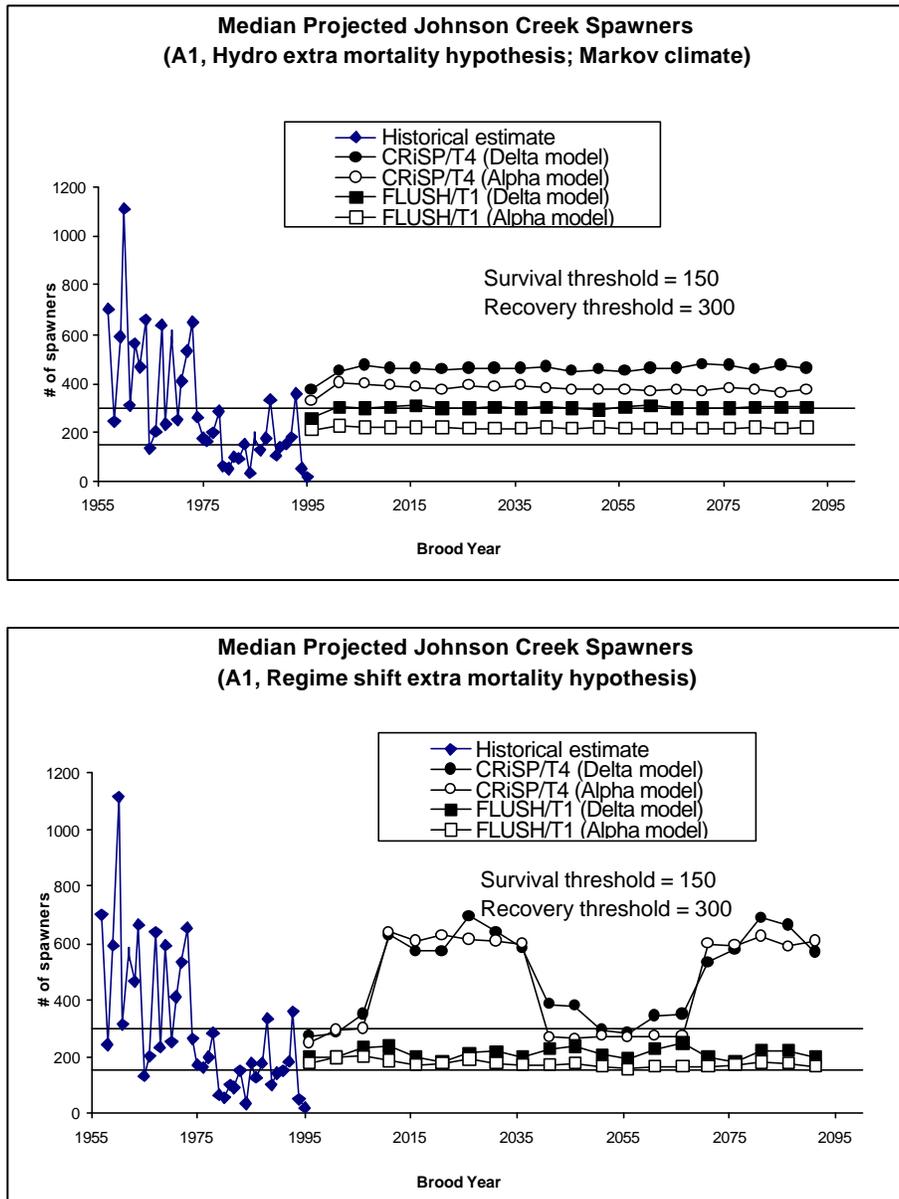


Figure 4-13: Projected escapements for Johnson Creek under hydrosystem-related and regime shift extra mortality hypotheses, relative to historical data.

4.3 Evaluation of Drawdown Hypotheses

Unlike the other key hypotheses, there is no direct historical element of the drawdown hypotheses because they apply to a set of actions and responses that may occur in the future. Although some pre-dam data exists, some would argue that a post-drawdown Snake River will not necessarily revert to the pre-dam condition because impoundment has caused permanent changes in the river's physical and biological characteristics.

Therefore, PATH has focussed its efforts on forming reasonable hypotheses about the duration of various time periods before, during, and after the actual removal of dams, and the juvenile and adult survival rates that are expected to occur during those time periods. Of these, we considered multiple hypotheses about the equilibrated juvenile survival rate and the duration and juvenile survival rate during the transition period. The alternative hypotheses for these values are intended merely to bracket the range of responses we expect to see following drawdown.

Evidence for these alternative hypotheses is limited, and in some cases we have been required to form reasonable guesses based on first principles. The current modeling work by Batelle on physical changes following drawdown should help to focus the current set of hypotheses.

4.3.1 Equilibrated Juvenile Survival Rate

This is defined as the survival rate of smolts through the new free-flowing reach (i.e., from above LGR to below IHR) after the river system has reached some equilibrium state. The PATH drawdown group developed two alternative hypotheses.

Hypothesis 1 (EJUV 1)

Although survival rates will increase following drawdown, they will not necessarily return to the same level as they were prior to construction of the Snake River dams because there are other permanent changes that have occurred since then that will continue to reduce smolt survival rates. These other changes include shoreline development, changes in upstream water regulation, and changes in ecological communities. The hypothesized survival rate through the free-flowing reach under this hypothesis is 0.86, based on recent reach survival studies in free-flowing reaches above Lower Granite pool.

Hypothesis 2 (EJUV 2)

Survival rates will increase to the same level as they were prior to the construction of the dams. Implicit in this hypothesis is the assumption that predator densities in the free-flowing river following drawdown will return to historical levels, and that other changes since 1968 have had negligible effects on Snake River spring/summer chinook. The hypothesized equilibrium survival rate under this hypothesis is 0.96, based on historical reach survival estimates prior to construction of 3 of the 4 Snake River dams.

Criterion 1) Clarity

Conceptually, both hypotheses are relatively clear and uncomplicated. One hypothesis says that survival rates will return to pre-dam levels, while the other does not. However, both hypotheses make some assumptions about what pre-dam survival rates were and what other changes in the river have occurred since then.

Criterion 2) Mechanism

The implied mechanism under EJUV1 is that there have been permanent changes in the river since the construction of the dams, and that these changes will continue to limit juvenile survival rates even after the river is returned to a free-flowing state. These changes include shoreline development, changes in upstream water regulation, and changes in ecological communities. The implied mechanism under EJUV2 is that all aspects of the river that affect juvenile survival rates will revert completely to pre-dam conditions.

These two hypotheses were intended to bracket the range of possible equilibrated juvenile survival rates, and are therefore somewhat extreme. Shoreline development and changes in ecological communities could reasonably be hypothesized to at least partially revert to pre-dam conditions as the free-flowing river forms a new channel within current reservoir shorelines and becomes less hospitable to species adapted to lacustrine conditions. The question is, to what extent?

Those who would support equilibrated survival rates at the lower end of this range need to provide evidence to show that these factors have had a significant effect on juvenile survival rates and that they will not be affected by returning the river to free-flowing conditions. Similarly, those who would support the upper end of this range need to provide evidence that aspects of the river that have changed since the dams were constructed either have no effect on juvenile survival rates or will change back to their pre-dam state.

Criterion 3) Consistency with empirical evidence

Hypothesized equilibrated juvenile survival rates are based on pre-dam survival estimates or on recent estimates of survival through free-flowing stretches of the Snake River. The lower estimate is the average survival rate of wild juveniles in free-flowing reaches between Whitebird and Imnaha traps and Lower Granite Dam from 1993-1996, expanded on a per-km basis to encompass the 210-km length of the free-flowing reach under drawdown (Table 4-7). The higher estimate is the average survival rate of smolts between Whitebird trap in the Salmon River and Ice Harbor Dam from 1966 to 1968, prior to completion of the other three Snake River dams (Table 4-7).

(i) Applicability: Score = 3

The question of whether the respective time periods of these two sets of reach survival estimates is applicable is central to the alternative hypotheses. If the changes that have occurred since the dams were built do indeed affect salmon survival, and will continue to do so, then estimates from the recent time period are the appropriate comparison. If survival rates will return to pre-dam states, then the pre-dam estimates are more applicable.

(ii) Clarity: Score = 2

Neither set of reach survival estimates are completely applicable to the free-flowing reach that would be created by drawing down Snake River dams (i.e., from above LGR pool to below IHR). Pre-dam estimates are from Whitebird trap in the Salmon River to Ice Harbor Dam. Recent estimates are from Whitebird or Imnaha traps to LGR tailrace. Both of these reaches overlap with the free-flowing reach under drawdown, but do not cover it completely. Difference between these two estimates could also be partially due to changes in methods, but we haven't checked this.

Table 4-17: Pre-dam and current estimates of survival of wild spring/summer chinook smolts in free-flowing reaches of the Snake River

Application of the reach survival estimates to free-flow survival is somewhat complicated by the inclusion of one dam in the study reach. Pre-dam estimates are to below Ice Harbor dam. Recent estimates are to below Lower Granite dam. Therefore, to determine a free-flowing survival rate requires an adjustment for the assumed project mortality in that time period. This has not been considered a major problem because the assumed project survival in each case is high (0.95 for Ice Harbor, 0.92 for Lower Granite). As a sensitivity analysis, changing LGR mortality to 0.9 and 0.94 (instead of 0.92) shifts estimated 210 km reach survivals in the 1996 Whitebird Trap – LGR tailrace study from 0.88 to 0.90 and 0.86 respectively. Without knowing the variances of reach survival estimates, it is not possible to assess the potential error in the 210 km estimates.

- (iii) Rigor: Score = 1
See earlier discussion of reach survival estimates.

Criterion 4) Validity of method of projection

Operationally, the implementation of these hypotheses is uncomplicated. Under drawdown scenarios, both passage models fix the juvenile survival rate through the newly-formed free-flowing stretch of river to one of the two hypothesized values (0.85 or 0.96). In FLUSH, the hypothesized survival rates are hard-wired, while CRiSP fixes the hypothesized survival rate by adjusting its predator densities through that portion of the river. Survival rates below the free-flowing reach are modeled in the usual way and are dependent on flows.

This approach has been criticized because the survival rates (in the drawdown reach only) are assumed to be constant in the future rather than varying from year to year depending on flows. Introducing some variability into the equilibrated juvenile survival rate (either random variation or based on flows) seems like a logical approach, if variability in these rates is expected to change the results of the decision analysis.

4.3.2 Transition Period – Duration and Juvenile Survival Rate

The transition period is defined as the period of time between the end of the construction period and when the free-flowing river attains some equilibrium survival rate for juveniles. Physical processes during this period include increased water velocities (reduced travel times), formation of a new channel, washing out of accumulated sediments, stabilization of banks, and re-establishment of riparian areas beside the new channel. Possible biological processes include changes in ecological communities. With respect to juvenile survival rates during the transition period, the primary biological consideration is what changes will occur in the density, abundance, activity, and distribution of predator species in the free-flowing river.

Currently, there is very little data on physical and biological effects of drawdown during the transition period. Although work on delineating these effects of drawdown are underway by the U.S. Geological Survey – Biological Resources Division and by Batelle laboratories, results are not expected until late 1998. In the absence of this information., the PATH drawdown workgroup identified two alternative scenarios for the duration of the transition period. These hypotheses were intended to bracket the range of possible effects for the sensitivity analysis based on some preliminary assessments of these effects.

Hypothesis 1 (TJUVa)

The transition period will be two years, with a linear increase in in-river survival rates current levels to the hypothesized equilibrium levels. This hypothesis represents a lower bound, and assumes a quick response of juvenile survival rates to short-term physical (scouring of main channel) and biological (redistribution of predators) effects.

Hypothesis 2 (TJUVb)

The transition period will be 10 years, with a linear increase in in-river survival rates from current levels to the hypothesized equilibrium levels. This hypothesis represents an upper bound on the duration of the transition period, based on long-term physical and biological effects.

Criterion 1) Clarity

Both hypotheses are intended to provide a simple means of summarizing the aggregate effects of a range of physical and biological processes. This approach is a reasonable first step given the amount of information currently available. A more mechanistic approach to modeling these effects was recommended by the PATH drawdown workgroup. However, because of time constraints this approach will not provide any information in time for the PATH final report in October 1998.

Criterion 2) Mechanism

The response of juvenile survival rates during the transition period is thought to be primarily a function of two processes:

- a) the response of predator populations to the change from reservoir to free-flowing conditions. Specifically:
 - lower water volumes may reduce predator carrying capacity (although initial increases in density are possible)
 - increased turbidity and decreased temperature may reduce consumption rate,
 - changes in channel morphology and microhabitat distribution may affect distribution of predators and juvenile chinook, which would affect encounter rates.
- b) Increased water velocities will reduce fish travel times, reducing the exposure of juvenile chinook to predation

The increase in water velocities under drawdown is generally accepted. The key question, therefore, is whether the effects of changes in predator populations will be sufficiently negative to counteract the positive effects of reduced travel times. A very limited amount of information is available on predator densities and predation rates in free-flowing sections of the Snake River (upstream of Lower Granite dam) and the Columbia River (below Bonneville dam). At both sites, predator densities and consumption rates were found to be higher than in mid-reservoir samples, but the applicability of these data to a free-flowing Snake River is tenuous, and the “data for making broad conclusions are sparse” (review by Petersen and Poe, 1998 proposal). Work is currently underway to study the effects of habitat change on predator densities and consumption rates.

In the absence of empirical evidence, we have hypothesized that juvenile survival rates will increase linearly during the transition period in both of these hypotheses. Since the Preliminary Decision Analysis report was published, Jim Anderson has proposed a new hypothesis in which there is an initial increase in

mortality because the increase in predator density is not fully compensated for by the reduction in travel times. This hypothesis is described more fully in Section 5 and in Submission 12.

Criterion 3) Consistency with empirical evidence

Empirical evidence regarding juvenile survival rates during the transition period is scarce. Therefore, transition period hypotheses were based largely on some preliminary assessments of relevant physical and biological processes during the transition period by the Corps of Engineers and PATH scientists.

The two-year estimate was based on three pieces of information:

- a) initial estimates by the Corps are that sediments in the main channel will be scoured out within 2 years of reservoir drawdown;
- b) redistribution of predators could take place relatively quickly;
- c) response to removal of dam mortality should be immediate

The 10-year hypothesis was based on the following preliminary estimates:

- a) sediments along the banks will be scoured out approximately 10 years after drawdown;
- b) banks will stabilize and be recolonized by vegetation approximately 10 years after drawdown
- c) based on time to maturity, population responses of predator species will take place at least 5 years after drawdown.

Because these are preliminary assessments rather than empirical evidence, we do not apply the “evidence criteria” (i.e., applicability, clarity, rigor). However, PATH participants who submit empirical evidence related to these hypotheses should do so with these criteria in mind.

We have recently received more detailed preliminary estimates of timing for various changes to occur after drawdown (Submission 13). In general, initial changes in channel and substrate are expected to become observable between 2 and 10 years, with more complex channel morphology reinitiating over longer time periods (20-50 years). Re-establishment of early successional riparian communities is expected to initiate within 10 years, with more diverse riparian communities developing over 3-20 years.

Criterion 4) Validity of method of projection

Hypotheses about the duration of the transition period are implemented within the structure of the life-cycle models. For ease of computation, in-river survival rates during the transition period are increased in a step-wise manner that approximates a linear increase from current levels to hypothesized equilibrium levels. New hypotheses about transition survival (Submission 12) would be implemented in a similar step-wise fashion. Although step-wise changes in survival only allow coarse projections of trajectories of juvenile survival rates over time, coarse projections are all that are warranted given the scarcity of empirical evidence.

5.0 Summary of New Hypotheses

A number of new hypotheses have been proposed since the Preliminary Decision Analysis Report was published in March. The text descriptions of these hypotheses that we have received are compiled in the submissions to this report. The purpose of this section is to:

1. provide a brief summary of the hypotheses with emphasis on what is different from existing hypotheses;
2. discuss the potential for affecting the results, in terms of both the ranking of actions and the ability of an action to meet the jeopardy standards. Given the limited amount of time, we will need to carefully consider whether these new hypotheses affect the results to minimize the amount of modeling required;
3. summarize the evidence listed in the summaries of the new hypotheses that were submitted to ESSA; and
4. resolve issues. We raise these as items for PATH participants to consider while compiling additional information for or against these hypotheses.

List of New Hypotheses (Submission number)

1. Multi-factor extra mortality hypothesis (Submission 11)
2. Hatchery extra mortality hypothesis (Submissions 1 and 2)
3. Reformulation of the Hydro extra mortality hypothesis (Submission 3)
4. Juvenile survival rate during the transition period (Submission 12)
5. Adult survival rate under drawdown (Submission 8)
6. Juvenile survival rate at equilibrium (Submission 8)
7. Productivity as a function of marine nutrients (Submission 8)
8. Additional habitat hypothesis (Submission 4 and 6)

As the first three hypotheses have already been examined in Section 4.3.3, these are only briefly discussed here.

5.1 Multi-factor Extra Mortality Hypothesis

5.1.1 Description

In this aggregate hypothesis a number of factors are hypothesized to affect extra mortality of Snake River chinook salmon. These factors include climatic effects, avian predators, hatchery interactions, and various hydrosystem effects. This hypothesis was discussed in more detail in Section 4.3.3.

5.1.2 Effects on Results

Given that existing extra mortality hypotheses are intended to bound the range of responses, it is likely that the results of a multi-factor hypothesis would be intermediate to those of existing extra mortality hypotheses.

5.1.3 Summary of Evidence

Between 1968 and 1985, when the hydropower system was either nearly complete or complete, SARs were equal or nearly equal to historic values in 7 of the 18 years (Submission 9, Figure 5). This suggests that factors than the hydrosystem may be important.

Populations of Caspian Terns and Double Breasted Cormorants in the lower Columbia estuary have increased since 1987. Evidence from feeding studies, bioenergetic calculations, and PIT-tag samples from the vicinity of major Caspian tern colonies support the hypothesis of substantial impacts by these predators on juvenile salmon (Submission 16, p. 9-11).

Alpha term in the Alpha model goes very negative in 1989, 1990, 1991 – coincident with increase in bird populations.

Counterpoints:

There is no decline in estimated delta during this period – what hypothesized mechanisms would affect Snake River fish differentially beyond common year effects?

5.1.4 Issues to Resolve

Key issues to be resolved are:

- whether a multi-factor hypothesis can be implemented. That is, can it have meaningful ranges on the parameter estimates given the colinearity in the independent variables. A key question is the extent to which the various factors interact (e.g., Is the abundance of birds actually a hatchery effect? Is flow at Astoria really a hydrosystem effect?).
- many multi-factor hypotheses are possible (e.g., effects of marine mammals, shad): will a multi-factor hypothesis affect the ranking of actions?

5.2 Hatchery Extra Mortality Hypothesis

This hypothesis was summarized in Section 4.3.3. The evidence presented in Submissions 1 and 2 is worth close scrutiny. It shows interesting correlations between extra mortality and numbers of hatchery smolts. The detrended correlations (Submission 2) are however very low, suggesting that other factors are also important in determining year to year changes in extra mortality. In addition, Paulsen (Submission 2) suggests that if stresses imposed by hatchery fish on transported wild fish were an important mechanism, then the correlation between extra mortality of transported fish and hatchery releases should be stronger than the correlation between extra mortality of in-river fish and releases. However, no consistent patterns of this sort are observed.

The conclusions of Williams et al. (Submission 1) suggest that the impacts of hatcheries are unlikely to change the ranking of actions. That is, the impacts of hatcheries would on balance be less under A3 than under A1/A2. The key question to be addressed by the group therefore is to what extent hatchery

activities are part of existing extra mortality, and are unlikely to change from hydrosystem actions alone. This affects the ability to meet standards.

It is important to note that effects of hatchery interactions are already implicitly incorporated into the prospective models. For transported fish, hatchery effects are captured in D values, while effects on in-river fish are implicit in the hydro extra mortality hypothesis. However, one could explicitly consider hatchery effects in terms of the weighting of BKD results and hydrosystem-related results. In the absence of changes in hatchery actions, one would expect historic levels of extra mortality that are due to hatchery effects to remain in the future, which is what the BKD extra mortality hypothesis assumes. For example, if 20% of the extra mortality is assumed to be here to stay without changes in hatcheries, then apply weights of 0.2 and 0.8 to these two hypotheses. However, such a weighting must be based on some rigorous representation of the impacts of hatchery fish. The next step would logically be to include total hatchery smolts in the MLE together with other variables such as WTT.

We explored a simple approach to modeling hatchery effects by looking at the effects of placing different weights on the BKD extra mortality hypothesis for a given combination of passage/transportation model, TURB assumptions, life-cycle model, and future climate hypotheses (FGE, predator removal, equilibrated juvenile survival rates, and transition period are weighted equally in these results). We looked at combinations of:

All weight placed on CRiSP/T4, TURB4, Alpha model, cyclical climate X varying weights on BKD and regime shift extra mortality; and

All weight placed on FLUSH/T1, TURB5, Delta model, Markov climate X varying weights on BKD and Hydro extra mortality.

In essence, these combinations of passage/transportation models, life-cycle models, and the BKD extra mortality hypothesis represent alternative aggregate hypotheses (see Section 4.1) These results are only for these specific combinations of passage/transportation models, TURB assumptions; other combinations are possible.

Results are shown in Figure 5-1 for the 24-year survival standard. Since this standard is hardest to meet, the ability to achieve this standard is a good approximation of the ability to meet all of the standards. Placing more weights on the BKD extra mortality results in lower jeopardy probabilities in all cases, and generally leads to a greater difference between A1/A2 and A3. Although the BKD extra mortality does not affect the ranking of actions, it does affect the ability of actions to meet all of the standards. With CRiSP/T4, TURB4, and Alpha model assumptions, the weight placed on the BKD hypothesis relative to the regime shift hypothesis must be close to 0 for A1 and A2 to achieve all of the standards. For A3, the weight on BKD must be 0.5 or less for a 3-year delay, and less than 0.2 for the 8-year delay. In the context of evaluating hatchery hypotheses, this implies that A1 and A2 will not meet all of the standards as long as the proportion of the extra mortality is ascribed to hatchery effects is greater than around 0.1. A3 will not meet the standards if 50% (or 20%, depending on the pre-removal period assumptions) or more of the extra mortality is ascribed to hatchery effects.

With FLUSH/T1, TURB5, and Delta model assumptions, A1 or A2 will not meet the standards as long as at least 10-20% of the extra mortality is assumed to be due to hatchery effects. A3 with an eight-year delay will not meet the all of the jeopardy standards if hatchery effects are assumed to account for 50% or more of the extra mortality. A3 with a three-year delay will meet all of the standards regardless of the hypothesized contribution of hatchery effects to extra mortality.

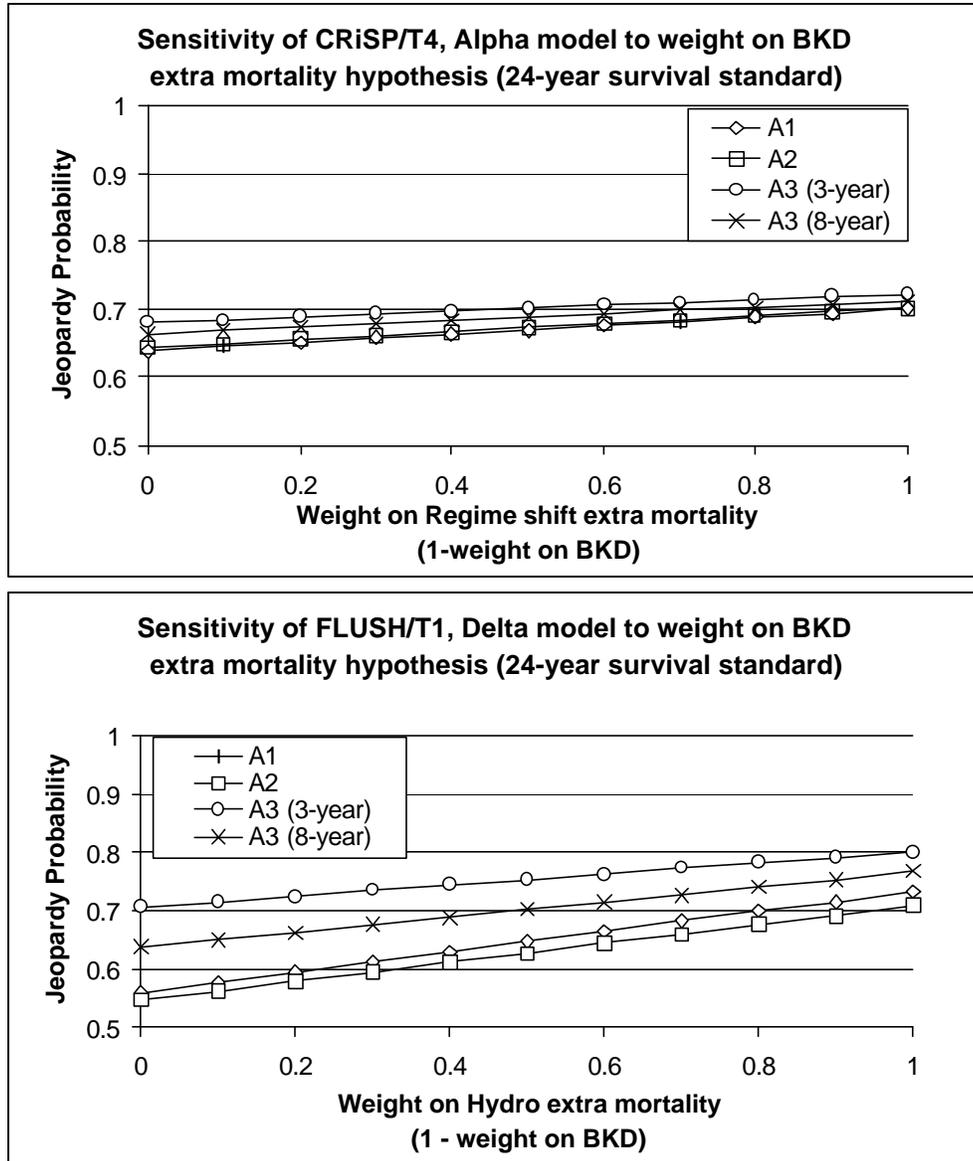


Figure 5-1: Sensitivity of the 24-year survival probability to weight placed on the BKD hypothesis.

5.3 Reformulation of the Hydro Extra Mortality Hypothesis

Ideas related to this issue are discussed in Section 4.3.3, and in Submission 3 from Hinrichsen and Paulsen. A sensitivity analysis to an alternative formulation of the hydro system is provided in Appendix H.

5.4 Juvenile Survival Rate During the Transition Period

5.4.1 Description

This hypothesis proposes an initial reduction in reservoir survival of smolts immediately following drawdown because of concentration of predators into a reduced water volume (Submission 12). The major difference in this hypothesis from the other transition period hypotheses is the assumption that the benefits of reduced travel time will not be sufficient to overcome an initial increase in predation mortality as the present abundance of predators are compacted into a smaller volume of water.

5.4.2 Effects on Results

The new hypothesis would predict lower jeopardy probabilities under A3 than the current set of transition hypotheses, because both of the current alternatives assume a linear increase in juvenile survival rates following drawdown. The new hypothesis would only be significant when coupled with the 10-year transition period hypothesis, because a 2-year transition period would result in only a short period of time where survival is reduced.

The largest effect that this could have on the results is on the margin of difference between 24-year survival probabilities for A1, A2, and A3 with CRiSP/T4 (virtually all of the cases where A1 or A2 was greater than or equal A3 were with the longer transition period) (Table 3-4). The magnitude of this effect will depend on a) the amount of the reduction in survival immediately following survival, and b) how long the reduced survival conditions will remain in effect. Given the small number of cases in which A1 or A2 performed better than A3 under CRiSP and the 24-year survival probability (106 out of 960), this effect would have to be very strong to substantially influence the results.

We did a limited sensitivity analysis to explore the effects of a reduction in juvenile survival following drawdown. We considered two scenarios: one in which overall in-river survival (V_n) decreases by 10% in the first 5 years after drawdown, and one where V_n decreases by 50% (Figure 5-2). The 10% and 50% figures were arbitrary values chosen to explore a range of possible population responses. Note that the actual decrease in reservoir survival is larger than this, because the decrease in V_n includes elimination of the dam mortality associated with the dams removed under drawdown. We have only done this analysis with CRiSP runs so far. The runs assume best-case passage assumptions (i.e., FGE1, TURB4, PREM3) and worst-case drawdown assumptions (8-year pre-removal, 10-year transition, and low equilibrated juvenile survival rate).

