Monitoring and Evaluation of Smolt Migration in the Columbia Basin Volume VIII

Comparison of the RPA Testing Rules

Technical Report
2002

DOE/BP-00004134-3 August 2002

This Document should be cited as follows:

Skalski, John, Roger Ngouenet, ''Monitoring and Evaluation of Smolt Migration in the Columbia Basin Volume VIII'', Project No. 1991-05100, 39 electronic pages, (BPA Report DOE/BP-00004134-3)

Bonneville Power Administration
P.O. Box 3621
Portland, Oregon 97208

This report was funded by the Bonneville Power Administration (BPA), U.S. Department of Energy, as part of BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries. The views in this report are the author's and do not necessarily represent the views of BPA.

MONITORING AND EVALUATION OF SMOLT MIGRATION IN THE COLUMBIA BASIN

VOLUME VIII

Comparison of the RPA testing rules provided in the 2000 Federal Columbia River Power System (FCRPS) Biological Opinion with new test criteria designed to improve the statistical power of the biological assessments

Prepared by:

John R. Skalski

and

Roger F. Ngouenet

School of Aquatic and Fishery Sciences

University of Washington

1325 Fourth Avenue, Suite 1820

Seattle, WA 98101-2509

Prepared for:

U.S. Department of Energy

Bonneville Power Administration

Division of Fish and Wildlife

P.O. Box 3621

Portland, OR 97208-3621

Project No. 91-051-00

Contract No. 00004234

August 2002

Monitoring and Evaluation of Smolt Migration in the Columbia Basin

Other Publications in this Series

Volume I: Townsend, R. L., J. R. Skalski, and D. Yasuda. 1997. Evaluation of the 1995 predictions of run-timing of wild migrant subyearling chinook in the Snake River Basin using program RealTime. Technical Report (DOE/BP-35885-11) to BPA, Project 91-051-00, Contract 91-BI-91572.

Volume II: Townsend, R. L., J. R. Skalski, and D. Yasuda. 1998. Evaluation of the 1996 predictions of run-timing of wild migrant subyearling chinook in the Snake River Basin using program RealTime. Technical Report (DOE/BP-91572-2) to BPA, Project 91-051-00, Contract 91-BI-91572.

Volume III: Townsend, R. L., J. R. Skalski, and D. Yasuda. 2000. Evaluation of the 1997 predictions of run-timing of wild migrant yearling and subyearling chinook and sockeye in the Snake River Basin using program RealTime. Technical Report (accepted) to BPA, Project 91-051-00, Contract 91-BI-91572.

Volume IV: Burgess, C., R. L. Townsend, J. R. Skalski, and D. Yasuda. 2000. Evaluation of the 1998 predictions of the run-timing of wild migrant yearling and subyearling chinook and steelhead, and hatchery sockeye in the Snake River Basin using program RealTime. Technical Report (submitted) to BPA, Project 91-051-00, Contract 96BI-91572.

Volume V: Burgess, C., and J. R. Skalski. 2000. Evaluation of the 1999 predictions of the run-timing of wild migrant yearling and subyearling chinook salmon and steelhead trout, and hatchery sockeye salmon in the Snake River Basin using program RealTime. Technical Report (submitted) to BPA, Project 91-051-00, Contract 96BI-91572.

Volume VI: Burgess, C., and J. R. Skalski. 2000. Evaluation of the 2000 predictions of the run-timing of wild migrant chinook salmon and steelhead trout, and hatchery sockeye salmon in the Snake River Basin, and combined wild and Hatchery Salmonids migrating to Rock Island and McNary Dams using program RealTime. Technical Report (submitted) to BPA, Project 91-051-00, Contract 96BI-91572.

Volume VII: Skalski, J. R., and R. F. Ngouenet. 2000. Evaluation of the compliance testing framework for RPA improvement as stated in the 2000 Federal Columbia River Power System (FCRPS) Biological Opinion. Technical Report (submitted) to BPA, Project 91-051-00, Contract 96BI-91572.

Other Publications Related to this Series

Other related publications, reports and papers available through the professional literature or from the Bonneville Power Administration (BPA) Public Information Center - CKPS-1, P.O. Box 3621, Portland, OR 97208.

Lady, J., and J. R. Skalski. 1998. Estimators of stream residence time of Pacific salmon (spp. *Oncorhynchus*) based on release-recapture data. *Canadian Journal of Fisheries and Aquatic Sciences* 55:2580-2587.

Lowther, A. B., and J. R. Skalski. 1998. A multinomial likelihood model for estimating survival probabilities and residualization for fall chinook salmon (*Oncorhynchus tschawytscha*) smolt using release-recapture methods. *Journal of Agricultural Biology and Environmental Statistics* 3:223-236.

Mathur, D., P. G. Heisey, E. T. Euston, and J. R. Skalski. 1996. Turbine passage survival estimates for chinook salmon smolt (*Oncorhynchus tshawytscha*) at a large dam on the Columbia River. *Canadian Journal of Fisheries and Aquatic Sciences* 53:542-549.

Ryding, K. E., and J. R. Skalski. 1999. Multivariate regression relationships between ocean conditions and early marine survival of coho salmon (*Oncorhynchus kisutch*). *Canadian Journal of Fisheries and Aquatic Sciences* 56:2374-2384.

Skalski, J. R. 1998. Estimating season-wide survival rates of outmigrating smolt in the Snake River, Washington. *Canadian Journal of Fisheries and Aquatic Sciences* 55:761-769.

Skalski, J. R., S. G. Smith, R. N. Iwamoto, J. G. Williams, and A. Hoffmann. 1998. Use of PIT-tags to estimate survival of migrating juvenile salmonids in the Snake and Columbia Rivers. *Canadian Journal of Fisheries and Aquatic Sciences* 55:1484-1493.