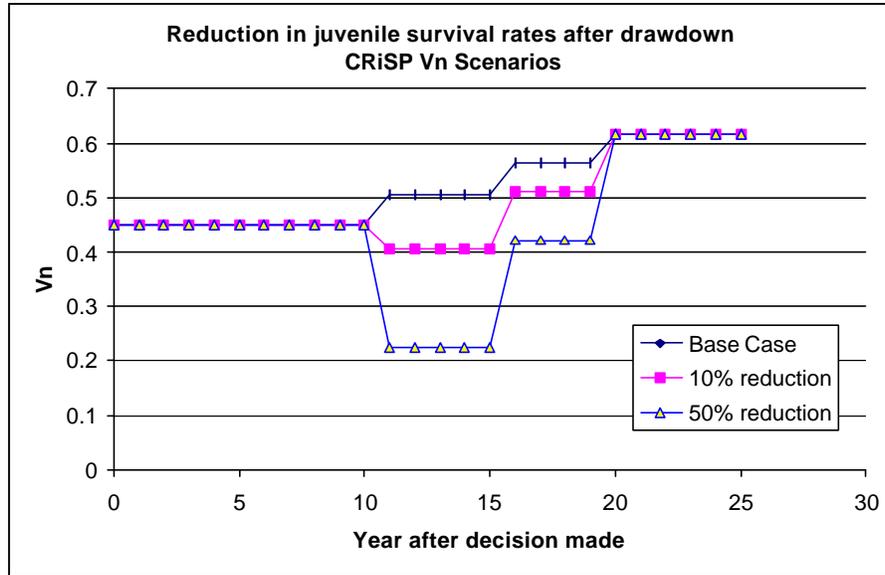


Figure 5-2: Trajectory of juvenile survival rates under different scenarios of survival rate response following drawdown.

Results for Marsh Creek stock are shown in Figure 5-3. Twenty-four-year survival standards are most sensitive, but the effects of the modeled decreases in survival are small. With the delta model, the 24-year survival probability declines by about 0.005 with a 10% reduction in survival, and by about 0.04 with a 50% reduction in survival. With the alpha model, the 24-year survival probability declines by about 0.01 and 0.03 for the 10% and 50% reduction in survival. These results are preliminary and need to be verified; we are not sure why reductions in survival would lead to **increases** in the 48-year recovery probability under the Alpha model. It may be that the small effects of the reductions in transition survival are masked by the small amount of numerical error introduced into the model results by stochastic processes.

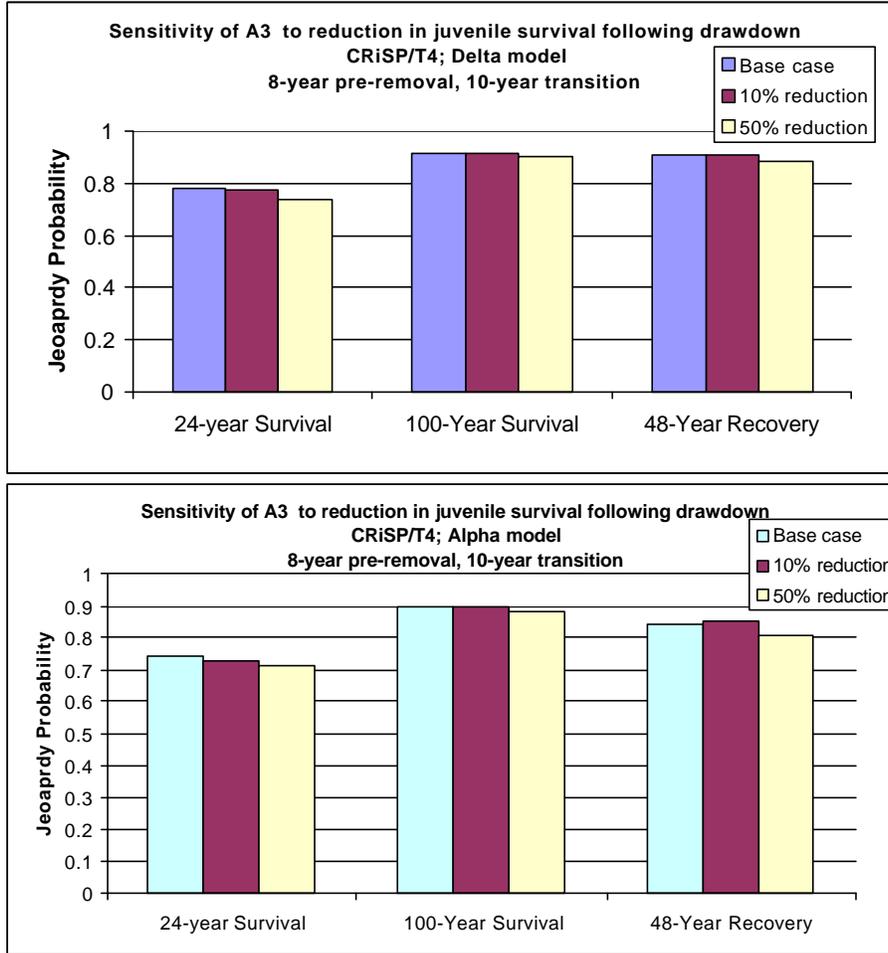


Figure 5-3: Sensitivity of 24-year survival probabilities for Marsh Creek to alternative scenarios of survival rate response following drawdown.

5.4.3 Summary of Evidence

The basis of this hypothesis is a mathematical argument that survival of smolts under drawdown will decrease given no change in predator abundance. As noted earlier in Section 4.3.2, there is no empirical evidence to say whether predator densities will increase or decrease.

5.4.4 Issues to Resolve

The direction and magnitude of impact depends on whether possible declines in predator activity (due to increased turbidity and changes from a reservoir to riverine environment) are sufficient to counteract the decline in cross-sectional area of the channel (Submission 12, p. 3).

5.5 Adult Survival Rate Under Drawdown

5.5.1 Description

This hypothesis says that adult migration survival through the drawdown reach will not be significantly different than survival through the existing dam and reservoir complex because adults migrate faster (and consequently have higher survival rates) through the hydrosystem than they do in free-flowing stretches of the river. Current simulations of drawdown scenarios do not consider multiple hypotheses for effects of drawdown on adult survival rates – all simulations assume that adult survival rates through the drawdown reach will increase to 97% (see Section 5.5.3 for description of how this estimate was derived).

5.5.2 Effects on Results

Since we have not done any sensitivity analyses of adult survival rates under drawdown, we can't say whether the hypothesis will affect the results. Proponents of this hypothesis need to supply results of such sensitivity analyses so we can assess whether the effects warrant simulation of this hypothesis. Given that A3 was preferred under the vast majority of cases, we suspect that the effect would likely have to be quite strong to change the results significantly.

5.5.3 Summary of Evidence

Four pieces of evidence are cited in support of this hypothesis:

1. Estimates used to derive the 97% adult survival rate used in the current analyses are not valid. The current estimate of 97% was based on Bjornn's estimate of average changes in adult survival to spawning grounds following completion of the three projects above Ice Harbor. The criticism of this approach is that there is some question as to whether or not the sample size in the pre-dam estimates was large enough to allow comparison to post-dam survival rates.⁴
2. Bjornn (1998; cited in Jim Geiselman's summary – Submission 8) used radio tags to estimate adult survival rates from Ice Harbor dam to spawning grounds and hatcheries. The survival rate through the free-flowing portion of this reach was similar to survival through the hydrosystem portion (Table 5-1). Tests of significance were not reported in the summary of this hypothesis that was submitted to ESSA.

Table 5-1: Adult survival rates for 1991-1993 using radio tags

Year	Overall survival from IHR – spawning/hatchery	Survival through hydrosystem portion	Survival through free-flowing portion (1)
1991	54%	74%	73%
1992	61%	77%	79%
1993	76%	84%	90%

⁴ J. Geiselman: That estimated increase is based on information presented by Bjornn (1990). In that paper a general comparison in adult survival indices for two periods was presented. From 1962-1968 Bjornn (1990) estimated that on average 55% of the wild spring/summer chinook passing Ice Harbor Dam survived to spawn, whereas the mean for the years 1975-1988 with all four dams in place averaged 46%. However, a recent report by Bjornn et al. 1998 questions whether that difference is really significant; referring to the early paper (Bjornn 1990) they caution, "we do not know if the dams and reservoirs have increased the rate of mortality because few survival estimates are available (Bjornn 1990) before all four of the Snake River dams were completed." Inspection of Figure 1 in Bjornn (1990) indicates that the variability around the mean values for 1962-90 was pronounced, with survivals ranging from 30 to 70 percent, supporting the Bjornn et al. (1998) cautionary note in interpreting those mean values.

- (1) *The majority of tagged fish were hatchery fish. The geometric mean for these three estimates is about 0.8, which is about what one would expect for hatchery fish. These estimates are likely to be a conservative estimate because fish that go to minor tributaries and regurgitated tags are counted as mortality. Charlie Petrosky, oral comment.*
3. Bjornn (1998) found that fish migrated faster through the hydrosystem than through free-flowing reaches.
 4. Bjornn (1998) also found a statistically significant negative relationship between travel time and survival in two of the three years.
 5. Paulsen (1998; cited in Jim Geiselman's summary – haven't seen this report) found negative relationships between flow and adult conversion rates.

5.5.4 Issues to Resolve

The major difficulty with implementing this hypothesis is that all of the run reconstructions and retrospective and prospective model runs would have to be redone to reflect the hypothesized lower conversion rates prior to completion of the dams. This is a task that we clearly do not have time to complete. In any case, revising the conversion rates in this way would only increase the difference between upstream and downstream stocks (Howard S., oral comment).

A critical issue to resolve is the sensitivity of jeopardy probabilities to adult survival rates. As Jim Geiselman suggested, we should also assess the significance of differences in pre-dam and post-dam average survival rates on which the current value of 97% is based.

5.6 Juvenile Survival Rate at Equilibrium

5.6.1 Description

This is not really an alternative hypothesis, just a suggestion to allow some variability in the equilibrated juvenile survival rates under drawdown. Variation could be based on some flow-survival relationship or simply random if such a relationship does not exist. Currently, equilibrated survival rates are fixed at their hypothesized values (either 0.86 or 0.95) for the duration of the equilibrium period.

5.6.2 Effects on Results

Variable equilibrated juvenile survival rates will have minimal effect on overall results. The frequency with which water years are selected is fairly uniform (Fig. 3-27), which suggests that particularly high or low flow years have only a small influence on overall results. If variability is random, over many simulations the variable juvenile survival rates should average out to the current values.

5.6.3 Summary of Evidence

The summary of the hypothesis cites various sources of evidence to support a (weak) relationship between flow and survival. The key evidence that is missing is data that would suggest that overall results are particularly sensitive to variability in this parameter.

5.6.4 Issues to Resolve

Given the limited time, we should not incorporate more variability into the model unless it is important to do so. Assuming that equilibrated survival rates are constant is not the most realistic approach, but it may be the most feasible if assuming a constant value has only small effects on results.

5.7 Productivity as a function of marine nutrients

5.7.1 Description

The hypothesis is that the productivity of Snake River salmon stock is a function of marine nutrients delivered to freshwater rearing habitats via salmon carcasses. Although it is not clear from the summary of this hypothesis exactly how this hypothesis would be implemented, we assume that some explicit relationship between productivity and spawner abundance is being proposed. Currently, there is no explicit consideration of this process, although it is implied in the stock-recruitment function both in the form of the Ricker function and the depensation parameter p (equation 4.2.2-1)..

5.7.2 Effects on Results

Deriso (1997; in FY97 Report) explored the sensitivity of jeopardy probabilities to an explicit depensation function in which the spawner-to-recruit survival rate declined at a greater rate below the minimum observed spawning abundances. His results showed that implementation of this depensation function had minor effects on jeopardy probabilities (Table 9 in Deriso 1997).

The marine nutrient hypothesis would probably have a greater effect on the ability of actions to meet the standards than on the ranking of actions, because the process would occur in all three management alternatives. However, given that there is generally a greater number of spawners projected under drawdown, the effects of low spawner numbers may be greater under A1 and A2.

5.7.3 Summary of Evidence

There are a number of reports cited in the hypothesis summary in support of the importance of marine nutrients delivered by salmon carcasses to the freshwater environment in salmon-bearing streams. However, such effects are generally not detected in spawner-recruit data. The stock-recruitment function we are using for life-cycle modeling contains an explicit depensation parameter. The empirical estimate of this parameter from the spawner-recruit data is very close to 0, which suggests that depensation is not apparent in the data. In a separate analysis, Botsford (1997) found a weak indication of depensation at low spawner values. A review of 128 harvested fish species by Myers et al. (1995) found evidence of depensation in only three of them (two of these were in Pink salmon). It is also important to recognize that the Ricker stock-recruitment function that we are using already explicitly models spawner to recruit survival as a function of spawner abundance.

The bottom line is that although there are experimental data that show the contribution of marine-derived nutrients to freshwater habitats, this and other depensatory mechanisms are difficult to detect in spawner-recruit data. This is an important point because the life-cycle simulations are all based on the spawner-recruit data set, and it would be difficult to impose another productivity effect into the life-cycle simulations.

5.7.4 Issues to Resolve

We need to know exactly what this hypothesis is in quantitative terms so that we can assess:

- a) whether the proposed mechanism is already captured in empirical estimates of stock-recruitment parameters; and if not;
- b) how feasible it is to implement the proposed approach within the current stock-recruitment framework

5.8 Additional Habitat Hypothesis

5.8.1 Description

The submission by Charlie Paulsen on habitat effects is in two parts. The first part describes the current implementation of these effects in the life-cycle model and how problems with the implementation caused model results under habitat enhancement to be worse than under current conditions. This was the reason why the results in Section 3 of this report do not include any habitat effects.

The second part of the submission describes an alternative habitat hypothesis based on an observed association between overwintering survival and indices of land use/vegetation and road density. It is suggested that the increases in survival associated with particular habitat clusters could be implemented by changes in Ricker a values based on the statistical models developed. This is a more specific method for modeling habitat effects than that used in the current analysis, where the habitat workgroup used expert judgement to make some associations between general habitat management actions and effects on Ricker a values.

5.8.2 Effects on Results

Any effects on survival in the freshwater rearing life stage will be independent of hydro actions, and therefore will not affect the relative ranking of the alternative hydro actions (A1, A2, and A3). However, habitat effects should improve the ability of actions to meet the standards to some degree, because of the increase in the Ricker a parameter. Preliminary results do not tell us much about how large those effects would be, since the habitat hypothesis was implemented incorrectly.

5.8.3 Summary of Evidence

The basis for the new habitat hypothesis is an analysis of recoveries of PIT-tagged spring/summer chinook parr at downstream projects the following spring. Statistical models were built to test for associations between broad-scale habitat variables and overwintering survival. These results suggest some associations between land-use/vegetation patterns and parr-smolt survival rates. The association is particularly strong between survival rates and road density, and between survival rates and private agricultural land. However, the authors caution that these associations do not necessarily imply causality.

5.8.4 Issues to Resolve

The point raised about the current implementation of habitat effects in BSM is a good one, and we agree that we should address this in the next round of modeling. Simply fixing the current approach as outlined in the hypothesis summary (Submission 4) seems like the preferred approach, given the caveats about implying causality from statistical association. Another issue related to the analysis is the extent to which land management and physical features are intermingled in the habitat clusters (e.g., elevation is a

component of the wilderness cluster). This reduces the ability to draw inferences about the effects of switching from one habitat cluster to another, because not all of these physical characteristics can be changed through habitat management.

Bibliography (incomplete)

- Adkison, M.D., R.M. Peterman, M.L. Lapointe, D.M. Gillis, and J. Korman.** 1996. Alternative models of climatic effects on sockeye salmon productivity in Bristol Bay, Alaska, and the Fraser River, British Columbia. *Fish. Oceanogr.* 5: 137-152.
- Anderson.** 1998. Critique on transport and extra mortality hypotheses. January 12, 1998.
- Anderson.** 1996 Chapter 12 In: D.R. Marmorek. and 21 co-authors. Plan for Analyzing and Testing hypotheses (PATH): Final report on retrospective analyses for fiscal year 1996. Compiled and edited by ESSA Technologies, Ltd., Vancouver, BC.
- Anderson, J. and R. Hinrichsen.** 1997a. Prospective Analysis for the Alpha Model. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.
- Barnthouse, L.B.** 1993. Expert Initial Review of Columbia River Basin Salmonid Management Models: Summary Report. Prepared by the Oak Ridge National Laboratory for the Bonneville Power Administration.
- Basham, L. and L. Garrett.** 1996. Historical review of juvenile fish descaling information collected at the Columbia and Snake River transportation projects. In: D. Marmorek, I.P. Parnell, and D. Bouillon (eds.). 1996. Plan for Analyzing and Testing Hypotheses (PATH); preliminary report on retrospective analyses. March 15, 1996. Section 8.3 .Compiled and edited by ESSA Technologies, Ltd., Vancouver, BC.
- Beamesderfer, R.C.P., H.A. Schaller, M.P. Zimmerman, C.E. Petrosky, O.P. Langness and L. LaVoy.** 1997. Spawner - recruit data for spring and summer chinook salmon populations in Idaho, Oregon, and Washington. In: D.R. Marmorek, and C. Peters. Plan for Analyzing and Testing Hypotheses (PATH): report of retrospective analysis for fiscal year 1997. Compiled and edited by ESSA Technologies, Ltd. Vancouver, BC.
- Beamesderfer, R.C., D.L. Ward, and A.A. Nigro.** 1996. Evaluation of the biological basis for a predator control program on northern squawfish (*Ptychocheilus oregonensis*) in the Columbia and Snake Rivers. *Can. J. Fish. Aquat. Sci.* 53: 2898-2908.
- Beamish, R.J. and D.R. Bouillon.** 1993. Pacific salmon production trends in relation to climate. *Can. J. Fish. Aquat. Sci.* 50: 1002-1016.
- Berkson, J.** 1991. Coded wire tag analysis on Snake River spring, summer, and fall chinook. Letter dated February 19, 1991 to National Marine Fisheries Service for Endangered Species Act record. Columbia River Inter-Tribal Fisheries Commission. Portland, OR.
- Bjornn, T.C., Tolotti, K.R., Hunt, J.P., Keniry, P.J., Ringe, R.R., and C.A. Peery.** 1998. Passage of Chinook Salmon through the Lower Snake River and Distribution into the Tributaries, 1991-1993. Part 1 of final report for: Migration of Adult Chinook Salmon and Steelhead Past Dams and Through Reservoirs in the Lower Snake River and into Tributaries. Prepared by the Idaho Cooperative Fish and Wildlife Research Unit for the U.S. Army Corps of Engineers, Walla Walla District and Bonneville Power Administration.

Botsford, L. 1997. Depensation, performance standards, and probabilities of extinction for Columbia River spring/summer chinook salmon. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.

Botsford. (pers. comm.)

Botsford and Paulsen. 1997. Covariability in abundance among index stocks of Columbia River Spring/Summer chinook salmon. PATH FY97 Report.

Bradford, Michael J. 1994. Trends in the abundance of chinook salmon (*Oncorhynchus tshawytscha*) of the Nechako River, British Columbia. Can. J. Fish. Aquat. Sci. 51: 965-973.

Cada, G.F, M. Deacon, S. Mitz and M. Bevelhimer. 1994. Review of information pertaining to the effect of water velocity on the survival of juvenile salmon and steelhead in the Columbia River Basin. ORNL. Report prepared for NPPC. U.S. Dept of Energy Contract DE-AC05-84OR21400. 70 pp.

CBFWA. 1991.

Chapman, D., A. Giorgi, M. Hill, A. Maule, S. McCutcheon, D. Park, W. Platts, K. Pratt, J. Seeb, L. Seeb, F. Utter, 1991. Status of the Snake River Chinook Salmon, Report to the Pacific Northwest Utilities Conference Committee, February 19, 1991. (I will send section E for you and the SRP)

Dawley, E.M. and W.J. Ebel. 1975. Effects of various concentrations of dissolved atmospheric gas on juvenile chinook salmon and steelhead trout. Fish. Bull. 73: 787-786.

Dawley, E.M., M. Schiewe, and B. Monk. 1976. Effects of long-term exposure to supersaturation of dissolved atmospheric gases on juvenile chinook salmon and steelhead trout in deep and shallow tank tests. In: D.H. Fickeisen and M.J. Schneider (eds.). Gas Bubble Disease. Tech. Information Center, Oak Ridge, TN. CONF-741033, pp. 1-10.

Deriso, R.B. 1997. Prospective analysis of spring chinook of the Snake River Basin. 33 pp. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.

Deriso, R., D. Marmorek and I. Parnell. 1996. Retrospective analysis of passage mortality of spring chinook of the Columbia River. In: D.R. Marmorek. and 21 co-authors. Plan for Analyzing and Testing hypotheses (PATH): Final report on retrospective analyses for fiscal year 1996. Chapter 5 .Compiled and edited by ESSA Technologies, Ltd., Vancouver, BC.

Ebbesmeyer et al. 1998.

Fagerlund, U.H.M., J.R. McBride and I.V. Williams. 1995. Stress and Tolerance. In: C. Groot, L. Margolis and W.C. Clarke (eds.). Physiological Ecology of Pacific Salmon. 1995. Chapter 8. UBC Press, Vancouver, Canada.

Fryer, J.K. and M. Schwartzberg. 1993. Age and length composition of Columbia Basin spring and summer chinook salmon at Bonneville Dam in 1992. CRITFC Technical Report 93-3. Columbia River Inter-Tribal Fisheries Commission. Portland, OR.

- Hanrahan, T.P., Neitzel, D.A., and M.C. Richmond.** 1998. Assessment of Drawdown From a Geomorphic Perspective Using Geographic Information Systems, Lower Snake River, Washington. Prepared by the Pacific Northwest National Laboratory for the US Army Corps of Engineers, Walla Walla District.
- Hanski.** I. 1991. Single species metapopulation dynamics: brief history and conceptual domain. *Biol. Jour. Linnean Soc.* 42: 3-16.
- Hare, S.R., N.J. Mantua, and R.C. Francis. (in press).** Inverse production regimes: Alaska and West Coast pacific salmon. *Fisheries*, in review.
(See http://www.iphc.washington.edu:80/PAGES/IPHC/Staff/hare/html/papers/inverse/abst_inv.html)
- Harrison S. and J.F. Quinn.** 1989. Correlated environments and the persistence of metapopulations. *Oikos* 56:293-298.
- Hinrichsen and Paulsen.** 1998. Submission 3 to WOE report.
- Hoar, W.S.** 1976. Smolt transformation: evolution, behaviour, and physiology. *J. Fish. Res. Board. Can.* 33: 1234-1252.
- Hymer, J.A., R. Pettit, M. Wastel, P. Hahn and K. Hatch.** 1992. Stock summary reports for Columbia River anadromous salmonids. Volume IV: Washington. Coordinated Information System. Prepared for US Department of Energy, Bonneville Power Administration. Project 88-108. Contract No. DE-FC79-89BP94402.
- ISG.** 1996. Return to the River: Restoration of Salmonid Fishes in the Columbia River Ecosystem. Report prepared by the Independent Scientific Group, 10 September 1996.
- Kiefer.** IDFG. (pers. comm.)
- Libero, F.E.de.** 1986. A statistical assessment of the use of the coded wire tag for chinook (*Oncorhynchus tshawytscha*) and coho (*Oncorhynchus kisutch*) studies. Ph.D. dissertation, University of Washinton.
- Lichatowich, J.** 1993. Ocean carrying capacity. Mobrاند Biometrics, Inc. Draft prepared for S.P. Cramer and Associates under contract to Bonneville Power Administration.
- Litchatowich, J, L. Mobrand, L Lestelle, T. Vogel,** 1996. Chinook Salmon in the Columbia River: the Components of Decline. In the Publication: Applied Ecosystem Analysis – Background. Submitted to the Bonneville Power Administration, May 1996. Report No 5
- Lindsay, R.B., W.J. Knox, M.W. Flesher, B.J. Smith, E.A. Olsen and L.S. Lutz.** 1986. Study of wild spring chinook salmon in the John Day River system. Oregon Department of Fish and Wildlife. 1985 Final Report. Prepared for Prepared for US Department of Energy, Bonneville Power Administration. Project 79-4. Contract No. DE-AI79-83BP39796.
- Marmorek, D.R. and C.N. Peters** (eds.) 1998. Plan for Analyzing and Testing Hypotheses (PATH): Preliminary decision analysis report on the Snake River spring/summer chinook. Draft report compiled and edited by ESSA Technologies Ltd., Vancouver BC.

Marmorek, D.R. and C.N. Peters (eds.) 1998. Plan for Analyzing and Testing Hypotheses (PATH): Recent Analyses of D Values. Draft report compiled and edited by ESSA Technologies Ltd., Vancouver BC.

Marmorek, D.R. and C.N. Peters (eds.) 1996. Plan for Analyzing and Testing Hypotheses (PATH): Conclusions of FY 96 Retrospective Analyses. Report compiled and edited by ESSA Technologies Ltd., Vancouver BC.

Matthews, G.M. and R.S. Waples. 1991. Status review for Snake River spring and summer chinook salmon. U.S. Dept. for Commer., NOAA Tech. Memo. NMFS F/NWC-200, 75 pp.

Matthews and Williams. 1985.

Mesa, M.G., J.J. Warren, K.M. Hans, and A.G. Maule. 1996. Progression and severity of gas bubble trauma in juvenile chinook salmon and development of non-lethal methods for trauma assessment. In: Gas bubble Trauma Monitoring and Research of Juvenile Salmonids. 1996 Annual report of the US Geological Survey, Biological Resources Division, Columbia River Research Laboratory, Cook, WA. Prepared for the Bonneville Power Administration, Portland, OR.

Mundy, P.R. and 9 co-authors. 1994. Transportation of juvenile salmonids from hydroelectric projects in the Columbia River Basin; an independent peer review. Final Report, US Fish and Wildlife Service, 911 N.E. 11th Avenue, Portland, OR 97232-4181.

Myers et al. 1995.

Paulsen. 1997. Chapter 4 update. Effects of climate and land use on stock recruitment.

Paulsen. 1996. Chapter 4. Level 2 Hypotheses. FY96 Retrospective Analyses.

Paulsen, C. 1995. Level 1 and 2 hypothesis testing: Preliminary results for chinook. In: D.R.Marmorek and I. Parnell. 1995. PATH Information Package for Workshop 1.

Paulsen, C. and T. Fisher. 1997. Update on Ocean Distribution of Coded Wire Tagged Spring/Summer Chinook. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.

Pearcy, W.G. 1992. Ocean Ecology of North Pacific Salmonids. University of Washington Press. Seattle.

Pearcy, W.G. 1988. Factors affecting survival of coho salmon off Oregon and Washington. In: W.J. McNeil (ed.). Salmon Production, Management, and Allocation: Biological, Economic, and Policy Issues. Oregon State University Press, Corvallis, OR.

Pearl, R. 1928. The Rate of Living. Knopf, New York.

Peterman et al. 1998. unpubl. data.

Peterman, R.M., F.J. Pyper, M.F. Lapointe, and M.D. Adkison. 1997. Patterns of covariation in components of recruitment among sockeye salmon stocks in British Columbia and Alaska. In: R.L. Emmett and M.H. Schiewe (eds.). Estuarine and Ocean Survival of Northeastern Pacific Salmon. Proceedings of the Workshop, March 1996, Newport, Oregon, US Department of Commerce, NOAA Tech. Memo. NMFS-NWFSC-29: 243-247.

Peters, C.N. and D. R. Marmorek. 1997. A generalized prospective modeling framework. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.

Peters, C.N. et al. 1997. Alternative structures for a hydro action decision tree. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.

Peterson, J.H. 1994. Importance of spatial pattern in estimating predation of juvenile salmonids in the Columbia River. *Trans. Amer. Fish. Soc.* 123: 924-930.

Petersen, J.H. and T.P. Poe. 1997. Predicting the effects of reservoir drawdown on juvenile salmonoids and their predators. Research proposal submitted 24 June 1997.

Petrosky and Schaller. 1996

PSC. 1994.

Raymond, H.L. 1988. Effects of hydroelectric development and fisheries enhancement on spring and summer chinook salmon and steelhead in the Columbia River Basin. *N. Amer. J. Fish. Manage.* 8: 1-24.

Raymond, H.L. 1979. Effects of dams and impoundments on migrations of juvenile chinook salmon and steelhead from the Snake River, 1966 to 1975. *Trans. Amer. Fish. Soc.* 108: 505-529.

Raymond, H.L. 1975. Snake River runs of salmon and steelhead trout: trends in abundance of adults and downstream survival of juveniles. Report of research financed by U.S. Army Corps of Engineers and National Marine Fisheries Service (CZES). Revised November 1975.

Raymond. 1978.

Raymond. 1977.

Rieman B.E. and J.D. McIntyre. 1993. Demographic and habitat requirements for conservation of bull trout. Ogden, US Forest Service, Intermountain Research Station. Utah: 38

Roby et al. 1998. Avian predation on juvenile salmonids in the Lower Columbia River. Oregon Cooperative Wildlife Research Unit. Report to BPA. March 1998.

Schaller, H.A., C.E. Petrosky and O.P. Langness. 1996. Contrasts in stock recruitment patterns of Snake and Columbia River spring and summer chinook populations. In: D.R. Marmorek and 21 co-authors. Plan for Analyzing and Testing hypotheses (PATH): Final report on retrospective analyses for fiscal year 1996. Chapter 3. Compiled and edited by ESSA Technologies, Ltd., Vancouver, BC.

Schaller et al. In: Prelim. Report (pg. A-107).

Skalski, J.R. 1998. Estimating season-wide survival rates of outmigrating salmon smolt in the Snake River, Washington. *Can. J. Fish. Aquat. Sci.* 55: 761-769.

Skalski et al. (in press) *Can. J. Fish. Aquat. Sci.*

Starr et al. 1998. *Can. J. Fish. Aquat. Sci.* 55: 529-537.

Steward 1994, Assessment of the Flow-Survival Relationship Obtained by Sims and Ossiander (1981) for Snake River Spring/Summer Chinook Salmon Smolts. Report to Bonneville Power Administration, DOE/BP-99654-13, April 1994.

Toole, C. 1997. Integration of PATH Chapter 6 Interim Survival Goals with Chapter 5 Passage Mortality Estimation Procedures

Toole, C., A. Giorgi, E. Weber and C. McConnaha. 1996. Hydro decision pathway and review of existing information. In: D.R. Marmorek and 21 co-authors. Plan for Analyzing and Testing Hypotheses (PATH): final report of retrospective analysis for fiscal year 1996. Chapter 6. Compiled and edited by ESSA Technologies.

Vigg, S. 1988. Functional response of the northern Squawfish predation to salmonid prey density in the McNary Tailrace, Columbia River. In: T. Poe and B. Rieman (eds.). Predation by Resident Fish on Juvenile Salmonids in John Day Reservoir, 1963-1986. Vol. 1.

Wahle, R.J., E. Chaney, and R.E. Pearson. 1981. Areal distribution of marked Columbia River basin Spring Chinook salmon recovered in fisheries and at parent hatcheries. Mar. Fish. Rev. 43 (12): 1-9.

Ware. 1995.

Weber, E., P. Wilson, H. Schaller, R.Beamesderfer and C.Petrosky. 1997. Internal PATH Review of Williams, Matthews and Myers. 1997. The Columbia River Hydropower System: Does It Limit Recovery of Spring/Summer Chinook Salmon? DRAFT Review Comments, 6/97.

West Coast Chinook Biological Review Team. 1996. Review of the status of chinook salmon (*Oncorhynchus tshawytscha*) from Wasington, Oregon, California, and Idaho under the US Endangered Species Act. November 1996 draft report sent to salmon co-managers for review.

Williams. 1989.

Williams, J.G. and Mathews. 1995. A review of flow and survivalrelationships for spring and summer chinook salmon, *Oncorhynchus tshawytscha*, from the Snake River Basin. Fish. Bull. 93: 732-740.

Wilson, P., Weber, E., Petrosky, C., and H. Schaller. 1997. Draft General Framework for Prospective Modeling, with One Proposed Hypothesis on Delayed Mortality. In: PATH Retrospective and Prospective Analyses of Spring/Summer Chinook Reviewed in FY 1997.

Glossary of Terms, Acronyms, Variables and Parameters

a: extra mortality in a given year for a given sub-region (i.e., Snake River, Lower Columbia River).

d: year – effect parameter for a given year (common year effects affecting both upstream and downstream stocks).

e: normally distributed mixed process error and recruitment measurement, which depends on year and sub-basin.

I: post-Bonneville survival factors for transported (I_t) and non-transported smolts (I_n).

m Incremental total mortality between the Snake River Basin and the John Day project in a specific year.

w: system survival ($e^{-M} + [DP + 1 - P]$).

a: Ricker a parameter.

b: Ricker b parameter.

A1, A2, A3: Management Actions (see Table 3-1).

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

AIC (Akaike Information Criterion): $-2 \ln(\text{Likelihood}) + 2p$, where $p = \#$ parameters.

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).

BIC (Bayesian Information Criterion): $-2 \ln(\text{Likelihood}) + p \cdot \ln(k)$, where $p = \#$ parameters and $k = \#$ observations.

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

BON (Bonneville Dam)

BPA (Bonneville Power Administration)

BRWG (Biological Requirements Working Group)

BSM (Bayesian Simulation Model)

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

CARTs (Categorical Regression Trees).

cp: Complexity parameter.

CPUE: (Catch Per Unit Effort)

CRFMP (Columbia River Fish Management Plan)

CRiSP (Columbia River Salmon Passage Model)

CWT (Coded wire tag): 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: Ratio of post-Bonneville survival of transported fish to 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.

E: Climate index variable (PAPA drift). Represents the latitude of a drifting object after three months drift starting at station PAPA.

EJUV: Equilibrated Juvenile survival rates.

EM (Extra Mortality): Extra mortality is any mortality occurring outside of the juvenile migration corridor that is not accounted for by either: 1) productivity parameters in spawner-recruit relationships; 2) estimates of direct mortality within the migration corridor (from passage models); or 3) for the delta model only, common year effects affecting both Snake River and Lower Columbia River stocks.

EMCLIM: Extra Mortality / future Climate.

ESA (Endangered Species Act)

ESBS (Extended Length Submersible Bar Screens)

ESU (Evolutionary Significant Unit): 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.

F: Average flow (in thousand cubic feet per second) at Astoria during April-June.

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

FLUSH (Fish Leaving Under Several Hypotheses): a passage model developed by the State and Tribal fish agencies

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

GBT (Gas Bubble Trauma): non-lethal or lethal effects of the growth of air bubbles in the cardiovascular systems of fish.

HAB: Habitat effects.

HYSER (): A U.S. Army Corps hydro-regulation model to predict monthly flows associated with a particular method of operating the hydrosystem.

IHR / IHR (Ice Harbor Dam)

ISAB (): Independent Scientific Advisory Board; provides independent advice and reviews to NMFS and the NPPC.

I.T. (Implementation Team): an inter-agency policy group to whom PATH reports.

In-river survival rate: 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.

JDA (John Day 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).

KCFS: A unit of measure for flowing water, expressed in thousands of cubic feet per second.

LGO/LGS (Little Goose Dam)

LGR (Lower Granite Dam)

LMO/LMN (Lower Monumental Dam)

m: Total direct passage mortality rate, including both passage and extra mortality.

Dm: Extra mortality rate, expressed as an instantaneous rate, which depends on year and region, and is calculated as the differences between total mortality (m) and passage mortality (M).

M: Direct instantaneous passage mortality rate of juvenile fish (both transported and non-transported) from LGR pool to below BON.

MCN (McNary Dam)

MLE (Maximum Likelihood Estimate)

NMFS (National Marine Fisheries Service)

NPPC (Northwest Power Planning Council)

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.

OSCURS (): An ocean circulation model.

p: Depensation parameter.

P or Pbt: The proportion of juvenile fish below BON that were transported.

PAPA (): an index of ocean currents

PATH (Plan for Analyzing and Testing Hypotheses)

PIT (Passive Integrated Transponder) tags: These 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.

PDO (Pacific Decadal Oscillation)

PMOD: Passage Model.

PREM: Predator Removal effectiveness.

PRER: Length of pre-removal period.

PROSP: Prospective model for the distribution of extra mortality (Alpha or Delta).

PSMFC ():

R: "Observed" returns (recruitment) originating from spawning.

R/S: Recruits per spawner is the number of adult fish returning to BON (R) divided by the number of spawners in the parent generation (S).

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).

s: FLUSH variable for survival to below BON of control (non-transported) fish.

S: "Observed" spawning.

SAR (Smolt-to-adult return rate): 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.

SRP (Scientific Review Panel)

STEP: Formulated to model the effect of a 1975 (brood year) climate regime shift, which has different effects in different subregions.

STS (Standard Length Submersible Travel Screens)

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.

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.

T:C or T/C or TCR: The Transport : Control ratio is the ratio of transported fish survival to in-river fish survival from juveniles at the collection point to adults at the same point.

TDD (The Dalles Dam)

TRANS: Transportation model.

TJUV: Transition period: Juvenile survival.

TURB: Historical turbine / bypass survival assumptions.

Vb: Survival of barge-transported fish.

Vc: Survival to below BON of control (non-transported) fish.

Vcl: CRiSP variable for survival to below BON of control (non-transported) fish.

Vn: Direct passage survival of in-river juvenile fish, measured from the head of LGR pool to the tailrace of BON, including reservoir and dam survival at each project.

Vpre-b: Survival of fish destined to be transported to the point of collection. (pg. D-9, line 33)

Vt: In-river survival of transported juvenile fish. This value takes into account in-river survival from the head of LGR to the point of collection, bypass survival at the collection project, and barge survival to BON tailrace.

WOE (Weight of Evidence)

WTT (Water Transit Time)

Appendix A: Detailed Results

A.1 Overall Results

The changes described in Section 3.1 result in less distinction between the two passage models in terms of which action is favored (Figure A-1). With both models, a higher proportion of runs (incorporating all alternative hypotheses) achieve all 3 of the jeopardy standards (i.e., 24-year survival, 100-year survival, and 48-year recovery) under action A3 (with either a three or eight year delay) than under A1 or A2. These proportions are higher using CRiSP T4 assumptions for all actions.

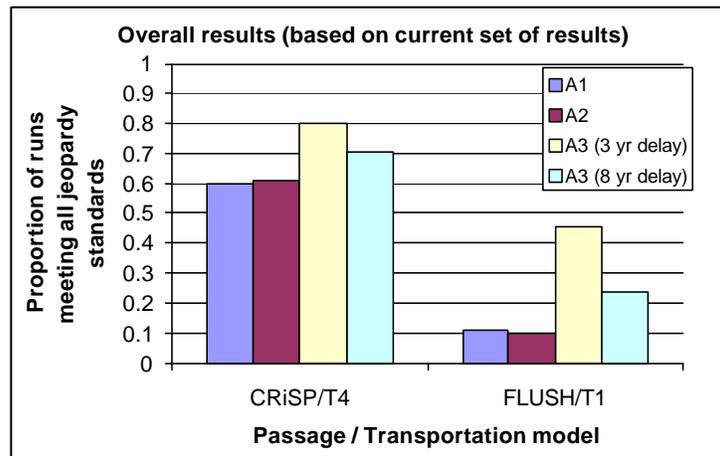


Figure A-1: Proportion of runs meeting all standards, by passage/transportation model.

In addition to the three official jeopardy standards, the Biological Requirements Working Group also developed a 24-year recovery “indicator”. Although the primary focus is on the ability of actions to meet the three official jeopardy standards, we also show the proportion of runs in which the 24-year recovery probability equals or exceeds 0.5 in Figure A-2.

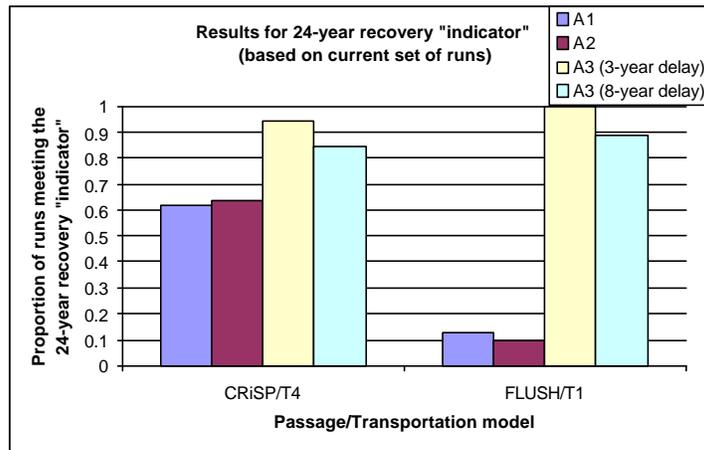
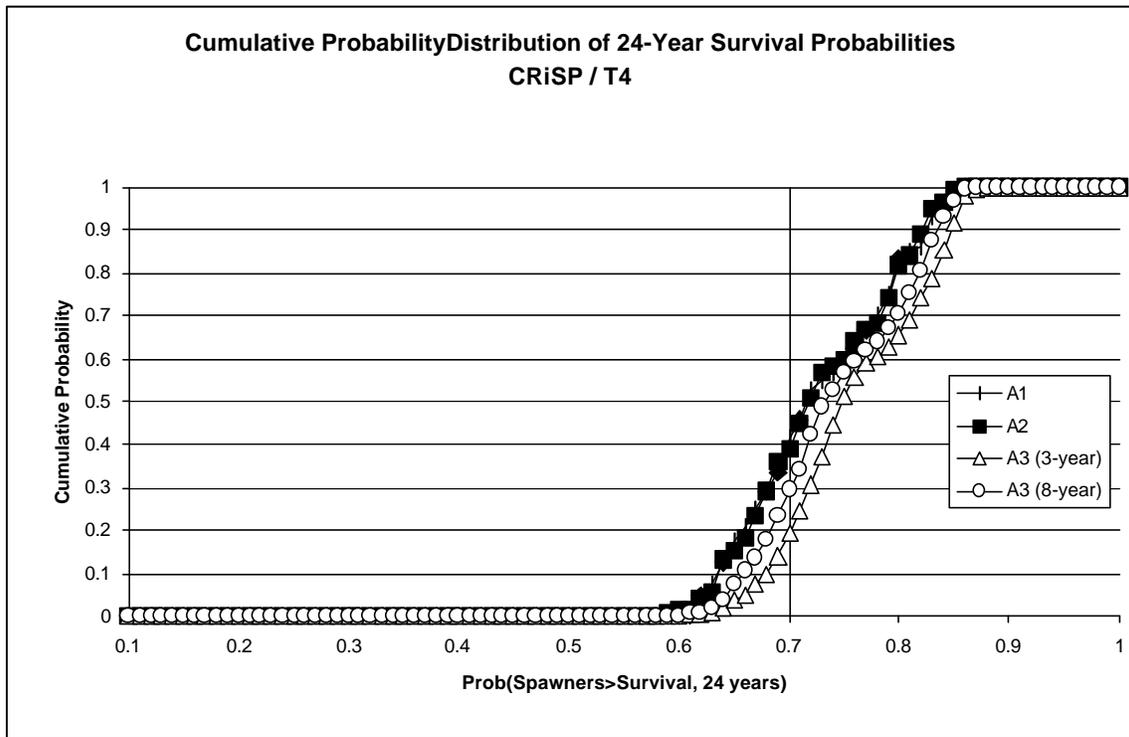
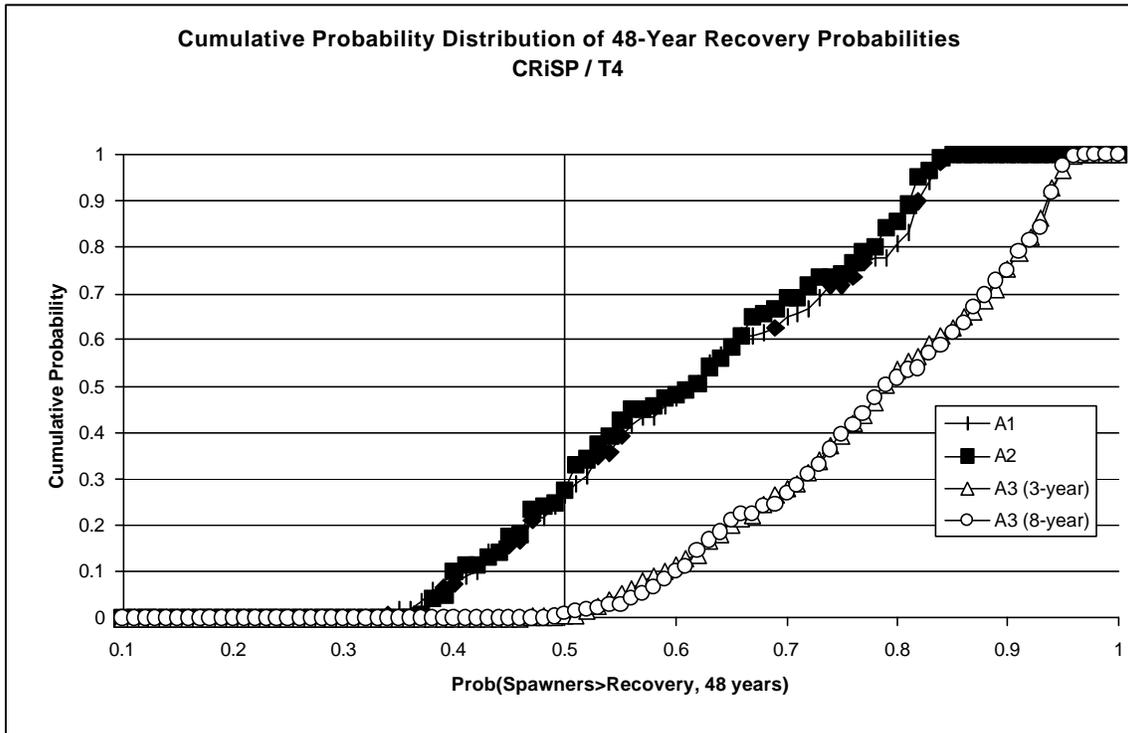
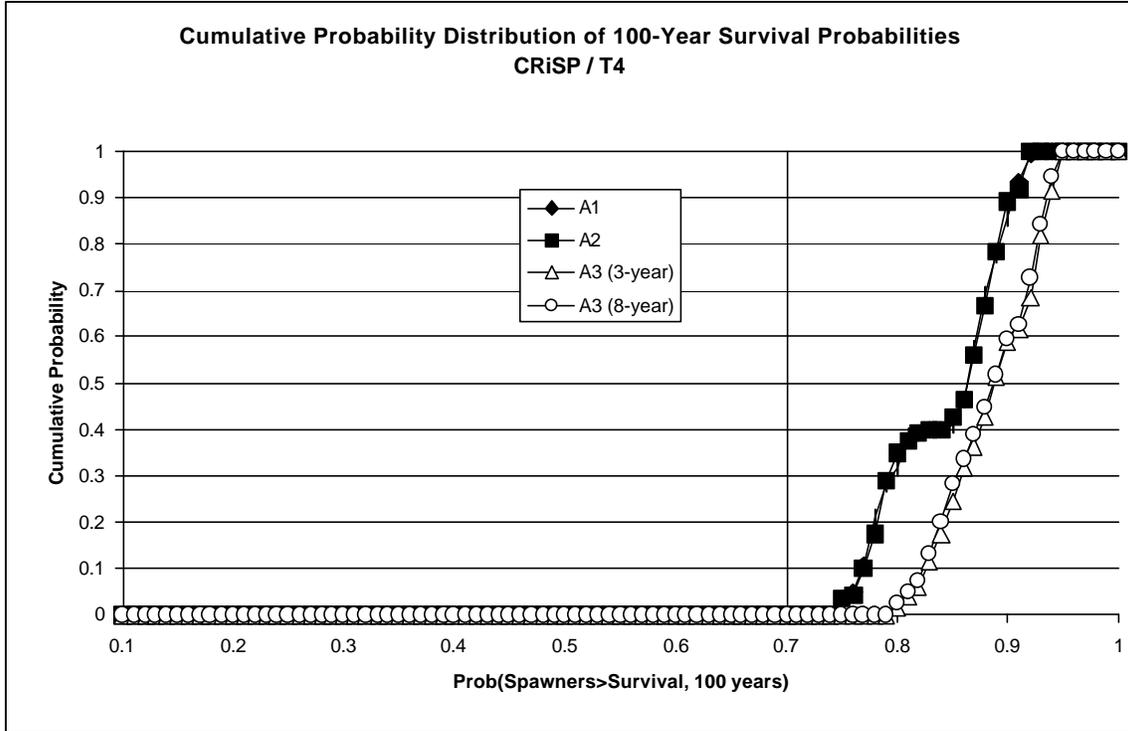
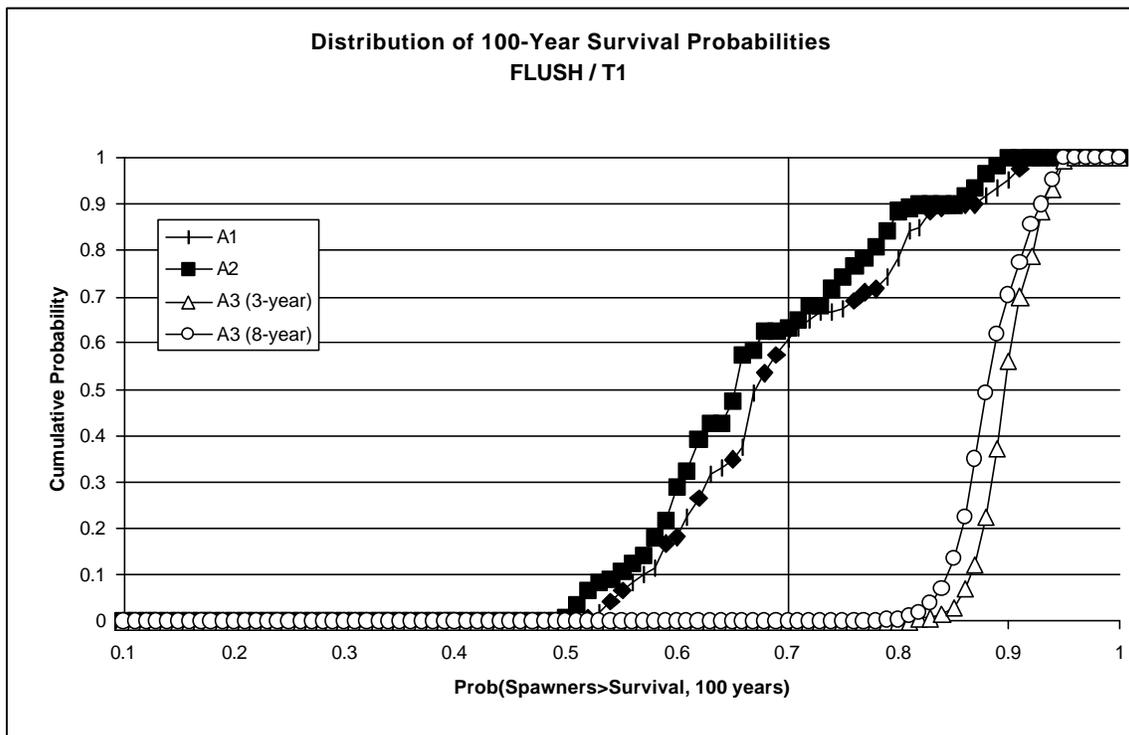
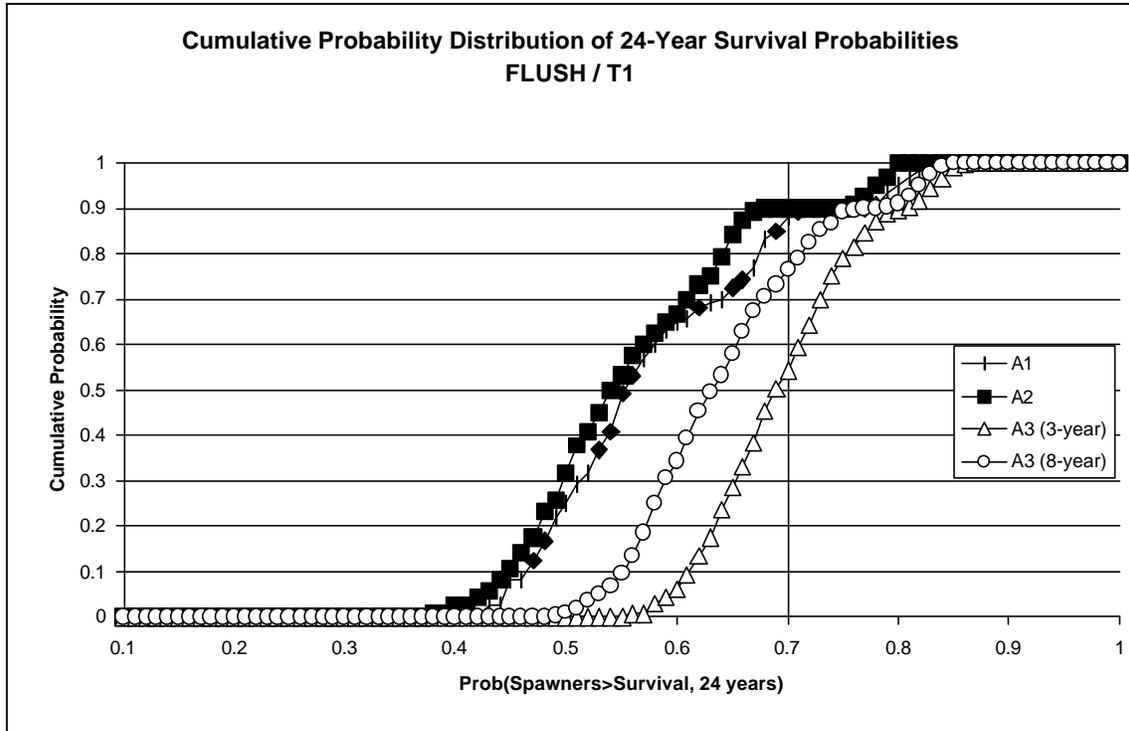


Figure A-2: Proportion of runs meeting the 24-year recovery indicator, by passage/transportation model.

The jeopardy standards are “knife-edged,” in the sense that a 24-year survival probability of 0.71, for example (i.e., the average probability over 24 years that the number of spawners will exceed the defined survival level of spawners), is deemed to be just as good as a probability of 0.98. Both of these exceed the defined criterion of 0.7, but obviously a higher probability would be preferred. Figure A-3 provides a more complete picture of the results. This figure shows the cumulative frequency distribution of jeopardy probabilities for each of the 3 jeopardy standards. These graphs are useful because they illustrate both the relative performance of the different actions (as measured by the horizontal distance between the lines) and the ability of the actions to meet the standards (standards are indicated by vertical lines). For example, with CRiSP/T4 and the 24-year survival standard, the A3 action performs slightly better than A1 or A2. Approximately 60% of A1 and A2 runs, and approximately 70% (80%) of A3 runs with a 3-year (8-year) delay meet the 24-year survival standard. In contrast, with FLUSH/T1 and the 48-year recovery standard, A3 outperforms A1 and A2 by a wide margin. For this passage model, the recovery standard is met by only 10-20% of A1 and A2 runs, but is met by 100% of the A3 runs.







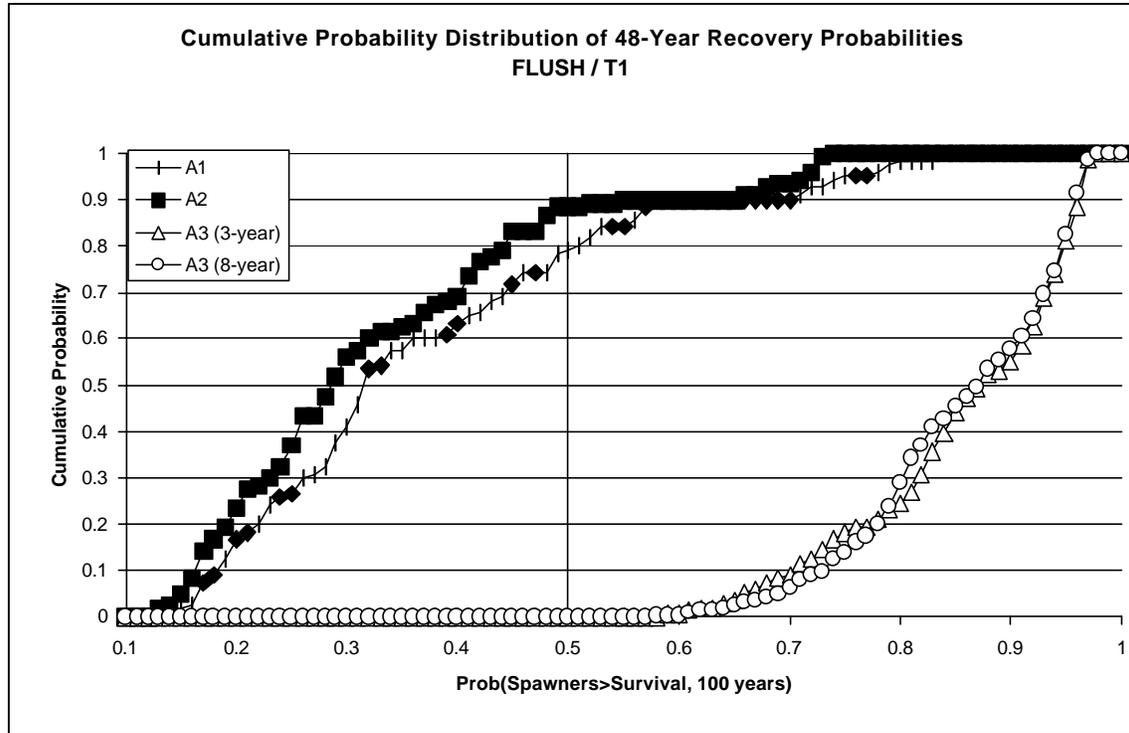


Figure A-3: Cumulative Probability Distribution of jeopardy probabilities. Jeopardy standards are indicated by vertical lines.

A.2 Effects of Hypotheses on Ability of Actions to Meet Standards

Section 3.3 summarized results of sensitivity analyses to determine which hypotheses had the greatest effect on the ability of actions to meet all of the Jeopardy Standards. This section provides more details on these analyses. We looked at the contribution of individual hypotheses to the runs that met all of the standards (0.7 for survival standards, 0.5 for recovery standards). If a hypothesis has a strong effect on the results, its proportional representation in the runs that meet the standards will be different from its proportional representation in all of the runs.

Results for individual hypotheses (broken down into CRiSP and FLUSH subsets) are presented in Tables A-1 and A-2. The structure of these tables is similar to Table 3-4. Column 2 shows the proportion of all runs that include each individual hypothesis. Columns 3-6 show the proportion of runs for each action that meet all of the jeopardy standards. Columns 7-10 show the percentage change from the overall proportion of all runs for that passage model, as a scaled indicator for comparison.

Table A-1: Effects of individual hypotheses on ability of actions to meet all of the standards (CRiSP passage model). Hypotheses are defined in Table 3-2.

Hypothesis	Proportion of all CRiSP runs	Proportion of CRiSP runs that meet all of the jeopardy standards				% change from proportion of all CRiSP runs			
		A1	A2	A3 (3-yr)	A3 (8-yr)	A1	A2	A3 (3-yr)	A3 (8-yr)
FGE1	0.5	0.54	0.56	0.52	0.55	+8.3	+12.3	+4.7	+9.7
FGE2	0.5	0.47	0.42	0.47	0.45	-5.6	-15.1	-5.7	-9.7
TURB1	0.333	0.31	0.29	0.30	0.28	-8.2	-13.6	-9.0	-15.8
TURB4	0.333	0.38	0.36	0.34	0.35	+12.6	+7.0	+1.1	+6.3
TURB5	0.333	0.33	0.36	0.35	0.36	-0.1	+7.0	+5.8	+8.1
PREM1	0.5	0.47	0.47	0.44	0.46	-5.6	-6.8	-11.4	-7.4
PREM3	0.5	0.54	0.52	0.55	0.54	+8.3	+4.1	+10.9	+7.4
PROSPA	0.5	0.39	0.41	0.43	0.40	-22.2	-17.8	-14.0	-20.4
PROSPD	0.5	0.64	0.60	0.57	0.60	+27.8	+20.5	+14.0	+19.2
BKD/Markov	0.2	0.10	0.08	0.14	0.10	-51.4	-58.9	-28.8	-48.4
BKD/Cyclic	0.2	0.07	0.05	0.13	0.10	-65.3	-72.6	-32.6	-51.3
Hydro/Markov	0.2	0.33	0.32	0.25	0.28	+66.7	+57.5	+23.1	+40.1
Hydro/Cyclic	0.2	0.33	0.32	0.25	0.28	+66.7	+57.5	+23.1	+40.1
Regime Shift	0.2	0.19	0.22	0.23	0.24	-2.8	+9.6	+12.7	+18.0
EJUV1/TJUVa	0.25			0.27	0.26			+6.7	+3.8
EJUV1/TJUVb	0.25			0.21	0.22			-15.0	-10.3
EJUV2/TJUVa	0.25			0.28	0.28			+13.0	+10.9
EJUV2/TJUVb	0.25			0.24	0.24			-3.6	-4.4

Table A-2: Effects of individual hypotheses on ability of actions to meet all of the standards (FLUSH passage model).

Hypothesis	Proportion of all FLUSH runs	Proportion of FLUSH runs that meet all of the jeopardy standards				% change from proportion of all FLUSH runs			
		A1	A2	A3 (3-yr)	A3 (8-yr)	A1	A2	A3 (3-yr)	A3 (8-yr)
FGE1	0.5	0.54	0.50	0.53	0.59	+7.7	0.0	+5.9	+17.5
FGE2	0.5	0.46	0.50	0.47	0.41	-7.7	0.0	-5.9	-17.5
TURB1	0.333	0.31	0.33	0.32	0.31	-7.6	0.1	-5.4	-7.8
TURB4	0.333	0.38	0.33	0.36	0.39	+15.5	0.1	+8.3	+15.9
TURB5	0.333	0.31	0.33	0.32	0.31	-7.6	0.1	-2.6	-7.8
PREM1	0.5	0.08	0.00	0.36	0.35	-84.6	-100.0	-27.9	-29.8
PREM3	0.5	0.92	1.00	0.64	0.65	+84.6	+100.0	+27.9	+29.8
PROSPA	0.5	0.00	0.00	0.20	0.23	-100.0	-100.0	-59.8	-54.4
PROSPD	0.5	1.00	1.00	0.80	0.77	+100.0	+100.0	+59.8	+54.4
BKD/Markov	0.2	0.00	0.00	0.14	0.00	-100.0	-100.0	-31.5	-100.0
BKD/Cyclic	0.2	0.00	0.00	0.11	0.00	-100.0	-100.0	-47.5	-100.0
Hydro/Markov	0.2	0.54	0.50	0.32	0.48	+169.2	+150.0	+59.8	+141.2
Hydro/Cyclic	0.2	0.46	0.50	0.32	0.52	+130.8	+150.0	+59.8	+158.8
Regime Shift	0.2	0.00	0.00	0.12	0.00	-100.0	-100.0	-40.6	-100.0
EJUV1/TJUVa	0.25			0.31	0.28			+22.4	+12.3
EJUV1/TJUVb	0.25			0.16	0.20			-34.2	-19.3
EJUV2/TJUVa	0.25			0.33	0.29			+31.5	+15.8
EJUV2/TJUVb	0.25			0.20	0.23			-19.6	-8.8

These tables show the effects of single factors on the ability of actions to meet the standards, but they do not show the effects of potential interactions. For example, a long transition period is important, but it may only be important if coupled with regime shift extra mortality. To identify those interactions, Tables A-3 to A-12 below show the number of runs for a specific combination of hypotheses out of the total of runs in which all actions met the jeopardy standards. The total number of runs for each action and passage model that met all of the standards were summarized in Table 3-3. Hypotheses are coded as follows (See Table 3-2 for a description of all the alternative hypotheses):

Passage-related hypotheses are grouped into 12 overall combinations, identified by Fish Guidance Efficiency (FGE) / Historical turbine/bypass survival assumptions (TURB) / Predator Removal Effectiveness (PREM). For example, passage hypothesis 143 identifies the combination of FGE1(high FGEs), TURB4 (lower turbine/bypass survival in the past), and PREM3 (25% reduction in reservoir mortality due to removal of predators).

Drawdown assumptions are grouped into 8 overall combinations, identified by length of pre-removal period (PRER) / equilibrated juvenile survival rates following drawdown (EJUV) / length of transition period (TJUV).

Prospective assumptions are grouped into 10 combinations of 2 prospective models (Alpha (A) or Delta (D)) X 5 Extra mortality/Future climate scenarios. These scenarios are numbered as follows* :

1. BKD extra mortality, Markov future climate
3. Hydro-related extra mortality, Markov future climate
5. BKD extra mortality, Cyclical future climate
7. Hydro-related extra mortality, Cyclical future climate
9. Regime shift extra mortality, Cyclical future climate

Prospective assumption A5, for example, identifies an Alpha model run, coupled with BKD extra mortality / Cyclical future climate.

Table A-3: Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A1 to meet all jeopardy standards (CRiSP/T4).

A1 Passage	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	0	1	0	1	0	0	1	0	1	1	5
113	0	1	0	1	0	1	1	0	1	1	6
141	0	1	0	1	0	1	1	0	1	1	6
143	0	1	0	1	1	1	1	1	1	1	8
151	0	1	0	1	0	1	1	0	1	1	6
153	0	1	0	1	0	1	1	1	1	1	7
211	0	1	0	1	0	0	1	0	1	1	5
213	0	1	0	1	0	0	1	0	1	1	5
241	0	1	0	1	1	0	1	0	1	1	6
243	0	1	0	1	0	1	1	1	1	1	7
251	0	1	0	1	0	0	1	0	1	1	5
253	0	1	0	1	0	0	1	0	1	1	5
Grand Total	0	12	0	12	2	6	12	3	12	12	71

* Even numbers identify the HABB variants on these hypotheses and are omitted from this analysis.

Table A-4: Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A2 to meet all jeopardy standards (CRiSP/T4).

A2 Passage	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	0	1	0	1	0	0	1	0	1	1	5
113	0	1	0	1	0	1	1	0	1	1	6
141	0	1	0	1	1	1	1	0	1	1	7
143	0	1	0	1	1	1	1	1	1	1	8
151	0	1	0	1	1	1	1	1	1	1	8
153	0	1	0	1	1	1	1	1	1	1	8
211	0	1	0	1	0	0	1	0	1	1	5
213	0	1	0	1	0	0	1	0	1	1	5
241	0	1	0	1	0	0	1	0	1	1	5
243	0	1	0	1	1	0	1	0	1	1	6
251	0	1	0	1	0	0	1	0	1	1	5
253	0	1	0	1	0	0	1	0	1	1	5
Grand Total	0	12	0	12	5	5	12	3	12	12	73

Table A-5: Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A3 to meet all jeopardy standards (CRiSP/T4).

A3 Passage	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	0	8	0	8	3	4	8	3	8	8	50
113	2	8	2	8	8	7	8	7	8	8	66
141	0	8	0	8	7	8	8	8	8	8	63
143	3	8	3	8	8	8	8	8	8	8	70
151	3	8	3	8	8	6	8	6	8	8	66
153	5	8	5	8	8	8	8	8	8	8	74
211	0	8	0	8	2	2	8	0	8	7	43
213	1	8	1	8	5	5	8	3	8	8	55
241	0	8	0	8	4	5	8	5	8	8	54
243	1	8	1	8	7	8	8	8	8	8	65
251	0	8	0	8	6	3	8	3	8	8	52
253	3	8	3	8	8	7	8	7	8	8	68
Grand Total	18	96	18	96	74	71	96	66	96	95	726

Table A-6: Effects of Interactions between Prospective hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (CRiSP/T4).

A3 Drawdown	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
11a	4	12	4	12	12	12	12	11	12	12	103
11b	1	12	1	12	7	7	12	6	12	12	82
12a	7	12	7	12	12	12	12	11	12	12	109
12b	3	12	3	12	9	9	12	9	12	12	93
21a	1	12	1	12	8	10	12	8	12	12	88
21b	0	12	0	12	9	4	12	4	12	11	76
22a	2	12	2	12	10	10	12	10	12	12	94
22b	0	12	0	12	7	7	12	7	12	12	81
Grand Total	18	96	18	96	74	71	96	66	96	95	726

Table A-7: Effects of Interactions between Prospective hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (CRiSP/T4).

A3 Drawdown	Passage Hypotheses												Grand Total
	111	113	141	143	151	153	211	213	241	243	251	253	
11a	8	8	8	10	10	10	7	8	8	8	8	10	103
11b	5	8	7	10	6	8	5	6	5	8	6	8	82
12a	8	10	8	10	10	10	7	10	8	10	8	10	109
12b	5	10	8	8	8	10	5	8	7	8	6	10	93
21a	6	8	8	8	8	10	5	6	8	8	5	8	88
21b	6	6	8	8	6	8	4	5	5	8	6	6	76
22a	7	8	8	8	10	10	5	6	8	8	8	8	94
22b	5	8	8	8	8	8	5	6	5	7	5	8	81
Grand Total	50	66	63	70	66	74	43	55	54	65	52	68	726

Table A-8: Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A1 to meet all jeopardy standards (FLUSH/T1).

A1 Passage	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	0	0	0	0	0	0	0	0	0	0	0
113	0	0	0	0	0	0	1	0	1	0	2
141	0	0	0	0	0	0	1	0	0	0	1
143	0	0	0	0	0	0	1	0	1	0	2
151	0	0	0	0	0	0	0	0	0	0	0
153	0	0	0	0	0	0	1	0	1	0	2
211	0	0	0	0	0	0	0	0	0	0	0
213	0	0	0	0	0	0	1	0	1	0	2
241	0	0	0	0	0	0	0	0	0	0	0
243	0	0	0	0	0	0	1	0	1	0	2
251	0	0	0	0	0	0	0	0	0	0	0
253	0	0	0	0	0	0	1	0	1	0	2
Grand Total	0	0	0	0	0	0	7	0	6	0	13

Table A-9: Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A2 to meet all jeopardy standards (FLUSH/T1).

A2 Passage	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	0	0	0	0	0	0	0	0	0	0	0
113	0	0	0	0	0	0	1	0	1	0	2
141	0	0	0	0	0	0	0	0	0	0	0
143	0	0	0	0	0	0	1	0	1	0	2
151	0	0	0	0	0	0	0	0	0	0	0
153	0	0	0	0	0	0	1	0	1	0	2
211	0	0	0	0	0	0	0	0	0	0	0
213	0	0	0	0	0	0	1	0	1	0	2
241	0	0	0	0	0	0	0	0	0	0	0
243	0	0	0	0	0	0	1	0	1	0	2
251	0	0	0	0	0	0	0	0	0	0	0
253	0	0	0	0	0	0	1	0	1	0	2
Grand Total	0	0	0	0	0	0	6	0	6	0	12

Table A-10: Effects of Interactions between Prospective hypotheses and passage hypotheses on the ability of A3 to meet all jeopardy standards (FLUSH/T1).

A3 Passage	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	0	0	0	0	0	2	8	2	8	2	22
113	0	7	0	7	0	3	8	2	8	2	37
141	0	0	0	0	0	2	8	2	8	2	22
143	0	8	0	8	0	4	8	4	8	4	44
151	0	0	0	0	0	2	8	2	8	2	22
153	0	6	0	7	0	3	8	2	8	2	36
211	0	0	0	0	0	2	6	1	6	1	16
213	0	3	0	4	0	2	8	2	8	2	29
241	0	0	0	0	0	2	8	1	8	2	21
243	0	6	0	6	0	3	8	2	8	3	36
251	0	0	0	0	0	2	6	1	6	1	16
253	0	3	0	5	0	3	8	2	8	3	32
Grand Total	0	33	0	37	0	30	92	23	92	26	333

Table A-11: Effects of Interactions between Prospective hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (FLUSH/T1).

A3 Drawdown	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
11a	0	6	0	6	0	12	12	9	12	10	67
11b	0	4	0	4	0	2	12	1	12	1	36
12a	0	6	0	6	0	12	12	12	12	12	72
12b	0	6	0	6	0	4	12	1	12	3	44
21a	0	4	0	4	0	0	12	0	12	0	32
21b	0	1	0	2	0	0	10	0	10	0	23
22a	0	4	0	5	0	0	12	0	12	0	33
22b	0	2	0	4	0	0	10	0	10	0	26
Grand Total	0	33	0	37	0	30	92	23	92	26	333

Table A-12: Effects of Interactions between passage hypotheses and drawdown hypotheses on the ability of A3 to meet all jeopardy standards (FLUSH/T1).

A3 Drawdown	Passage Hypotheses												Grand Total
	111	113	141	143	151	153	211	213	241	243	251	253	
11a	5	7	5	7	5	7	3	7	4	7	3	7	67
11b	2	4	2	7	2	5	2	2	2	4	2	2	36
12a	5	7	5	7	5	7	5	7	5	7	5	7	72
12b	2	5	2	7	2	4	2	4	2	6	2	6	44
21a	2	4	2	4	2	4	2	2	2	4	2	2	32
21b	2	2	2	4	2	3	0	2	2	2	0	2	23
22a	2	4	2	4	2	4	2	2	2	4	2	3	33
22b	2	4	2	4	2	2	0	3	2	2	0	3	26
Grand Total	22	37	22	44	22	36	16	29	21	36	16	32	333

A.2 Effects of Interactions Among Hypotheses on Relative Ranking of Actions

Section 3.4 looked at the effects of individual hypotheses on the relative ranking of actions. Again, this analysis is useful for identifying single factors, but it does not say anything about the effects of interactions between the individual hypotheses. To test for potential interactions among factors, we constructed “interaction tables” similar to those in Tables A-3 to A-12. Instead of looking at runs in which all of the standards were met, here we look at the 106 CRiSP/T4 runs in which A1 or A2 ranked higher than A3 with the 24-year survival standard. Table A-13 shows interactions between prospective and passage assumptions, Table A-14 shows interactions between prospective and drawdown assumptions, and Table A-15 shows interactions between passage and drawdown assumptions.

Overall, the results suggest that there are no combinations of hypotheses that dominate the results. Some general patterns from the tables:

- a long transition period (TJUVb) is an important factor, and tends to favor A1/A2 under CRiSP regardless of the other passage and prospective hypotheses with which it is combined
- Alpha model runs make up a greater proportion of runs that favor A1/A2 than Delta model runs, regardless of other passage hypotheses
- The effects of TURB5 on A1/A2 vs. A3 are stronger when coupled with FGE1
- The effects of regime shift extra mortality on runs that favor A1/A2 do not depend on passage assumptions

Table A-13: Effects of interactions between Prospective hypotheses and passage hypotheses on ranking of A1/A2 vs. A3 for CRiSP/T4 and the 24-year survival standard.

Passage Hypotheses	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
111	2		2		3						7
113						1	1	1	1	1	5
141	1	1	1	1	2	1				1	8
143		3	1	1							5
151	2	2	2	1	4	2	2	2	2	6	25
153	2	3	3	2	3		2	1	2	5	23
211	2	2	2	2	3	1	1	1	1	2	17
213											0
241	2	1	1	1	5						10
243	1				2						3
251	1		1		1						3
253											0
Grand Total	13	12	13	8	23	5	6	5	6	15	106

Table A-14: Effects of interactions between Prospective hypotheses and drawdown hypotheses on ranking of A1/A2 vs. A3 for CRiSP/T4 and the 24-year survival standard

Drawdown Hypotheses	Prospective Hypotheses										Grand Total
	A1	A3	A5	A7	A9	D1	D3	D5	D7	D9	
11a					1						1
11b	5	3	5	2	6	1	1	1	1	4	29
12a											0
12b					1					1	2
21a		1			1					2	4
21b	4	6	4	4	6	4	4	3	4	4	43
22a					1					2	3
22b	4	2	4	2	7		1	1	1	2	24
Grand Total	13	12	13	8	23	5	6	5	6	15	106

Table A-15: Effects of interactions between passage hypotheses and drawdown hypotheses on ranking of A1/A2 vs. A3 for CRiSP/T4 and the 24-year survival standard

Drawdown Hypotheses	Passage Hypotheses										Grand Total
	111	113	141	143	151	153	211	241	243	251	
11a								1			1
11b	3		4		9	6	6	1			29
12a											0
12b					1			1			2
21a				1	2	1					4
21b		5	4	1	10	7	10	5	1		43
22a	1				1	1					3
22b	3			3	2	8	1	2	2	3	24
Grand Total	7	5	8	5	25	23	17	10	3	3	106

Appendix B: Changes since the March 1998 Preliminary Decision Analysis Report

A number of significant changes in hypotheses and model results have occurred since the Preliminary Decision Analysis Report for Spring/Summer Chinook was completed in March 1998. A number of new hypotheses were received after May 22 (described in Section 5), but are not reflected in these results because most of these new hypotheses have not yet been quantitatively defined sufficiently to allow incorporation into the models.

Changes since the Preliminary Decision Analysis Report

- CRiSP D values (ratio of post-Bonneville survival of transported fish: non-transported fish) have changed based on Anderson 1998. Revised D values are lower than those used in the preliminary report (Fig. B-1). We have designated the revised set of D values the “TRANS4” hypothesis to distinguish them from the earlier set (which was designated TRANS3).

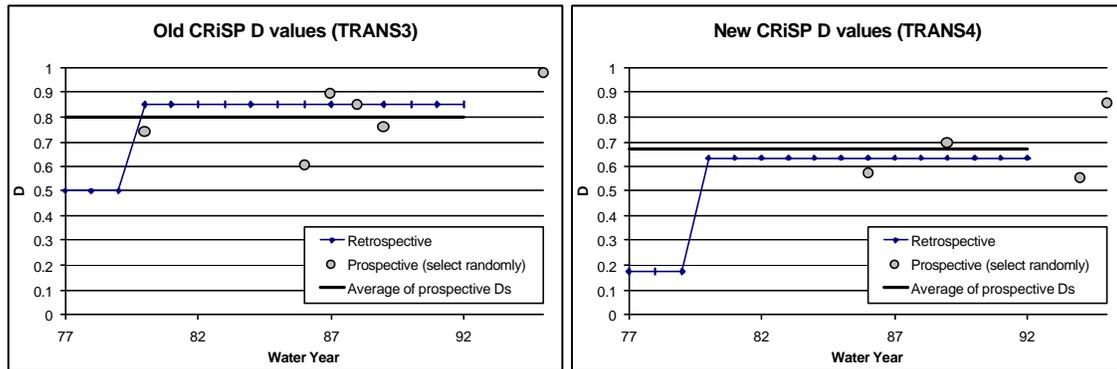


Figure B-1: Old (TRANS3) and new (TRANS4) transportation assumptions used with CRiSP.

- CRiSP in-river survival rates of non-transported fish (V_n) under A3 have changed. Previous CRiSP output was based on different travel time assumptions. New CRiSP A3 V_n 's are higher than those used in preliminary report, and are similar to A3 V_n 's projected by FLUSH (Fig. B-2).

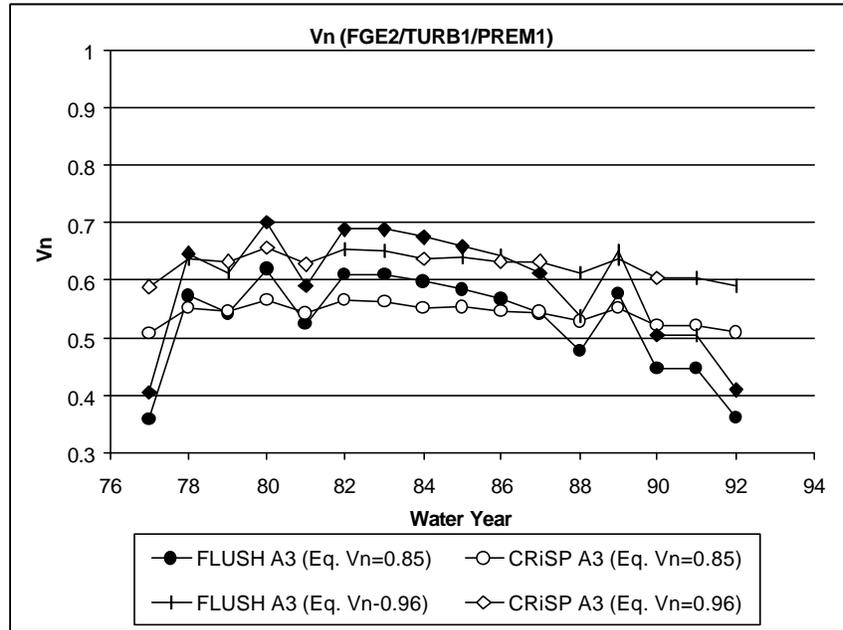


Figure B-2: FLUSH and CRiSP (revised) in-river survival under A3, for both the high (0.96) and low (0.85) hypothesized values of equilibrated juvenile survival rates through the drawdown reach.

3. Uses corrected set of Astoria flows. Previous results used the wrong values.
4. Addresses imbalanced design and associated problems (described by Hinrichsen and Paulsen 1998) by:
 - getting a complete set of passage model runs from CRiSP.
 - omitting FLUSH TURB6 and TRANS2 runs (these hypotheses were not run by CRiSP). Sensitivity analysis on TURB6 results show that results are intermediate to those of TURB1, 4, and 5, so effects of the TURB6 hypothesis are subsumed by results for the others. Sensitivity analysis on TRANS2 shows that TRANS2 (which results in lower D values than TRANS1) tends to decrease the jeopardy probabilities under A1 and A2 relative to A3 (Appendix C). Because A3 is favored by virtually all FLUSH runs, omitting TRANS2 will have no effect on the relative ranking of actions.
 - making it clear that we are not assuming that all 3 of the extra mortality hypotheses are weighted equally, nor are the 2 future climate. Because not all combinations of these two uncertainties were run, such equal weightings are impossible.
5. We have omitted HAB B hypothesis (effects of maximum protection / enhancement) because of problems with implementation in the preliminary draft. Results shown here are based only on HAB 0 (no change to Ricker a).

Effects of Changes on Overall Results

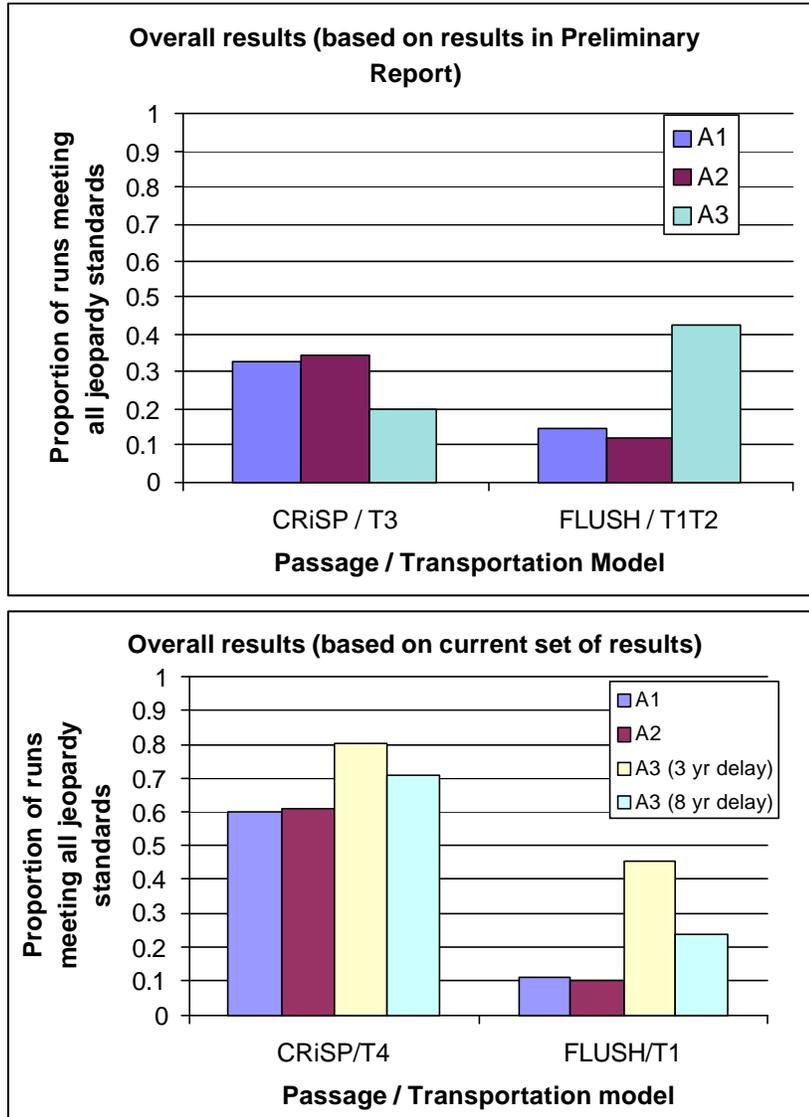


Figure B-3: Proportion of runs meeting all standards, by passage/transportation model. Top panel – results from Preliminary Decision Analysis Report. Bottom: Results incorporating changes listed above.

Effects of Changes in CRiSP D values and A3 In-river Survival Estimates

Jeopardy probabilities for CRiSP changed substantially from the Preliminary Report (Figure B-3). Changes in A3 results were due to both the change in D values and the changes in A3 Vn (Figure B-4). The purpose of this section is to explore the effects of changes in D’s. Because the D values are critical factors in determining model results, we need to understand how changes in these values have affected the current set of results.

Old (TRANS3) and new (TRANS4) CRiSP D values are compared in Figure B-1. Both pre-1980 and post-1980 D values declined from the previous set of results, although the pre-1980 values declined more.

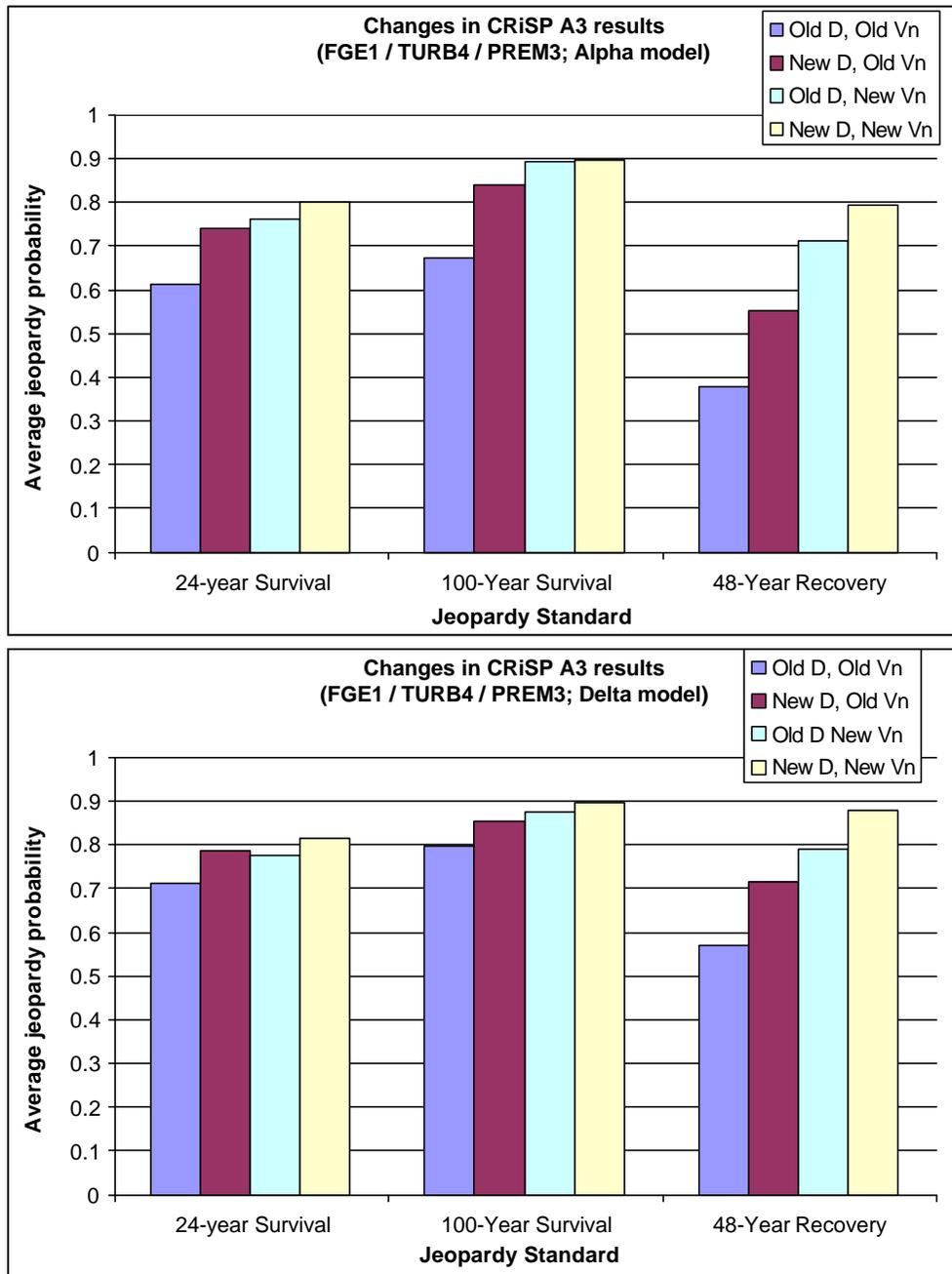


Figure B-4: Changes in CRiSP A3 results due to changes in D and Vn values. Probabilities shown are for a single passage model run, averaged over the 5 extra mortality/future climate hypotheses.

Changes in D's affect life-cycle modeling through changes in the ratio of prospective:retrospective system survivals. Because the M and P components of system survival have not changed from the preliminary report, changes in the system survival ratio are entirely due to the different set of D values. Therefore, it is instructive to look at the ratio of prospective:retrospective D (Figure B-5). The ratio of the

new D values is higher than the ratio of the old values in all water years. The increase is particularly significant in the pre-1980 period, because the reduction in the pre-1980 D values from the old set to the new set was much greater than the reduction in the post-1980 values (Figure B-1). The ratio of prospective:retrospective D values for post-1980 is slightly greater than 1 because the average of the 4 post-1980 D values (used for prospective analyses) is slightly larger than the median (used for retrospective analyses).

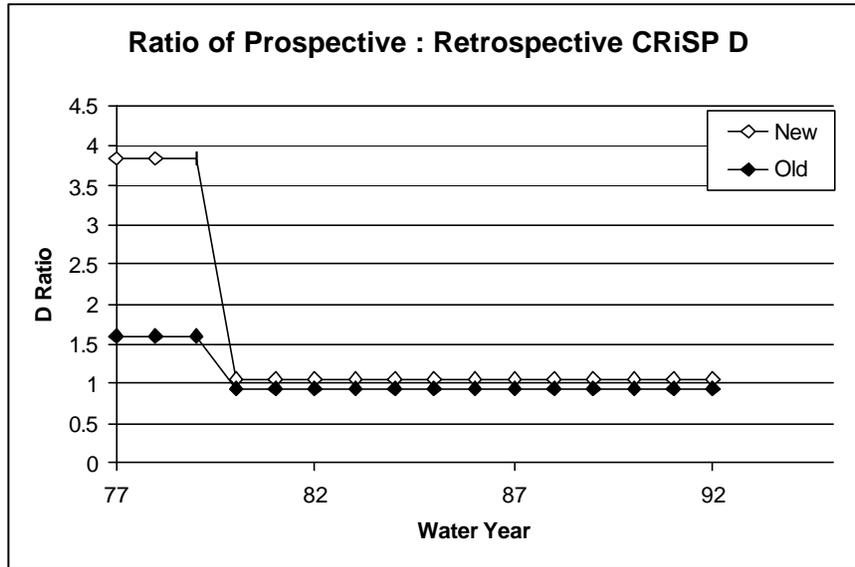


Figure B-5: Ratio of prospective:retrospective CRiSP/T4 D values.

This pattern in the D values is repeated in the ratio of prospective:retrospective system survivals. Figure B-6 shows this ratio as it would be implemented in the Delta model (i.e., \ln prospective system survival : retrospective system survival). This ratio is higher using the new set of D values than with the older set, particularly in the pre-1980 period. For example, in 1977 the new \ln ratio is almost 1.5 times larger than the old ratio; this equates to a 2.2-fold increase in system survivals in non-logarithmic terms. The ratio of prospective:retrospective system survivals for A3 appears to be more sensitive to changes in D than A1 or A2. The differences in the ratio of A3 system survivals is entirely due to the changes in retrospective D values, since there is no transportation (i.e., prospective system survivals for A3 were not affected by the change in D).

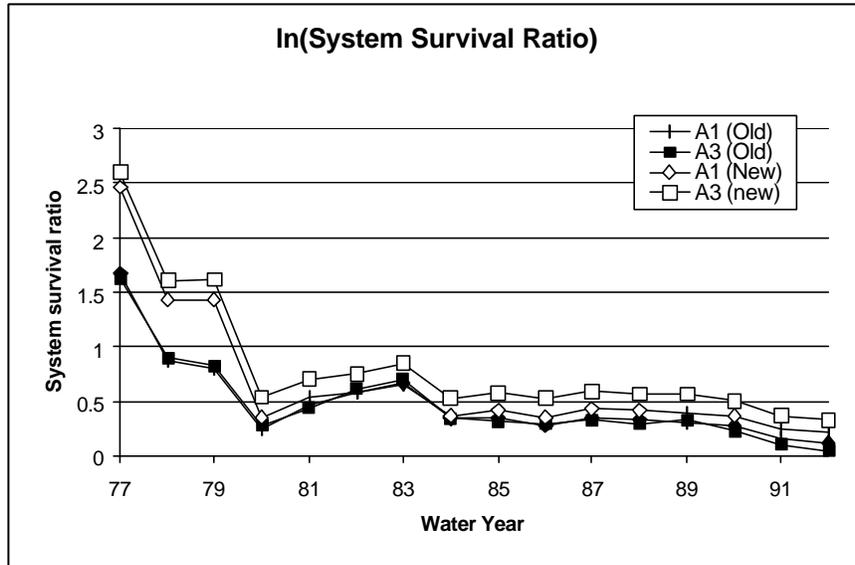


Figure B-6: Ln(pro prospective:retrospective system survival) using Old and New CRiSP/T4 values (only A1 and A3 are shown; results for A2 are virtually identical to A1).

Water years from 1977 to 1992 are selected based on the frequency of unregulated water transit times in the 1929-1992 historical data set (Figure B-7). Using the probabilities of selecting water years, we calculated weighted averages of ln(system survival ratios) over all water years (Figure B-8). Changes in D resulted in an increase in this ratio by about 34% for A1 and A2. These results explain the large increases in the proportion of runs meeting all of the standards under A1 and A2 (Figure B-3). The increase in ln(system survival ratio) was about 69% for A3, although the change in A3 Vn values also contributed (Figure B-4).

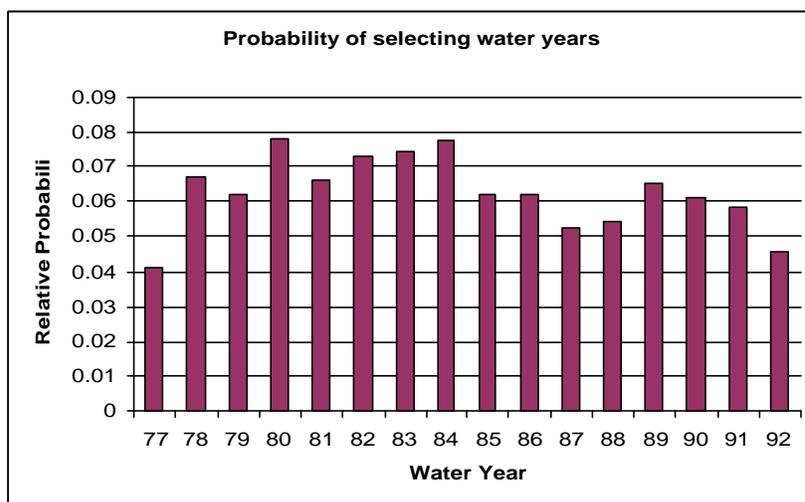


Figure B-7: Probabilities of selecting a water year from 1977-1992 for prospective analysis.

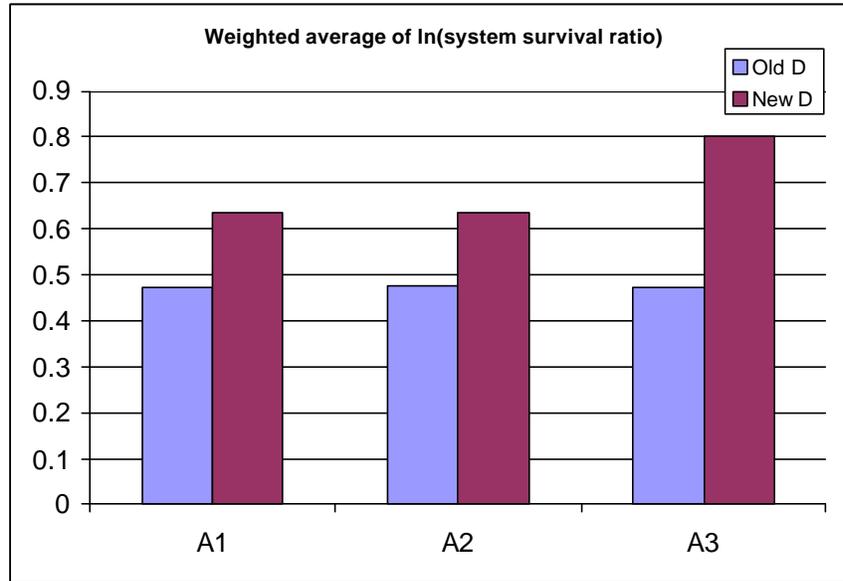


Figure B-8: Weighted average $\ln(\text{prospective system survival}:\text{retrospective system survival})$.

The contribution of the old vs. new difference in $\ln(\text{system survival ratio})$ in each water year to the difference in weighted average over all water years is shown in Figure B-9. Differences in the pre-1980 D values are the most significant factor in the weighted average $\ln(\text{system survival ratio})$ for A1 and A2, contributing over 60% of the overall difference. Differences in weighted average $\ln(\text{system survival ratio})$ for A3 are due to both pre-1980 D values and changes in V_n . The pre-1980 water years are not selected particularly often (probability of selecting 1977, 1978, or 1979 is only 0.17; Fig. B-7). However, there are large increases in the ratio of system survivals in those years, primarily because the retrospective D values in those years are much lower than they were.

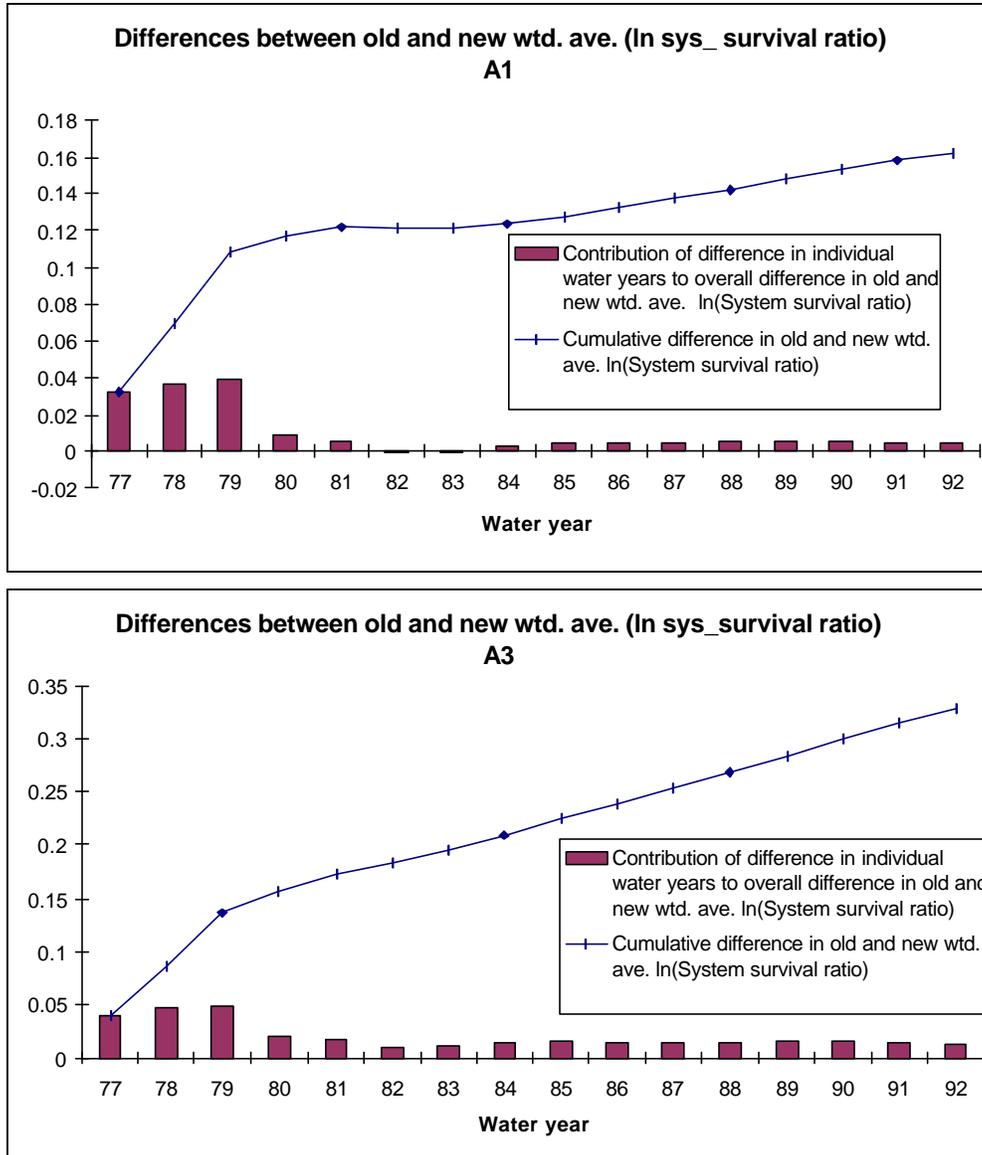


Figure B-9: Contribution of differences in individual water years to difference in weighted average ln(system survival ratio) for A1 (top) and A3 (bottom) over all water years.

Appendix C: Sensitivity Analysis of FLUSH TRANS 2 Transportation Model

The Transport:Control ratios (TCR) used in the FLUSH/TRANS1 and CRiSP/TRANS4 transportation models are derived from transport studies conducted at the mainstem dams. The FLUSH/TRANS2 transportation model adjusts these TCRs downward to reflect mortality incurred between the study dams and spawning areas (see Preliminary Decision Analysis Report, Appendix A p. 79 for details).

Lower TCRs equate to lower values for D, the ratio of post-Bonneville survival of transported fish : post-Bonneville survival of non-transported fish. Thus one would expect that the jeopardy probabilities for those actions that emphasized transportation (i.e., A1 and A2) would be lower under the TRANS2 model than they would under the TRANS1 model.

We ran a limited set of 40 model runs with the TRANS2 model and compared them to the equivalent runs with the TRANS1 model. Results of these runs confirms that average jeopardy probabilities (averaged over extra mortality/future climate hypotheses and life-cycle models) under A1 and A2 are lower with the TRANS2 model (Fig. C-1). For A1, the difference in average jeopardy probability is 0.02; for A2 the average probability with the TRANS2 model is 0.05 lower. The alternative transportation models have virtually no effect on average probabilities under A3.

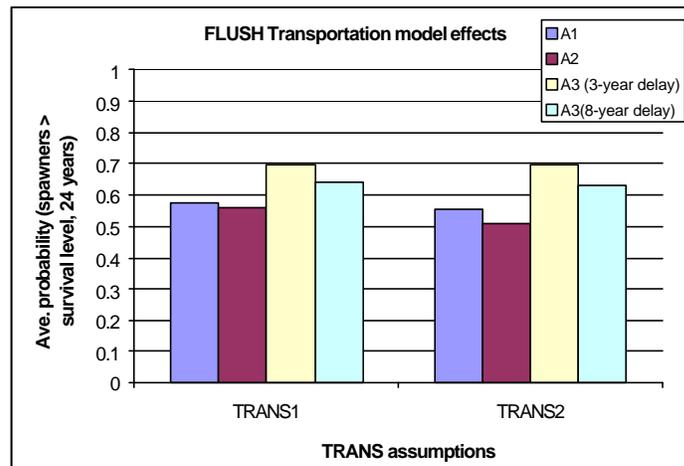


Figure C-1: Effects of FLUSH Transportation models on proportion of runs meeting 24-year survival standard.

Based on these results, the alternative TRANS models appear to have small effects on jeopardy probabilities relative to the effects of other key uncertainties. Overall, TRANS2 results would tend to result in slightly lower abilities of the actions, particularly A1 and A2, to achieve the jeopardy standards, but would have no effect on the relative ranking of actions.

Appendix D: Additional Diagnostic Analyses of passage / transportation models

D.1 Introduction and Definition of Terms

Section 3 concluded that the passage models and their associated transportation models are important in determining both the relative rank of actions and the ability of those actions to meet the jeopardy standards. Because these models are themselves complexes of hypotheses and sub-models, we have attempted to determine the critical components of the models that explain why the models produce different results. It is particularly important to separate out the effects of the passage and transportation models. These two components are linked, because the estimates of transportation effects are based partly on passage model estimates of juvenile survival, as explained below, and partly on results of transportation studies, which are independent of passage models.

The following observations identify some specific areas where passage mortality and transportation assumptions differ the most, based on our analyses of the passage model input files to the life cycle model. The passage model input files provide estimates of the following values:

V_n = survival rate of non-transported fish from head of LGR to below BON

P_{bt} = proportion of smolts below BON that were transported

D = ratio of post-Bonneville survival of transported: non-transported fish. D values are estimated with the equation:

$$D = T:C * \frac{\text{Survival of control fish through migration corridor}}{\text{Survival of transported fish through migration corridor}}$$

Where T:C is the ratio of returns of transported and control (non-transported) fish from transportation studies, and the survival of control fish through the migration corridor is estimated from the passage models, and is dependent on various passage assumptions.

Retrospectively, CRiSP uses a constant D of 0.174 for pre-1980 water years (median of D estimated in pre-1980 transportation studies), and a constant D of 0.633 for 1980-1992 (median of D estimated from post-1980 studies). CRiSP prospective D values are selected randomly from D 's estimated in 1986, 1989, 1994, and 1995 transportation studies. The CRiSP prospective system survivals shown below were calculated using the average of these 4 D values. FLUSH calculates a D estimate in every year based on the V_n for that year and an estimated T:C ratio (TCR).

ω = system survival = $\exp(-M)(D*P + 1 - P)$

M = weighted average instantaneous mortality rate of all smolts; a function of V_n , proportion transported from various projects, and V_t , the survival of transported fish in the barge (V_t is assumed to be 0.98 in both passage models)

We looked specifically at water years 1977-1992, because those are the water years that are sampled from in the prospective simulations.

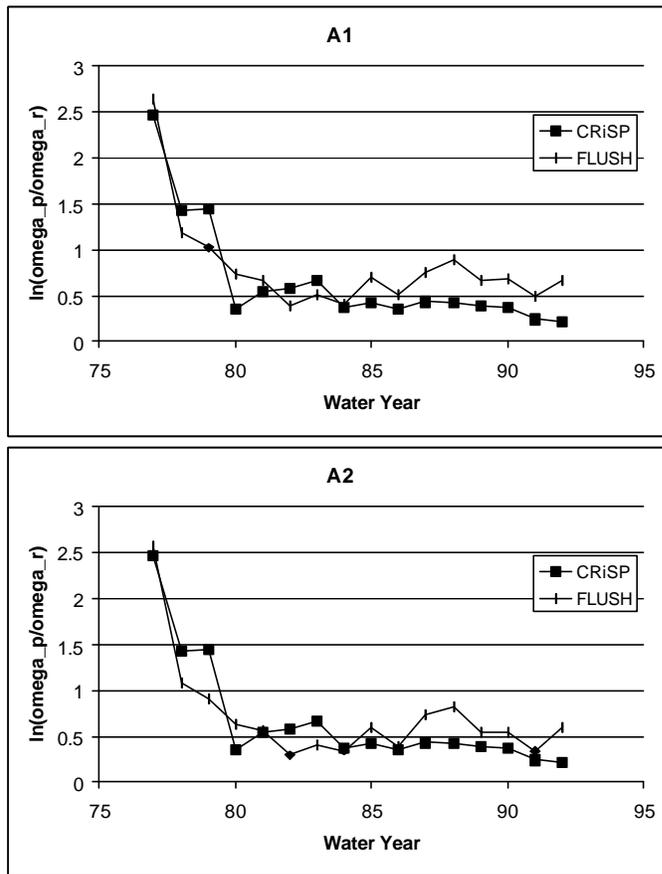
The comparisons below are for a single set of passage assumptions (FGE1, TURB4, PREM3). This set of passage assumptions was chosen because it was one of the sets that was run by both passage models for all three actions. Also, the T:C and D value information provided by Anderson 1998 was for TURB4

assumptions. For A3, we compared equilibrium values of in-river survival (there is no transportation after drawdown occurs in A3, so $P_{bt} = 0$).

D.2 Comparison of $\ln(\text{prospective}:\text{retrospective system survivals})$

Prospective and retrospective values of the above parameters are incorporated into prospective simulations in the following manner (further details are given in Section A.3.2 of the Preliminary Decision Analysis Report). In the delta model, retrospective values are considered by including the ratio of prospective and retrospective system survivals in the generalized Ricker model; in the alpha model, the retrospective $\ln(DP + 1 - P)$ (averaged over water years 1977-1992) is a term in the generalized Ricker model. In addition, for both models the retrospective passage model inputs are used to estimate each stock's Ricker parameters.

The first place to examine differences, then, is in the $\ln(\text{ratio of prospective:retrospective system survival values})$ derived from passage model input (Fig. D-1).



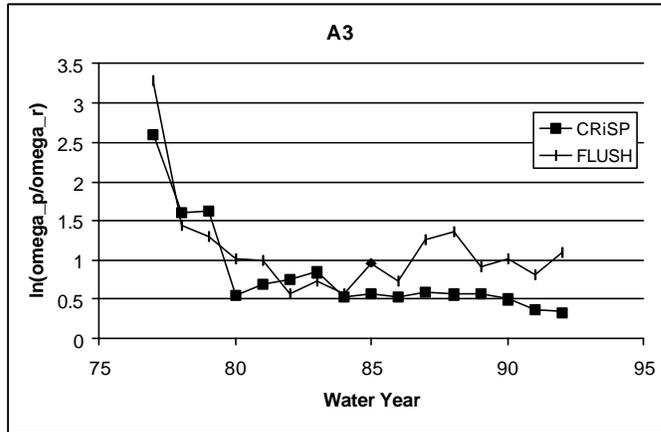


Figure D-1: $\ln(\text{prospective system survival} : \text{retrospective system survival})$ (FGE1, TURB4, PREM3).

To understand the differences in these ratios, the following sections look at differences in retrospective and prospective system survivals, and differences in their components. System survival is the product of $\exp(-M)$ (the weighted average survival rate of all smolts through the passage corridor) and $(DP+1-P)$, which is basically a survival “penalty” adjustment for additional mortality that is assumed to occur for transported fish relative to non-transported fish (Figure D-2). Under FLUSH, there is a larger penalty for transported fish than under CRiSP, so A1 and A2 system survivals are lower. There is no survival penalty under A3 because fish are not transported ($P=0$; $DP+1-P = 1$). Note that $\exp(-M)$ drops under A3 for both models, because the survival of in-river fish in a drawdown system with 4 projects remaining is still lower than the survival of transported fish in the barge.

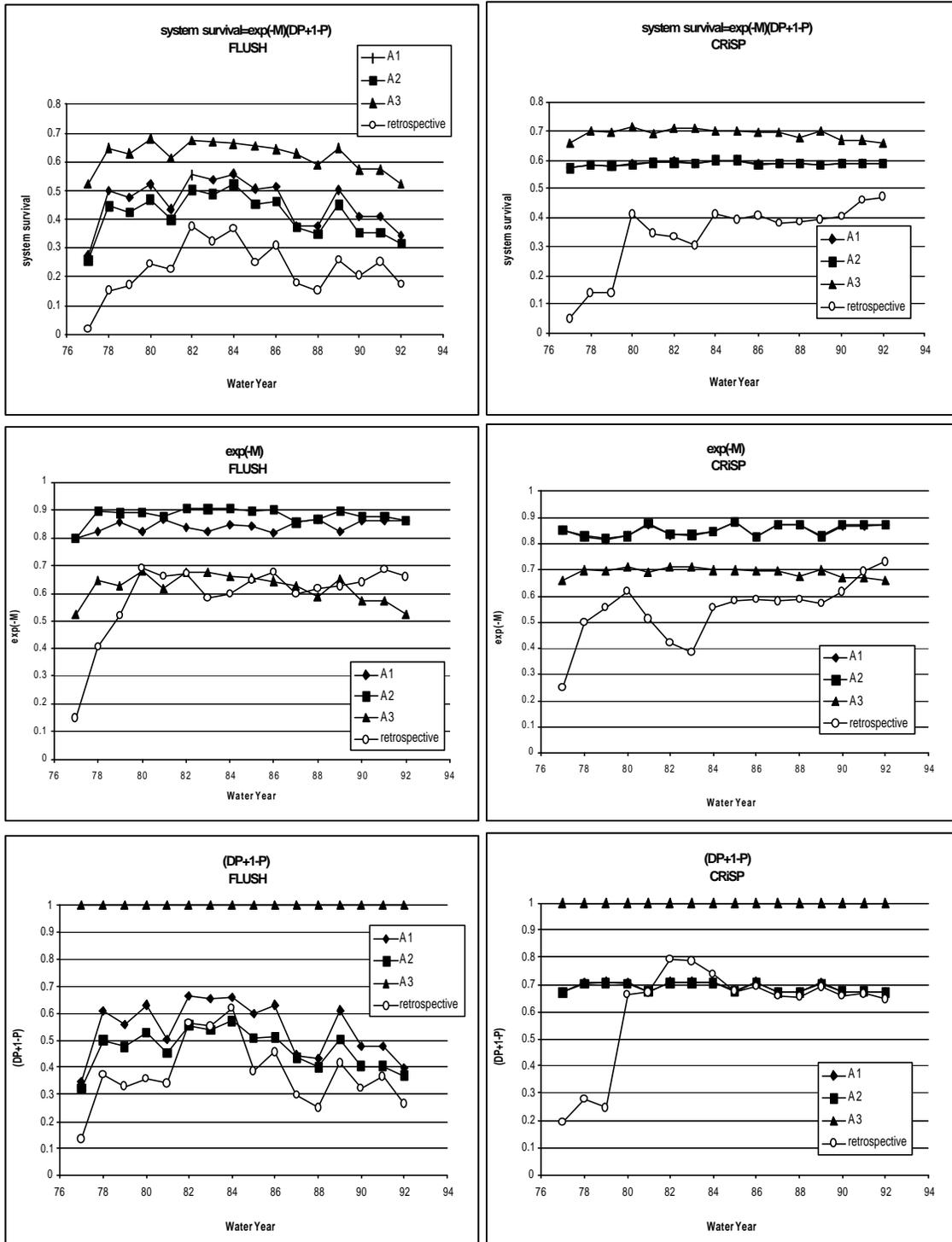


Figure D-2: Retrospective and prospective system survival and its components (FGE1, TURB4, PREM3).

Plotting system survival against water year is useful for comparing the differences between passage/transportation models in trends over time. It is also useful to look at differences in system survival as a function of fish travel time (Fig. D-3). This shows the differences in the assumed

relationships between survival and travel time, as well as the differences in the travel times themselves (see Fig. A.2.1-9 in the Preliminary Decision Analysis Report for a further illustration of this). These differences are possibly due to differences in the structures of the two models (e.g., CRiSP operates on a finer time scale than FLUSH).

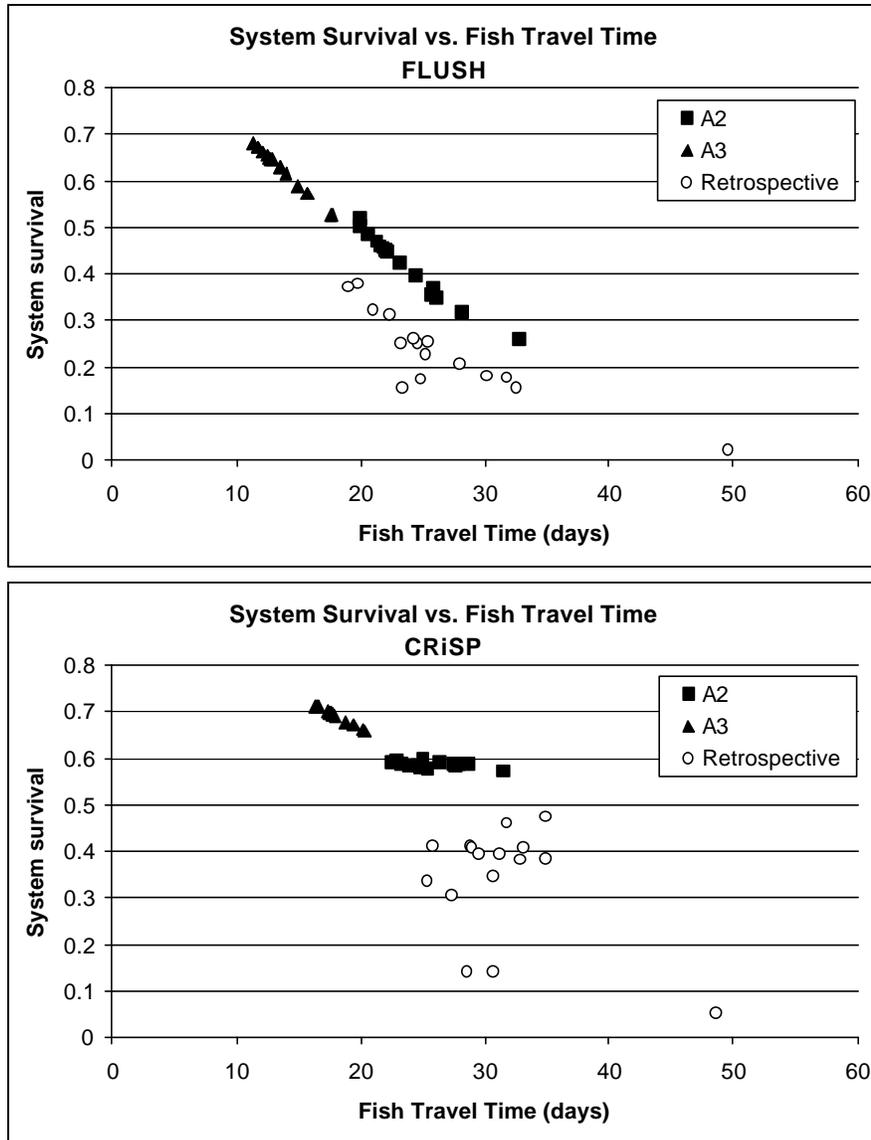


Figure D-3: Retrospective and prospective system survival vs. fish travel time (FGE1, TURB4, PREM3).

D.3 Differences in Retrospective System Survival

System survivals calculated using different M's but the same set of D values are virtually identical, while system survivals using different sets of D values but the same M estimates are markedly different (Figure D-4). Retrospective Pbt, M and Vn are fairly similar between CRiSP and FLUSH after 1977, except for 1981-1983 (Figure D-5). This suggests that different D assumptions by the two models have the most effect on differences in retrospective system survival.

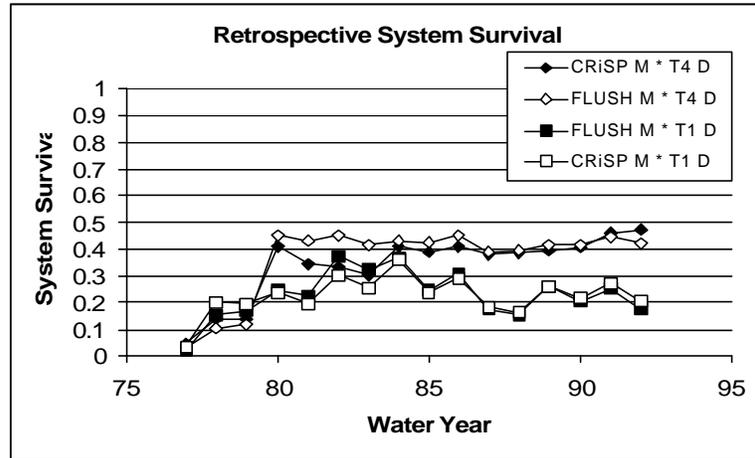


Figure D-4: Sensitivity of system survival to M and D (TURB4).

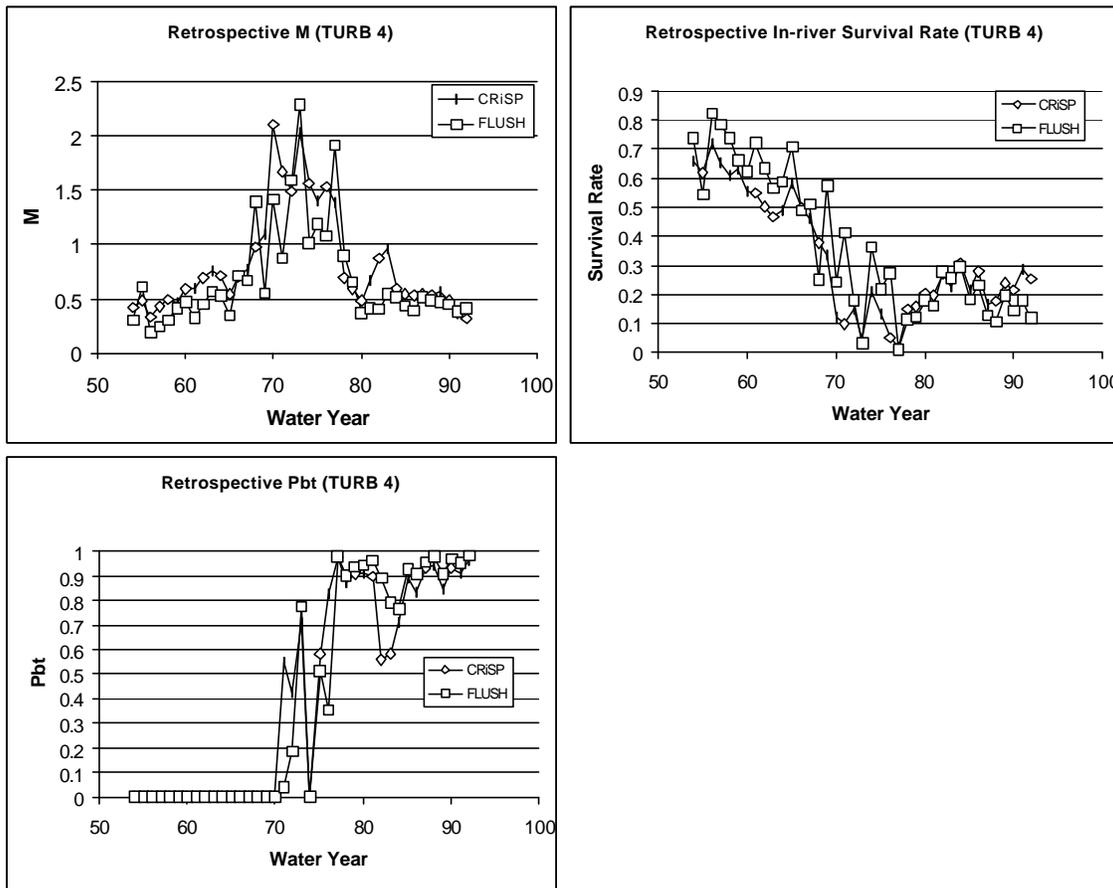


Figure D-5: Retrospective M, in-river survival (Vn), and Pbt.

The sensitivity of system survival to D can also be seen by comparing the CRiSP D and system survival from the March 1998 Preliminary Decision Analysis Report to the newer set of D and system survival values. The lower fraction of fish transported (Pbt) in water years 1981-1983 in CRiSP (relative to

FLUSH) is coincident with higher values of M. Because CRiSP and FLUSH in-river survivals (Vn) are similar in these three years, it is worth checking that similar transport and spill assumptions were made in those years.

D.4 Differences in Prospective A1 and A2 System Survival

M is not an important difference when there is transportation (i.e., in A1 and A2). Prospectively, Vn and Pbt tend to negate one another - when Vn is high (as in the CRiSP runs), the proportion of fish below BON that were transported (Pbt) is by definition relatively lower (Figure D-6). Also, prospective Pbt's are quite high in both models (> 0.85, and around 1 for FLUSH A2), so the weighted average mortality M is dominated by barge mortality, which is the same for both models. Differences in prospective system survival in A1 and A2 are therefore due to differences in D values for these prospective scenarios. This is apparent when CRiSP and FLUSH M and D values are exchanged in the calculation of system survivals (Figure D-7).

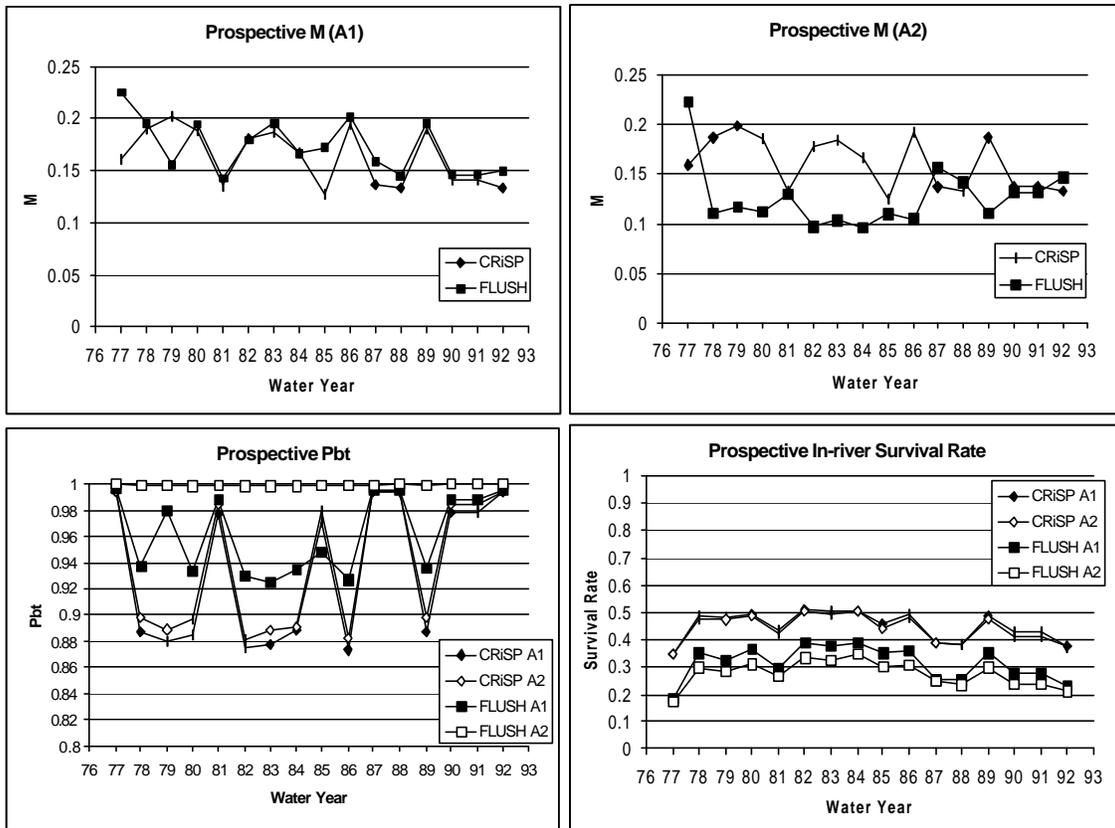


Figure D-6: Prospective (A1 and A2) M, Pbt, and Vn (FGE1, TURB4, PREM3).

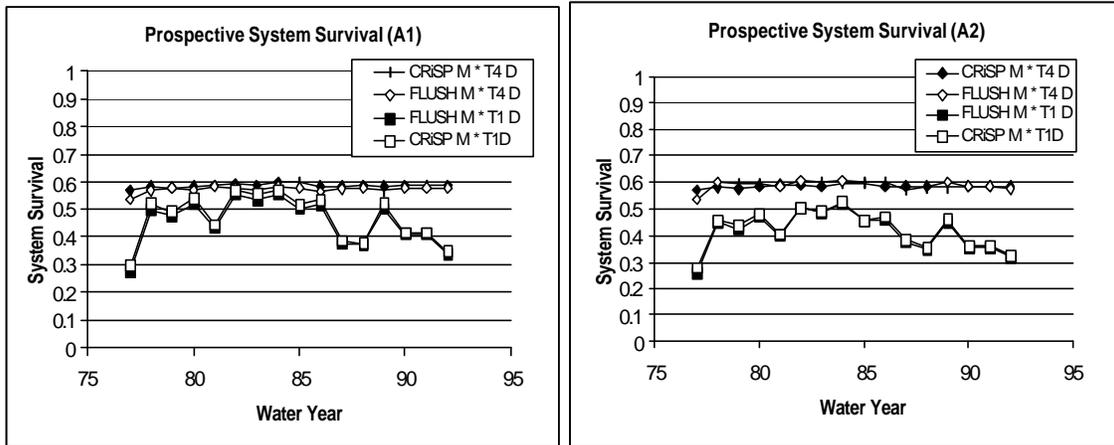


Figure D-7: Sensitivity of prospective system survival to M and D (FGE1, TURB4, PREM3).

D.5 Differences in Prospective A3 System Survival

Differences in prospective M are potentially a more important factor for A3, where there is no transportation after drawdown is completed. Since $P_{Bt} = 0$, all of the differences in M are due to differences in prospective V_n between CRISP and FLUSH. However, these differences are not large (Figure D-8), which leads to similar prospective systems survivals for A3 for the two models (see top panel, Figure D-2).

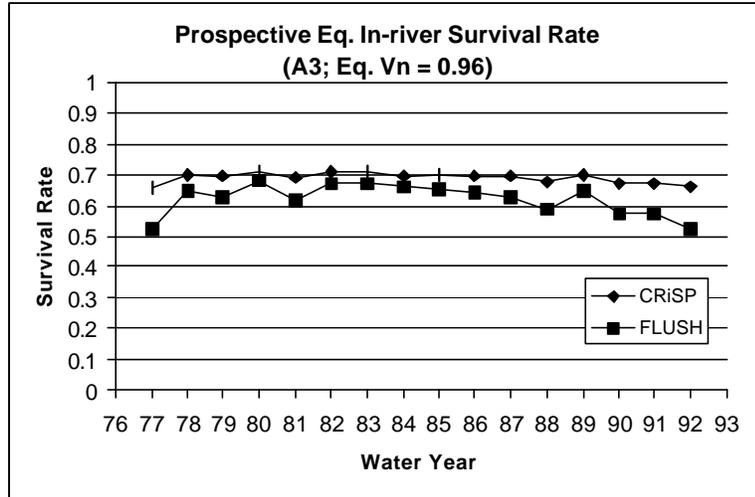


Figure D-8: Prospective (A3) V_n , under the hypothesis that the equilibrated juvenile survival rate through the drawdown reach is 0.96 (FGE1, TURB4, PREM3).

D.6 Differences in System Survival – Summary

Differences in ratios of prospective : retrospective system survival can be attributed to differences in D values associated with the two passage models. Retrospectively, differences in estimates of in-river survival and proportion transported are not large enough to account for differences in system survival. In

A1 and A2 scenarios, almost all fish are transported, so even though there are large differences in in-river survival between the two models, these differences are not the cause of differences in system survival. In A3, where there is no transportation, differences in retrospective D give rise to differences in the ratio of prospective : retrospective system survivals. Prospective Vn estimates, and therefore prospective system survival estimates, are similar in the two models for A3.

The above analysis shows that it is the differences in the transportation component of the passage/transportation models, not the differences in their Vn and Pbt estimates, that produce the differences in observed outputs between CRiSP/T4 and FLUSH/T1 runs.

D.7 Differences in Transportation Component

Differences between CRiSP T4 and FLUSH T1 estimates of D are substantial in most years (Figure D-9). Since these differences are largest in retrospective years (particularly after 1979), the analyses below focus on differences in retrospective D values.

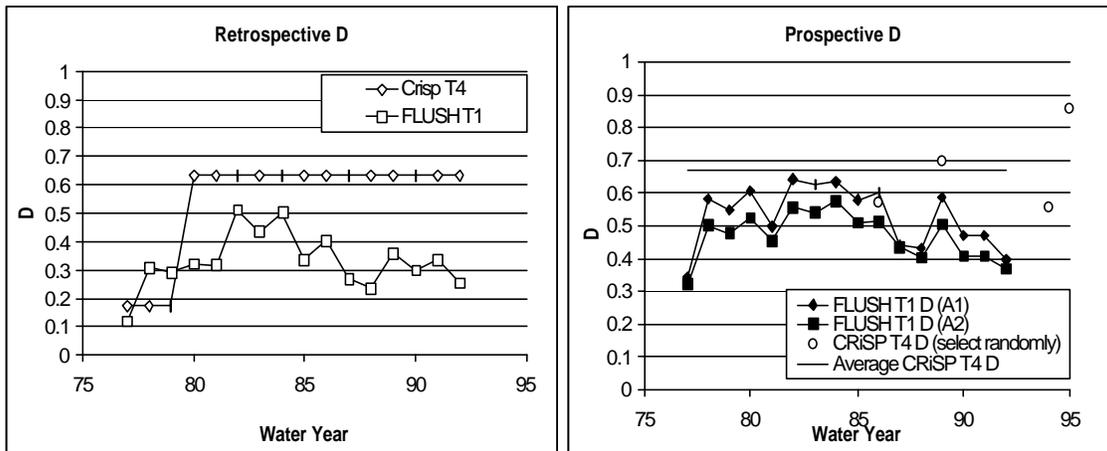


Figure D-9: Retrospective and prospective D's (TURB4).

D values are estimated in both models using the same basic equation:

$$D = T:C * Adj. * \text{Surv}_{\text{control}} / V_t$$

where:

D = ratio of post-Bonneville survival of transported fish : non-transported fish

T:C = Transport: Control ratio, estimated from transport studies.

$\text{Surv}_{\text{control}}$ = survival to below Bonneville Dam of control (non-transported) fish. In FLUSH, this is referred to as "s"; in CRiSP it is called "Vc1".

V_t = survival to below Bonneville Dam of transported fish. This is the product of the survival of fish destined to be transported to the point of collection ($V_{\text{pre-b}}$) X the survival of transported fish in the barge (V_b , assumed to be 0.98). Because we don't have $V_{\text{pre-b}}$ estimates from the passage modelers, we have used $V_t = V_b = 0.98$ in the analyses below. This means that what we refer to as Vc is actually the ratio of $V_c:V_{\text{pre-b}}$. Since most fish were collected at either

LGR or LGO in the transportation studies, CRiSP and FLUSH V-pre-b estimates should be reasonably similar in most years. The important point is that both Vc and Vpre-b depend on assumptions about survival of non-transported fish through the hydrosystem. Therefore, differences in the term Vc:Vpre-b are due to differences in these assumptions.

Adj. = adjustments to account for various treatment effects on control (non-transported) fish used in the transport studies.

The two transportation models use different T:C's, Vc's, and adjustments, and have different approaches to dealing with multiple transport studies within years and interpolating estimates of D to years when there were no transportation studies. Table D-1 summarizes these differences:

Table D-1: Differences between T1 (FLUSH) and T4 (CRiSP) transportation models.

Assumption	FLUSH/T1 (from information in Table A.3.1-1 and A.3.1-2)	CRISP/T4 (from information in Table 1 of Anderson 1998)
1. Transportation studies used	LGO, 1971-1989	LGR, 1968-1995
Use different raw T:C estimates from 1978 LGR studies		
2. Adjustment for mortality of control fish passing through project twice in certain years	Adjusted in 71, 72, 73, 75, 76, 78 (one group only)	Adjusted in 72, 73, 75, 76, 78
But passage models have different estimates of dam mortality		
3. Adjustment for control fish that were transported at lower projects	No	Yes – 75, 76, 79, 86, 89
4. Within year aggregation of T:C estimates	Weighted geometric mean, weighted by variance of individual studies	No
5. Method for filling in retrospective years with no transportation study	Fit T:C vs. s function to 1971-1989 TC data (s = survival of control fish from tailrace of transport project)	Calculate separate D for each study and each year; use 1968-1979 median D for 1977-1979 retrospective water years; 1986-1995 median D for 1980-1992 retrospective water years.
6. Method for projecting D in prospective scenarios	Project s using passage model, calculate T:C from T:C vs. s function, calculate D from T:C	Randomly select D values estimated from 4 post-1980 T:C studies
7. Estimates of in-river survival of control fish	FLUSH: "s" CRISP: Vcl	FLUSH – uses per-project expansion of control survival estimates from NMFS survival studies for 1970's; passage model for 86 and 89 (i.e., if study reach was 6 projects and full reach was 7 projects, control survival = observed survival ^(7/6)). CRISP – uses passage model for all years

Transportation studies used

FLUSH T1 uses all transportation studies conducted at Little Goose (LGO) and Lower Granite (LGR) dams between 1971 and 1989. CRiSP also uses these studies plus earlier studies conducted at Ice Harbor dam (IHR) in 1968-1970, and studies at LGR in 1994 and 1995. Unadjusted T:C's are identical in the studies used by both models, except for the three LGR studies in 1978 (Figure D-10). It's not clear why these raw, unadjusted T:C's are different for those years.

Adjustments for treatment effects on control fish

Adjusted T:C's for FLUSH are listed in Table A.3.1-2 of the Preliminary Decision Analysis Report. In CRiSP, the adjustments were integrated into the equation for D (eq. [2] of Anderson 1998) and were not explicitly applied to T:C's. However, we wanted to isolate these adjustments and apply them to the T:C ratios as was done in FLUSH so that we could compare the raw Vn's directly for CRiSP and FLUSH.

Adjustments for CRiSP were calculated from equation [2] and [3] in Anderson 1998. Based on these equations, the overall adjustment that was applied to CRiSP T:C's was:

$$\text{Adj. T:C} = \text{Raw T:C} * (1-f+Df) * \exp(-(0.0058*T)x)$$

where T and x are defined as in Anderson 1998.

Differences in adjusted observed T:C's are relatively minor after 1980 (Figure D-12). Differences are greatest in 1978 and 1979. Note, however, that the effect of the differences in observed T:C in 1978 and 1979 is dampened by estimating T:C's from the T:C vs. s function in FLUSH, and by calculating a median D in CRiSP. After these two smoothing algorithms are implemented, the difference in the actual D values in 1978 is much smaller than the difference in the adjusted T:C would suggest (left panel, Figure D-9).

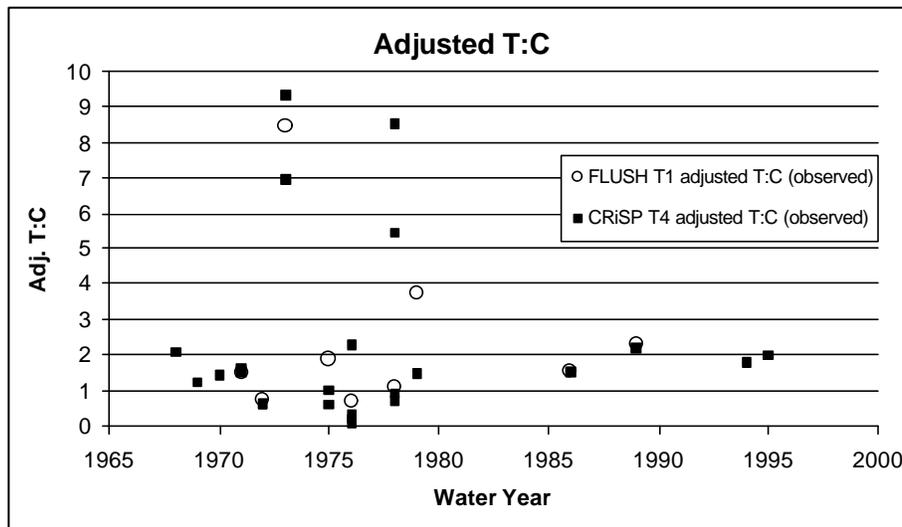


Figure D-12: Adjusted T:C's.

Within-year aggregation

Differences in within-year aggregation methods (i.e., weighted geometric mean in FLUSH vs. no aggregation in CRiSP) have only a small effect on differences in pre-1980 D values (there were no years with multiple studies after 1978). To check this, we calculated the weighted geometric mean of CRiSP adjusted T:C's in each year with multiple studies using the same method used for FLUSH, then used the CRiSP-estimated Vcl for that year to calculate a single estimate of D in each year with a transportation study. The pre-1980 median using these weighted geometric mean annual estimates was 0.211, which is slightly higher than the median of 0.174 when studies were kept separate.

Methods for estimating D in years with no transportation studies

We explored the effects of the different methods used to fill in years with no T:C studies by using the CRiSP method with FLUSH D data. Instead of aggregating T:C studies within a year (by calculating a weighted geometric mean), we kept each T:C study separate, and adjusted them using the same adjustment factor as was applied to the weighted geom. mean of T:C studies. Then, using the FLUSH “s” for that year, we calculated a D for each study. The pre-1980 median of these D values is 0.55; the post-1980 median is 0.49 (Figure D-13). These medians were used retrospectively in BSM. Prospectively, we mimicked the CRiSP method by taking an average of the 1986 and 1989 D values, and applying that average in each prospective year (Figure D-14). The point here is that although there are differences between the two transportation models in the method of estimating D values for years with no transportation studies, these differences are not large enough to account for the differences in life-cycle model results from the passage/transportation models.

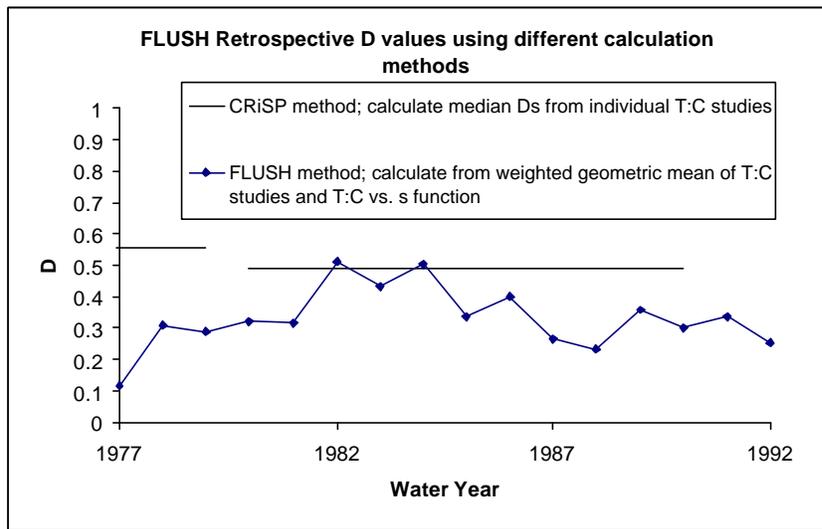


Figure D-13: Comparison of two methods for estimating retrospective D in years with no transportation studies.

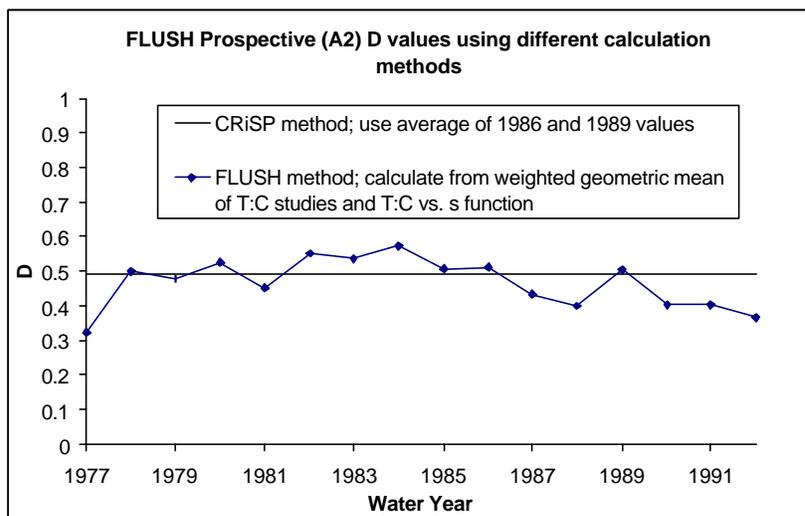


Figure D-14: Effects of using two methods for projecting prospective D values (only A2 is shown).

Jeopardy standards using the CRiSP method (“average D”) to calculate FLUSH D’s were not that different than when the FLUSH method (“base case”) was used (Figure D-15).

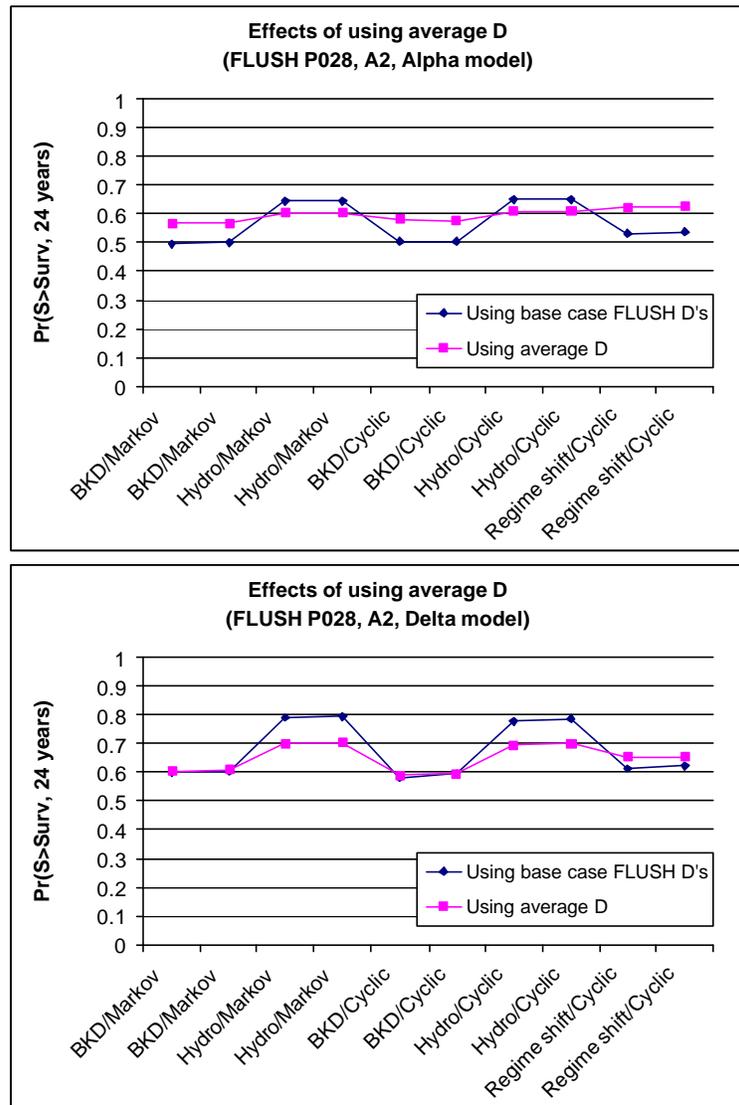


Figure D-15: Effects of two different methods for calculating FLUSH D’s on 24-year survival probabilities.

It is difficult to calculate D values by applying the FLUSH method to CRiSP because CRiSP doesn’t estimate survival of control fish prospectively. However, the effects of the methods on the T:C’s that are used to calculate D can be simulated by calculating a weighted geometric mean of CRiSP T:C’s for each retrospective year, then fitting the same functional relationship as the FLUSH T:C vs. s function to the CRiSP weighted geometric mean T:C’s and Vcl values. The resulting relationship is very similar to that of FLUSH, generating similar T:C’s (Figure D-16). Again, this demonstrates that the differences between the two transportation models in the method of estimating D values for years with no transportation studies are not large enough to account for the differences in overall results.

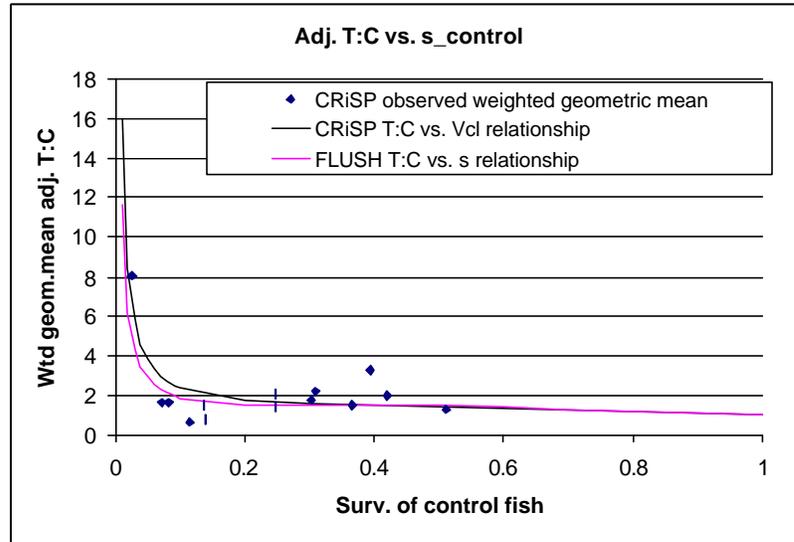


Figure D-16: Weighted geometric mean T:C vs. survival of control fish function.

In-river survival of control fish

The above observations suggest that differences in T:Cs and associated adjustments are not the cause of differences in retrospective D values. Therefore, differences in D's are mostly due to different in-river survival estimates for control fish that are combined with T:C ratios (FLUSH "s", CRiSP "Vcl") (Figure D-17). FLUSH "s" values were estimated from TCRs provided in the retrospective diagnostics and retrospective D values. CRiSP Vcls are from Table 1 in Anderson 1998. We note again that these values are actually the ratio of the survival of control fish to the survival of fish destined to be transported to the point of collection. Therefore, not all of the differences in Figure D-17 are due to survival of control fish – differences in Vpre-b will also contribute to the differences between s and Vcl.

Because CRiSP in-river survival estimates are generally much greater than FLUSH after 1978, this leads to greater D values estimated by CRiSP. Survival rates of control fish are generally higher than the survival of in-river fish in general (Vn, Figure D-18), and the general patterns in survival of control fish are also apparent in Vn – FLUSH estimates are generally higher than CRiSP estimates until around 1980. However, the difference between s and Vcl is larger than the difference between Vns, especially between 1970 and 1976.

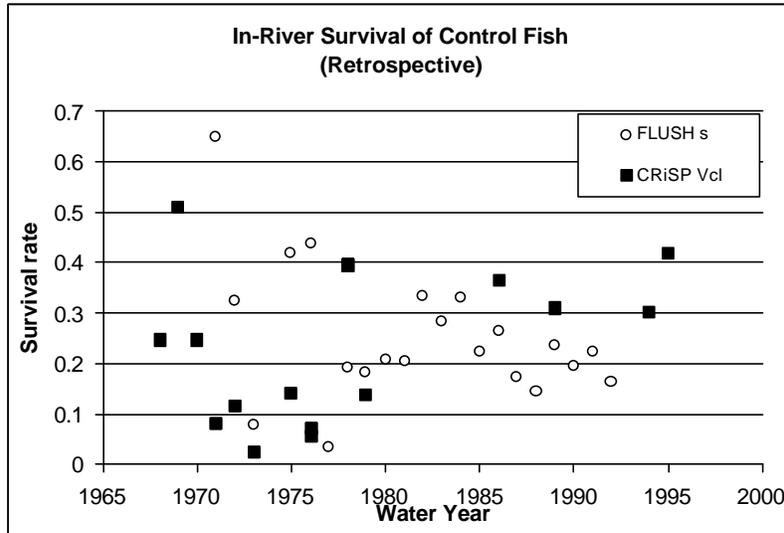


Figure D-17: In-river survival of control fish used to calculate D values.

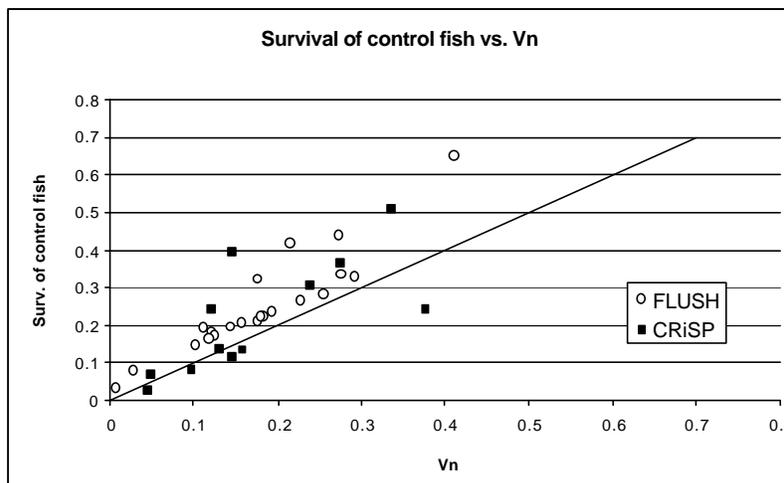


Figure D-18: In-survival of control fish vs. Vn

Appendix E: Using Categorical Regression Trees to Identify Key Uncertainties

PATH modeling results were subjected to a regression tree analysis to determine which factors had the greatest influence on results. The following description of regression trees is extracted from Watters and Deriso (1997):

We fitted regression trees to the example data using the RPART library of functions for S-Plus (Therneau and Atkinson 1997). The RPART library implements many of the concepts developed by Breiman et. al (1984).

We summarize the basics behind fitting a tree model in the following paragraphs. Readers interested in additional detail are referred to Breiman et. al (1984), Therneau and Atkinson (1997), and Venables and Ripley (1998). The basic idea behind fitting a tree-based model can be described as follows:

1. consider all of the predictor variables and find the single predictor variable that best splits the response into two groups (“best” will be defined later),
2. separate the data into two subgroups,
3. repeat Step 1 for each subgroup,
4. continue to find splits and separate subgroups until no improvement can be made (“improvement” will also be defined later),
5. determine whether a tree with fewer subgroups is as good as the tree developed at the end of Step 4.

Steps 1-4 can be thought of as the tree-growing phase; Step 5 can be thought of as the tree-pruning stage. The result of this process is a binary tree that naturally incorporates interactions between the predictor variables.

A sum of squares criterion is used to find the “best” split when a continuous response like CPUE is modeled. The criterion is

$$SS_T - (SS_L + SS_R) \quad (1)$$

where

$$SS_T = \sum_{i=1}^{n_T} (y_{T,i} - \bar{y}_T)^2 \quad (2)$$

for observations ($y_{T,i}$) contained in parent node T . The mean CPUE in node T is \bar{y}_T , the number of observations in node T is n_T , and SS_L and SS_R are, respectively, the sums of squares for the left and right sons of that parent. The criterion described in Expression 1 chooses splits that maximize the sum of squares between the left and right sons. Note that, when determining whether to split a node, this criterion is computed across the levels of all the predictor variables. These computations occur during Steps 1-3 of the tree-growing process.

The tree-growing phase is stopped (Step 4) on a node-by-node basis. Growth at a node is stopped when a further split at that node would not improve the overall fit of the tree beyond an arbitrary threshold. This threshold is called the “complexity parameter” (cp). This parameter has a direct interpretation: if splitting a node does not increase the overall R^2 of the tree model by at least cp , then that node is not split. R^2 has the usual definition of $1 - (\text{error SS}/\text{total SS})$. When growth is

stopped at a node, that node is called a terminal node. The value of cp is inversely related to the number of terminal nodes in a tree; when cp is small, the tree has a large number of terminal nodes. During the growth phase, cp should be set to a small value so that the tree can become quite large. The number of splits in a tree (number of terminal nodes – 1) is the number of parameters in the model.

Pruning a fully-grown tree (Step 5) is necessary because the marginal improvement in the overall fit of the tree decreases as the number of terminal nodes increases. Fully-grown trees can be pruned with a cross-validation technique. During cross-validation, the data are randomly assigned to ten separate groups. For a single value of cp , trees are grown on a combination of nine data-groups and the overall R^2 is calculated for the tenth group. This procedure is repeated ten times (each of the ten groups is held out as the validation sample) for each value of cp . This cross-validation procedure provides an estimate of the expected value and standard error of the overall R^2 for each value of cp (number of nodes). The expected value of the cross-validated $1-R^2$ (“xerror”) first decreases and eventually increases as the number of nodes is increased (cp is decreased). The “1-SE rule” (Breiman et. al 1984) can be used to find an appropriately sized tree from the cross-validation results. According to the 1-SE rule, an “appropriate parameterization” is obtained by the largest value of cp with an xerror that is within one standard error of the minimum xerror. This rule is easy to visualize graphically. The result of applying the 1-SE rule is a pruned tree that can explain nearly as much variation in the response variable as the fully-grown tree but with fewer parameters.

This procedure was applied to the current set of PATH modeling results to determine the relative influence of the different factors on the results. Graphs of R-squared vs. the number of splits are shown in Figure E-1, with the optimal number of splits indicated by the vertical line. The resulting tree using the 24-year survival standard probability as the dependent variable was shown in Figure 3-2. The extra mortality hypotheses, the passage/transportation model, and the life-cycle model account for the vast majority of the variation in 24-year survival probabilities (vertical lengths of the branches are proportional to the proportion of the variance accounted for by that factor). Interestingly, the actions themselves generally account for a small proportion of the variance in results for the 24-year survival standard. For some combinations of factors (e.g., Hydro extra mortality, CRiSP/T4 passage/transportation model, delta model), the effects of the actions are not significant at all. However, with the 100-year survival and the 48-year recovery standard the split between A1/A2 and A3 accounted for the majority of the variance (Figure 3-3, 3-4).

We also constructed a regression tree for the effects of factors on the geometric mean of the 24-year survival standard and the 48-year recovery standard. This value can be thought of as an indication of whether the stocks recover without going extinct. The tree is shown in Figure E-2; the graph of R-squared vs. number of splits is in Figure E-3.

Figure E-1: R-squared vs. # of splits, with the optimal number of splits indicated with vertical lines.

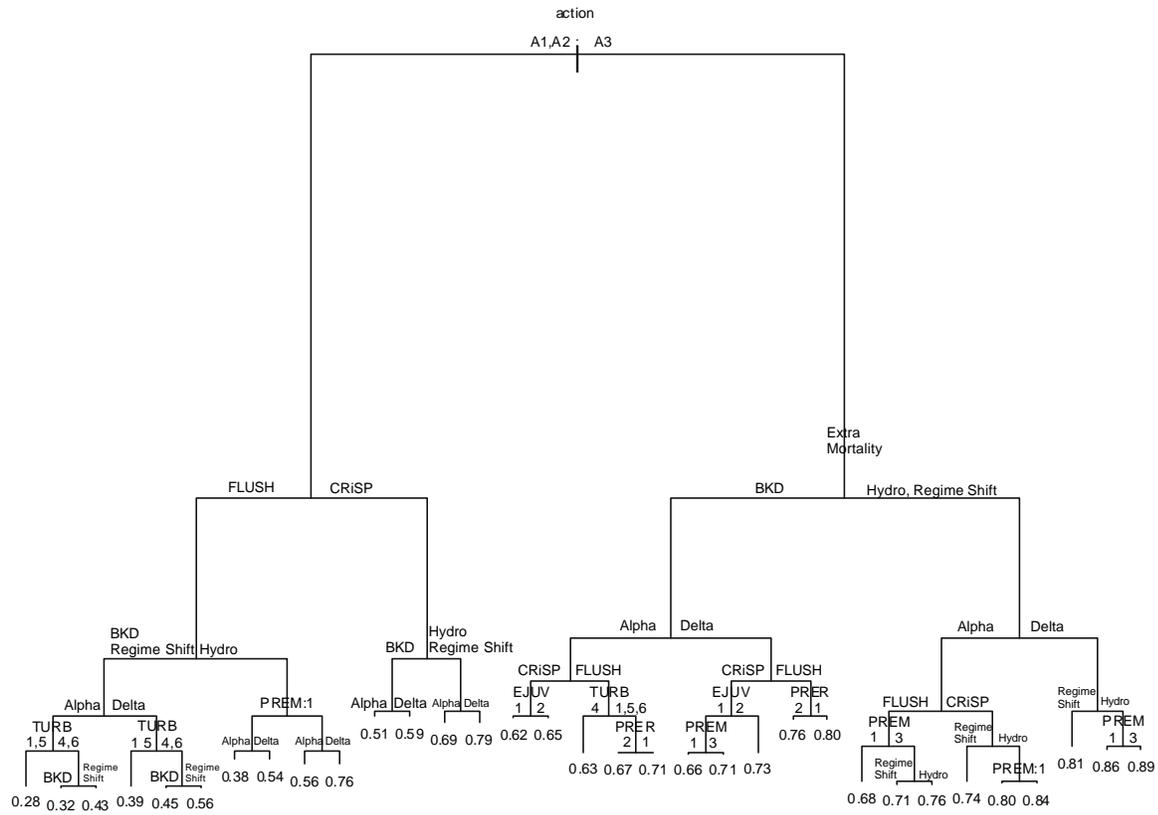


Figure E-2: Regression tree for geometric mean of jeopardy standards 1 (24-year survival) and 4 (48-year recovery) for all factors, truncated at R-squared > 0.95 (32 splits).

Geometric Mean of Jeopardy Stds. 1 & 4

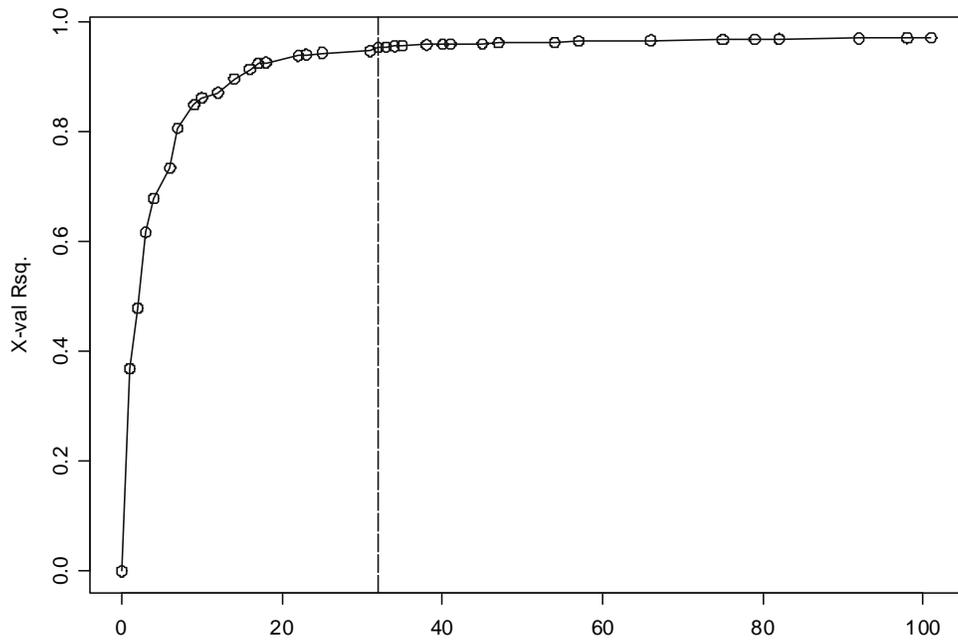


Figure E-3: R-squared vs. number of splits for the geometric mean of jeopardy standards 1 and 4.

Appendix F: Sensitivity Analysis of D Values

Introduction

A number of issues have been raised with respect to the implementation of the Transportation models. These issues are:

- a) Assumptions about historical turbine/bypass survival affect both retrospective estimates of passage survival (V_n) and retrospective estimates of the survival of control fish in transportation studies, which in turn determines the calculation of retrospective D values (see Appendix G).

All CRiSP D values in the current set of results are based on TURB4 assumptions about historical turbine/bypass survival. This means that those runs in which retrospective passage survival were based on TURB5 assumptions are not consistent with the TURB assumptions used to calculate D. See submission 23, pg. 7 for further discussion.

- b) Retrospective CRiSP D values are currently calculated as the median of D's estimated from individual transportation experiments. Medians are calculated in two time periods – pre-1980 and 1980-present. Some PATH participants have suggested that the geometric mean is a more appropriate estimate of the central tendency of D values than the median (Submission 23, pp. 1-4). They have also questioned the selection of 1980 as the “break point” between the two periods. They suggest that a break point of 1978 is more consistent with the pattern of D's over time (Submission 23, pp. 3-6).

Methods

We have done further sensitivity analyses to determine how important these issues are in determining the model results. TURB values used are in Table F-1. To address point a), we re-ran a subset of runs in which V_n estimates were based on TURB5 assumptions using the TURB5 D values so that the TURB assumptions used to estimate V_n and D were consistent. The runs we compared were:

- A1 using best-case passage assumptions – high FGE (FGE1), high predator removal effectiveness (PREM3);
- A2 using best-case passage assumptions;
- A2 using worst-case passage assumptions - low FGE (FGE2), low predator removal effectiveness (PREM1); and
- A3 using best-case passage assumptions and best-case drawdown assumptions – 3-year pre-removal period (PRER1), high equilibrated juvenile survival rate (EJUV2), 2-year transition period (TJUVA).

Table F-1: CRiSP Retrospective D Values for different TURB hypotheses.

	TCR	TURB1	TURB4	TURB5	Descaling
1968	2.07	0.519	0.519	0.519	3.1
1969	1.26	0.655	0.655	0.655	3.1
1970	1.45	0.356	0.356	0.356	3.1
1971	1.54	0.129	0.129	0.129	12
	1.66	0.139	0.139	0.139	12
1972	1.08	0.104	0.072	0.103	16
	1.12	0.107	0.075	0.107	16
1973	13.76	0.467	0.184	0.39	19.6
	18.39	0.624	0.247	0.515	19.6
1975	1.63	0.122	0.084	0.102	13
	2.59	0.208	0.14	0.197	13
1976	0.97	0.026	0.022	0.022	11.5
	6.06	0.201	0.164	0.166	11.5
	0.36	0.006	0.004	0.005	7
	1.01	0.014	0.011	0.011	7
1978	0.74	0.298	0.298	0.298	20
	8.51	3.43	3.43	3.43	7.5
	5.48	2.208	2.208	2.208	7.5
	0.94	0.378	0.378	0.378	7.5
1979	3.74	0.294	0.209	0.292	5.3
1986	1.58	0.571	0.571	0.571	3.7
1989	2.34	0.695	0.695	0.695	2.3
1994	1.8	0.554	0.554	0.554	3.6
1995	1.76	0.855	0.855	0.855	2.7

We compared average 24-year survival probabilities for these runs to the same runs in the current set of results in which TURB4 was used to estimate D values.

To address point b), we ran a series of runs where we varied the estimate of central tendency (median, geometric mean, or weighted average mean⁵) and the year in which the “break point” is assumed to occur (1978 or 1980). We looked at action A2 (maximum transportation), since that is where the effects of different assumptions about the relative post-Bonneville survival of transported and non-transported fish are likely to be the largest, and a smaller set of runs for A3 for comparison. All runs assumed FGE1, TURB4, and PREM3; drawdown runs assumed best-case drawdown assumptions. We also looked at the effects of including or excluding the 1973 and 1976 D estimates, because the CRiSP group has suggested that these points be omitted from the calculation (see Anderson 1998). Finally, we looked at the effects of using a CRiSP D value based on descaling, using D values in each retrospective year calculated from the

⁵ Where the D value from each transportation experiments is weighted by the inverse of the variance of the T:C ratio calculated from each experiment.

estimate of descaling in that year using a regression equation supplied by Jim Anderson (Aug. 3 e-mail)⁶. No estimate of D or descaling was supplied for 1977, so we used an average of the 1976 and 1978 D values for that year.

Results

a) Effects of using consistent TURB assumptions

Median D values for the pre-1980 and 1980-present time periods using TURB4 and TURB5 assumptions are compared in Table F-2.

Table F-2: Median D values using TURB4 and TURB5 assumptions

	TURB4	TURB5
Pre-1980	0.174	0.245
Post-1980	0.633	0.633

Results using consistent TURB assumptions is shown in Figure F-1 (24-year survival standard, Alpha model) F-2 (24-year survival standard, Delta model) and F-3 (48-year recovery standard, Alpha model). The values in those tables are jeopardy probabilities for the Marsh Creek stock, which was the 6th best stock for the 24-year survival standard in about 85% of all of the runs. The probabilities are averaged over the 5 extra mortality/future climate hypotheses.

Jeopardy probabilities using TURB5 assumptions are lower than when TURB4 was used. The differences are small (the largest difference was around 0.03 with A2 under the Alpha model), particularly with the Delta model, and are relatively uniform across actions and across extra mortality/future climate hypotheses (Fig. F-4). This suggests that overall, using TURB5 assumptions consistently for CRiSP Vn and D estimates would result in no change in the ranking of actions, but would slightly lower the ability of all actions to achieve the jeopardy standards.

⁶ Jim Anderson's e-mail provided retrospective D estimates for TURB1, 4, and 5 for each year in which there was a transportation study. TURB 4 D values in two of those years, 1975 and 1995, differed slightly from the estimates provided in the Anderson January 1998 D document, which were used in the current set of results. To make these sensitivity analyses comparable to the current set of base results, I used the values from the Anderson 1998 document for those years. The effects of using these different values for 1975 and 1995 D's on the estimates of D's in Table F-2 are minor.

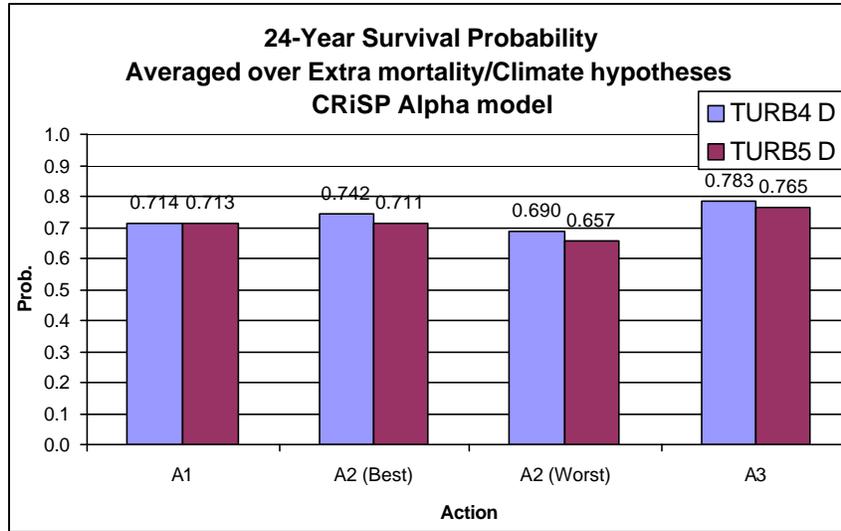


Figure F-1: Comparison of 24-year survival probabilities using different TURB assumptions (Alpha model).

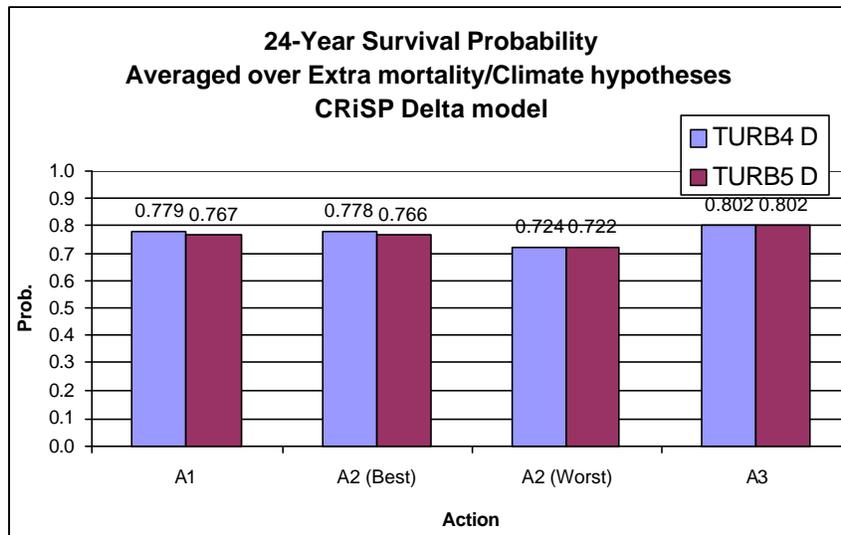


Figure F-2: Comparison of 24-year survival probabilities using different TURB assumptions (Delta model).

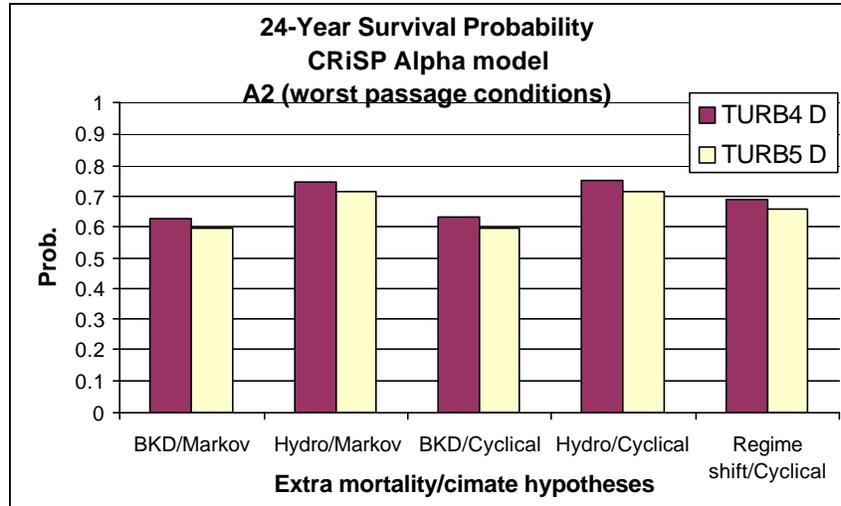


Figure F-3: Comparison of 24-year survival probabilities using different TURB assumptions, shown for each extra mortality/future climate hypothesis.

b) Effects of different estimates of central tendencies, break points, and inclusion/exclusion of 1973 and 1976 D estimates.

D values for each of these combinations of assumptions are shown in Table F-3. Average 24-year survival probabilities using these D values for the Marsh Creek stock are shown in Table F-4. Effects of the different D estimates are generally small. The largest difference is between the A3 median with a 1980 break point and 1973/1976 values included (jeopardy probability = 0.800) and the A3 geometric mean with a 1978 break point and 1973/1976 values included (jeopardy probability = 0.734). D values based on the descaling regression are intermediate to other D values. Patterns are consistent across actions, jeopardy standards, and extra mortality/future climate hypotheses (Fig. F-4 to F-6). These results are for the Alpha model; we did not do a full set of runs with the delta model because of time constraints. However, we did compare a geometric mean with a 1978 break point and 1973/1976 values included to the weighted arithmetic mean with a 1978 break point and 1973/1976 values excluded for A2 (this comparison yielded the largest difference in jeopardy probabilities for A2 under the Alpha model in Table F-4). With the delta model, the jeopardy probabilities were closer together than the Alpha model (0.75 for the geometric mean D, 0.767 for the weighted arithmetic mean D). The general conclusion is that these alternative D values will have slight effects on the ability of actions to meet the standards, but will probably not affect the relative ranking of actions.

Table F-3: D values for various assumptions about estimates of central tendencies (median, wtd. Arithmetic mean, geomean), break points for defining pre/post time periods, and inclusion or exclusion of 1973 and 1976 values. The first value in each cell is the D estimate for the pre-break point period, the last value is for the post-break point period.

Break Point	Median		Wtd. Arithmetic mean		Geometric mean	
	+73/76	-73/76	+73/76	-73/76	+73/76	-73/76
1978	.139 / .571	.139 / .571			.100 / .686	.173 / .686
1980	.174 / .633	.254 / .633		.292 / .634	.163 / .658	.286 / .658

Table F-4: Effects of assumptions about estimates of central tendencies (median, wtd. arithmetic mean, geomean), break points for defining pre/post time periods, and inclusion or exclusion of 1973 and 1976 values on average 24-year survival probabilities for Marsh Creek (Alpha model).

A2	Median		Wtd. Arithmetic mean		Geometric mean	
	+73/76	-73/76	+73/76	-73/76	+73/76	-73/76
Break Point	+73/76	-73/76	+73/76	-73/76	+73/76	-73/76
1978	0.725	0.725			0.673	0.700
1980	0.726	0.730		0.737	0.723	0.729
Decaling	0.712					
A3	Median		Wtd. Arithmetic mean		Geometric mean	
	+73/77	-73/33	+73/77	-73/33	+73/77	-73/33
Break Point	+73/77	-73/33	+73/77	-73/33	+73/77	-73/33
1978	0.775				0.734	0.75672
1980	0.800	0.780		0.786		0.779

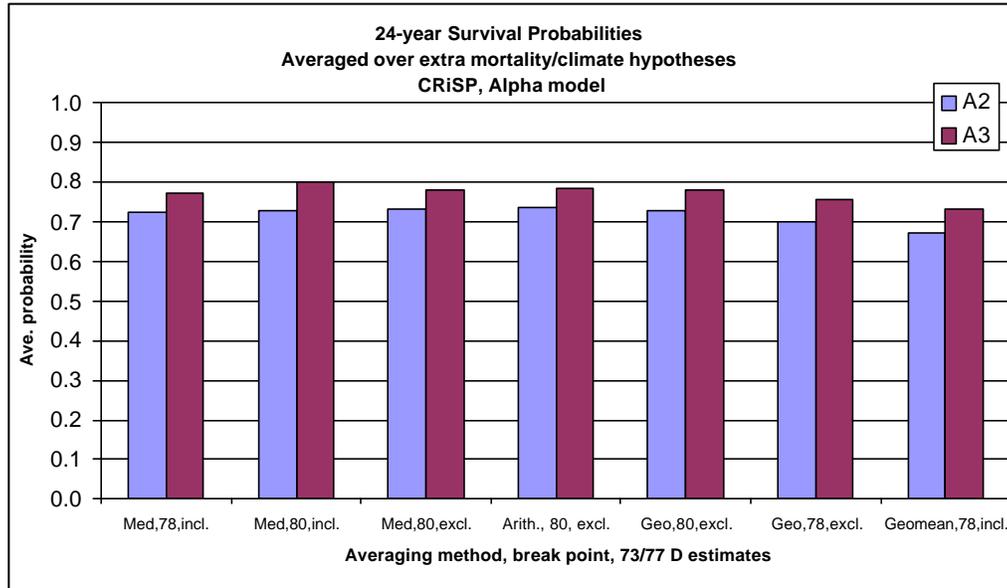


Figure F-4: Comparison of 24-year survival probabilities using different methods of calculating D.

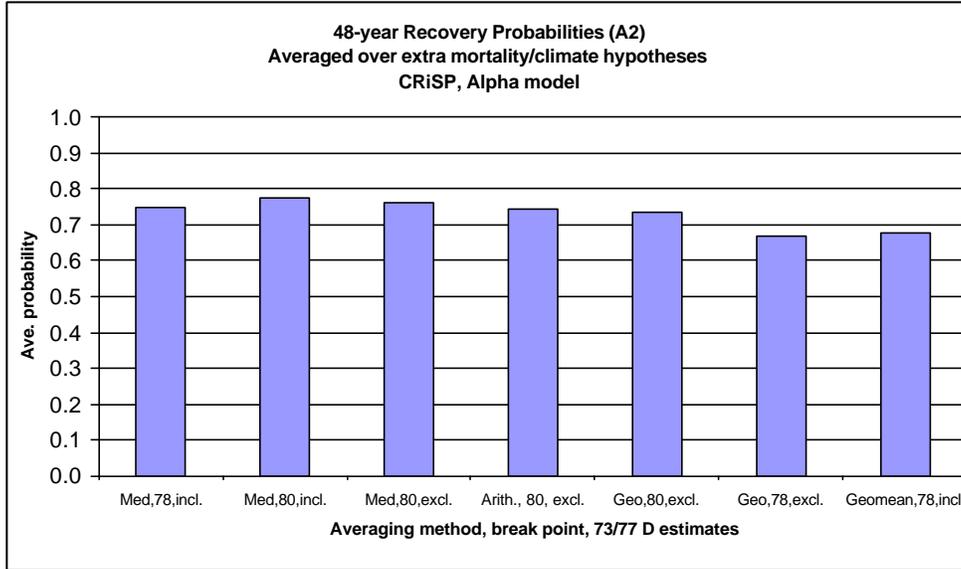


Figure F-5: Comparison of 48-year recovery probabilities using different methods of calculating D.

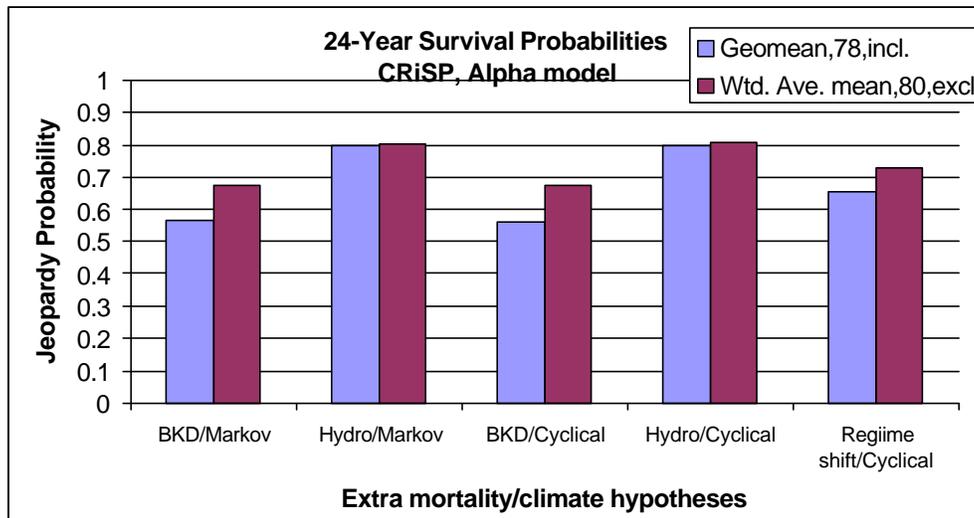


Figure F-6: Comparison of 24-year survival probabilities using different methods of calculating D, shown for each extra mortality/future climate hypothesis.

Appendix G: Sensitivity Analyses of Delta Model

Concerns have been raised about the sensitivity of the life cycle models to certain years of spawner-recruit data for the Middle Fork John Day stock and to the X-dam parameter. In response, we have conducted a sensitivity analysis of the Delta model to these factors. The results are summarized in the tables below. They show the jeopardy probabilities averaged over a better than average and worse than average set of passage model runs, and the five extra mortality/future climate hypotheses. These analyses were based on the runs compiled for the Preliminary Decision Analysis Report (i.e., they are based on the old set of CRiSP D and A3 Vn values), but they still provide an accurate indication of the sensitivity of the model to the various factors.

The results show that the delta model is not very sensitive to any of these factors (Table G-1, G-2). The largest change in jeopardy probabilities is around 0.07, which is not very big given the other uncertainties that affect these values. Furthermore, the different assumptions do not affect the overall ranking of actions.

Table G-1: Sensitivity of average probabilities to spawner-recruit data and X-dam parameter (CRiSP)

Jeopardy Standard	Action	Model					
		delta	delta no 1963	delta no M Jday	delta fix "X"	delta no 1963 fix "X"	delta no M Jday fix "X"
P(S>survival; 24 yrs)	A1	0.71	0.70	0.72	0.71	0.69	0.70
	A2	0.70	0.69	0.70	0.69	0.68	0.70
	A3	0.70	0.67	0.70	0.69	0.68	0.67
P(S>survival; 100 yrs)	A1	0.79	0.78	0.77	0.79	0.79	0.79
	A2	0.78	0.77	0.75	0.78	0.78	0.79
	A3	0.71	0.71	0.69	0.71	0.70	0.72
P(S>recovery; 48 yrs)	A1	0.62	0.62	0.58	0.62	0.61	0.61
	A2	0.60	0.59	0.53	0.60	0.58	0.60
	A3	0.48	0.48	0.43	0.46	0.46	0.45

Key:

delta	=	delta model
no 1963	=	exclude 1963 brood year data except for stock #13 (one stock needed to avoid reparameterization)
no M Jday	=	exclude pre-1974 brood year data for middle fork John Day stock
fix "X"	=	fix "X" dam parameter to equal passage model input

Table G-2: Sensitivity of average probabilities to spawner-recruit data and X-dam parameter (FLUSH)

Jeopardy Standard	Action	Model					
		delta	delta no 1963	delta no M Jday	delta fix "X"	delta no 1963 fix "X"	delta no M Jday fix "X"
P(S>survival; 24 yrs)	A1	0.62	0.60	0.66	0.63	0.62	0.66
	A2	0.61	0.56	0.65	0.60	0.57	0.62
	A3	0.68	0.67	0.71	0.69	0.68	0.71
P(S>survival; 100 yrs)	A1	0.71	0.70	0.71	0.71	0.69	0.72
	A2	0.68	0.67	0.69	0.69	0.68	0.70
	A3	0.89	0.89	0.87	0.89	0.88	0.88
P(S>recovery; 48 yrs)	A1	0.42	0.41	0.43	0.38	0.36	0.41
	A2	0.37	0.35	0.40	0.36	0.36	0.40
	A3	0.90	0.90	0.84	0.89	0.90	0.90

Key:

- delta = delta model
- no 1963 = exclude 1963 brood year data except for stock #13 (one stock needed to avoid reparameterization)
- no M Jday = exclude pre-1974 brood year data for middle fork John Day stock
- fix "X" = fix "X" dam parameter to equal passage model input

Appendix H: Sensitivity Analysis of an Alternative Formulation of the Hydro Extra Mortality Hypothesis

An alternative hydro hypothesis is that the I_n (post-Bonneville survival factor for non-transported fish) changed in recent years due to the Snake River dams. If the Snake River dams were not present then I_n would return to values observed in the historical data (taken to be pre-1970 for modeling purposes).

In the prospective modeling of the draw-down scenario (A3), I_n under the alternative hydro hypothesis is assumed to equal values selected according to matching water years from the pre-1970 values. This alternative hydro hypothesis is modeled very similar to the way the regime shift is modeled, except rather than cycling, I_n is permanently altered under Action A3. Thus we again use equation [A.3.3.2-1] in the Preliminary Report):

$$\ln(R_y) = (1 + p)\ln(S_y) + a - bS_y - m_r + \ln[\mathbf{w}_y / \mathbf{w}_r] + \mathbf{d}_y + \mathbf{e}_y \quad [\text{Eq. A.3.3.2-1}]$$

In this equation, the retrospective water year chosen for a given prospective year is one which occurred during the pre-Snake dam years (pre-1970 brood year). The only other difference between this alternative hydro implementation is that the D values (for Flush only) prospectively are values read from the input prospective file. Crisp D values are selected at random.

For actions A1 & A2, the alternative hydro hypothesis is modeled very similar to the way the current BKD hypothesis is modeled -- that is, it's here to stay. In the alternative view, the hydro hypothesis is that the I_n was altered by building of Snake River dams and if the dams stay (e.g., in A1 & A2) then the recent I_n stays as well. The only variant of this hydro hypothesis from the current BKD implementation is that, for FLUSH only, the D's prospectively could be selected as input values for the hydro hypothesis, rather than the random selection of recent D values as in the current BKD hypothesis implementation. CRiSP is an exact analog because D's are chosen randomly in both.

We did a short sensitivity study to examine the consequence to an A3 scenario of the alternative hydro assumption versus the one in the decision document. We assumed a best case passage and drawdown scenario: FGE1, TURB4, PREM3, PRER1, EJUV2 AND TJUVa (runs P516-C3 and P048-F3).

Results

The alternative representation of the hydrosystem extra mortality hypothesis generally produces higher probabilities of remaining above all three jeopardy standards (Fig. H-1). This occurs under the alpha model for both CRiSP and FLUSH, and under the delta model for CRiSP. With the Delta model, FLUSH generates approximately equal probabilities using the two formulations.

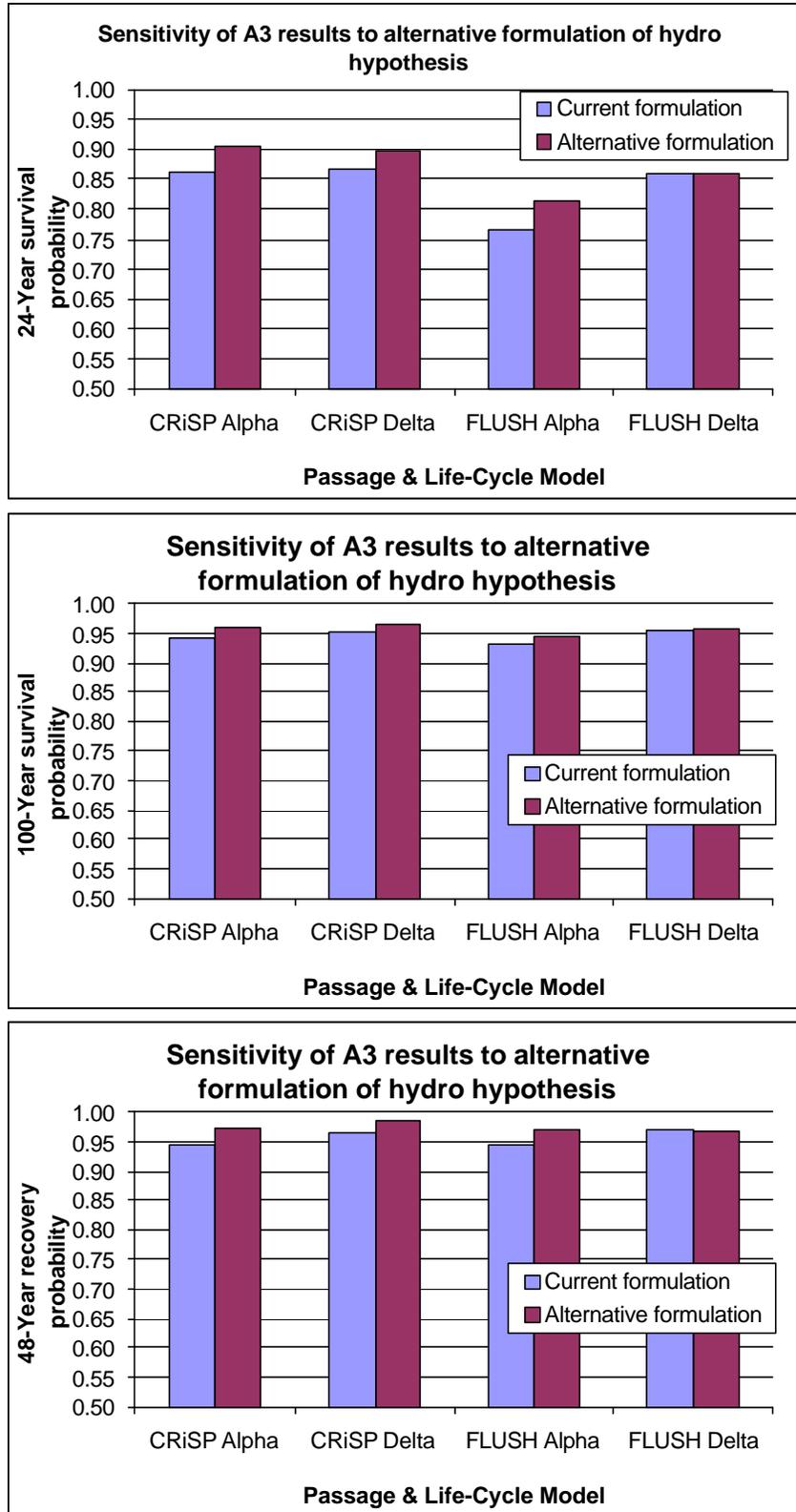


Figure H-1: Sensitivity of A3 jeopardy probabilities to an alternative formulation of the hydro extra mortality hypothesis.

Appendix I: Sensitivity Analysis of Alternative FLUSH D Values

The FLUSH group has calculated D values for individual T/C studies between 1980 and 1989 (Table I-1), similar to the approach used by CRiSP (except CRiSP also used 1968, 1969, 1994 and 1995 T/C studies as well). The geometric means of these values, for different TURB assumptions and time periods, are shown in Table I-2.

Table I-1: Alternative FLUSH D values

Year	Control Release Point 1	TURB1	TURB5	TURB4	TURB6
1968	4				
1969	4				
1970	4	0.409	0.453	0.631	0.411
1971	2	0.757	0.780	0.785	0.758
1971	2	0.813	0.838	0.843	0.814
1972	2	0.360	0.391	0.279	0.355
1972	2	0.376	0.408	0.291	0.370
1973	2	0.423	0.486	0.969	0.406
1973	2	0.565	0.650	1.295	0.542
1975	1	0.333	0.360	0.405	0.327
1975	1	0.530	0.573	0.644	0.520
1976	2	0.293	0.318	0.351	0.298
1976	2	1.834	1.986	2.196	1.862
1976	1	0.103	0.109	0.110	0.104
1976	1	0.291	0.309	0.312	0.294
1978	2	0.113	0.127	0.176	0.120
1978	1	0.470	0.515	0.588	0.456
1978	1	0.303	0.332	0.379	0.294
1978	1	0.052	0.057	0.065	0.051
1979	1	0.429	0.491	0.686	0.425
1986	2	0.359	0.406	0.560	0.362
1989	2	0.408	0.469	0.742	0.411

1/ 1 = LGR, 2 = LGS, 4 = IHR

Table I-2: Geometric mean alternative FLUSH D values

Year	TURB1	TURB5	TURB4	TURB6
70-79	0.357	0.387	0.449	0.353
78-89	0.244	0.275	0.357	0.244
86-89	0.383	0.436	0.644	0.387

Two sensitivities were run. In one sensitivity, prospective D's were selected from a log-normal distribution with geo-mean and variance of 1978-1989 values. In the other, prospective D's were selected from a log-normal distribution with geo-mean and variance of 1986-1989 values. Retrospectively, the runs fixed the D's at the geometric mean of pre-1980 and 1980-present values, similar to CRiSP. The runs used were:

- A1 – FGE1, PREM3, TURB5
- A2 – FGE1, PREM3, TURB5
– FGE1, PREM3, TURB4
- A3 – FGE1, PREM3, TURB5, best-case drawdown assumptions

Figure I-1 shows results for 24-year survival probabilities, averaged over the extra mortality hypotheses, for Marsh Creek. Figure I-2 shows the results for the extra mortality hypotheses for the A2(TURB5), Delta model case. Figure I-3 shows the results for the 48-year recovery standard. In all cases, jeopardy probabilities in the base case (current set of FLUSH results, with D's based on T:C vs. s relationship) are substantially higher than those calculated by using the 1978-1989 geometric means (geo78) prospectively, but are not that different from those calculated by using the 1986-1989 geometric means (geo86). Differences between D assumptions are greater with the 48-year recovery standard.

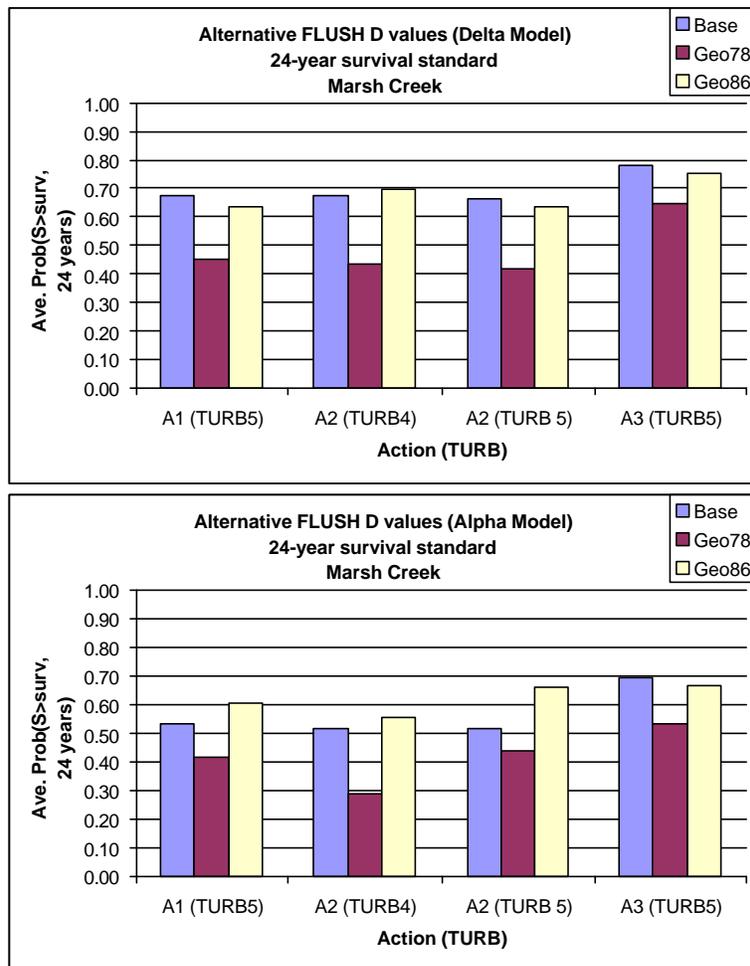


Figure I-1: Sensitivity to alternative FLUSH Ds, 24-year survival probability.

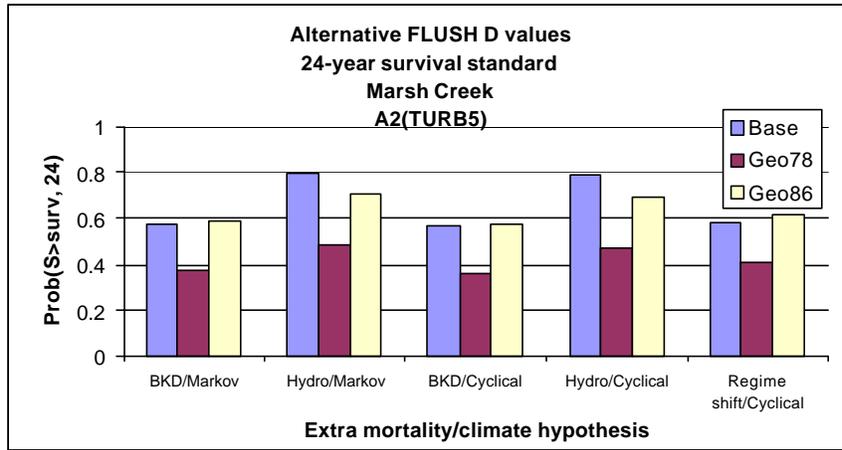


Figure I-2: Sensitivity to alternative FLUSH D's for different extra mortality/climate hypotheses (Delta model).

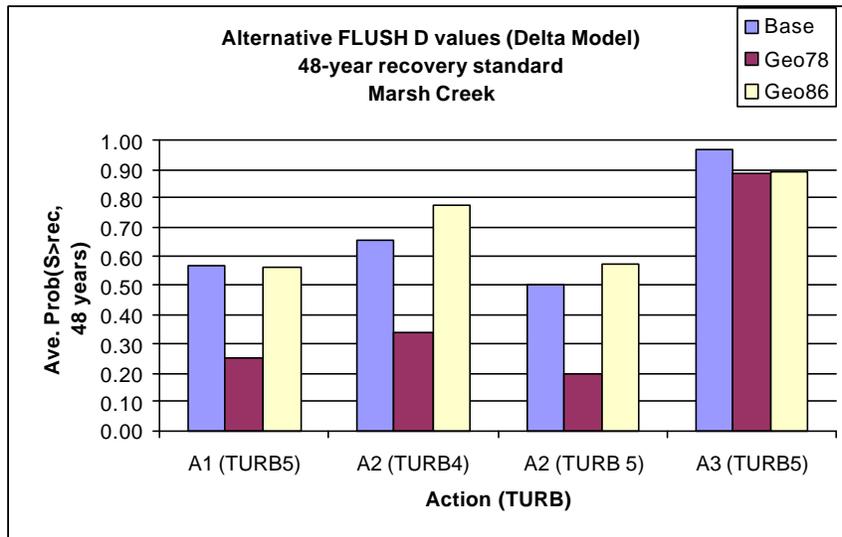


Figure I-3: Sensitivity to alternative FLUSH Ds, 48-year recovery probability.