Skalski, J. R. 1996. Regression of abundance estimates from mark-recapture surveys against environmental covariates. *Canadian Journal of Fisheries and Aquatic Sciences* 53:196-204.

Townsend, R. L., D. Yasuda, and J. R. Skalski. 1997. Evaluation of the 1996 predictions of run timing of wild migrant spring/summer yearling chinook in the Snake River Basin using Program RealTime. Technical Report (DOE/BP-91572-1) to BPA, Project 91-051-00, Contract 91-BI-91572.

Townsend, R. L., P. Westhagen, D. Yasuda, J. R. Skalski, and K. Ryding. 1996. Evaluation of the 1995 predictions of run timing of wild migrant spring/summer yearling chinook in the Snake River Basin using Program RealTime. Technical Report (DOE/BP-35885-9) to BPA, Project 91-051-00, Contract 87-BI-35885.

Townsend, R. L., P. Westhagen, D. Yasuda, and J. R. Skalski. 1995. Evaluation of the 1994 predictions of run-timing capabilities using PIT-tag databases. Technical Report (DOE/BP-35885-8) to BPA, Project 91-051-00, Contract 87-BI-35885.

Skalski, J. R., G. Tartakovsky, S. G. Smith, P. Westhagen, and A. E. Giorgi. 1994. Pre-1994 season projection of run-timing capabilities using PIT-tag databases. Technical Report (DOE/BP-35885-7) to BPA, Project 91-051-00, Contract 87-BI-35885.

Skalski, J. R., and A. E. Giorgi. 1993. A plan for estimating smolt travel time and survival in the Snake and Columbia Rivers. Technical Report (DOE/BP-35885-3) to BPA, Project 91-051-00, Contract 87-BI-35885.

Smith, S. G., J. R. Skalski, and A. E. Giorgi. 1993. Statistical evaluation of travel time estimation based on data from freeze-branded chinook salmon on the Snake River, 1982-1990. Technical Report (DOE/BP-35885-4) to BPA, Project 91-051-00, Contract 87-BI-35885.

PREFACE

Project 91-051 was initiated in 1991 in response to the Endangered Species Act (ESA) listings in the Snake River Basin of the Columbia River Basin. Primary objectives and management implications of this project include: (1) to address the need for further synthesis of historical tagging and other biological information to improve understanding and identify future research and analysis needs; (2) to assist in the development of improved monitoring capabilities, statistical methodologies, and software tools to aid management in optimizing operational and fish passage strategies to maximize the protection and survival of listed threatened and endangered Snake River salmon populations and other listed and nonlisted stocks in the Columbia River Basin; (3) to design better analysis tools for evaluation programs; and (4) to provide statistical support to the Bonneville Power Administration and the Northwest fisheries community.

The following report addresses measure 4.3C of the 1994 Northwest Power Planning Council's Fish and Wildlife Program with emphasis on improved monitoring and evaluation of smolt migration in the Columbia River Basin. In this report, statistical models are used to evaluate the framework for compliance testing of the RPA improvements using the information provided in the Federal Columbia River Power System (FCRPS) 2000 Biological Opinion (BO). The main concern is to evaluate the anticipated performance of two statistical hypothesis tests proposed in the 2000 FCRPS BO. It is hoped that assessing the compliance rules before actual data are collected will help avoid any unpleasant surprises concerning the statistical behavior of the proposed decision rules for compliance evaluation in 2005 and 2008. Having the capability to correctly assess the outcome of the RPAs should improve the public confidence in the recovery process and should also contribute to the regional goal of increasing juvenile passage survival through the Columbia River System.

ABSTRACT

The 2000 FCRPS Biological Opinion (BO) suggested two statistical hypothesis tests to assess the RPA compliance by the years 2005 and 2008. With the decision rules proposed in the BO, Skalski and Ngouenet (2001) developed a compliance framework based on classical t-tests and used Monte-Carlo simulations to calculate power curves. Unfortunately, the two-sample t-tests proposed in the BO only have moderate-to-low probability of correctly assessing the true status of the smolt survival recovery. We have developed a superior two-phase regression statistical model for testing the RPA compliance. The two-phase regression model improves the statistical power over the standard two-sample t-tests. In addition, the two-phase regression model has a higher probability of correctly assessing the true status of the smolt survival recovery. These classical statistical power curve approaches do not incorporate prior knowledge into the decision process. Therefore, we propose to examine Bayesian methods that complement classical statistics in situations where uncertainty must be taken into account. The Bayesian analysis will incorporate scientific/biological knowledge/expertise to thoroughly assess the RPA compliance in 2005 and 2008.

EXECUTIVE SUMMARY

Objectives

The 2000 Federal Columbia River Power System (FCRPS) Biological Opinion (BO) recommended performance measures and system goals to help recover listed salmonid species. Under the BO, the National Marine Fisheries Services (NMFS) will evaluate the RPA performance in 2005 and 2008 by comparing post-2000 smolt survivals with pre-2000 smolt survival estimates using standard statistical t-test. These models do not account for the possibility of gradual improvement in smolt survival, nor do they utilize information on a sudden change in juvenile survival.

The objectives of this report were to compare the RPA testing rules provided in the 2000 BO with a new two-phase, regression-based (Beckman and Cook 1979, Hinkley 1971, Searle 1971) test criteria designed to improve the statistical power of the biological assessments.

Results

The two-phase regression test model was applied to the survival data of three selected stocks: yearling chinook, subyearling chinook, and steelhead. The results of the power calculations were interpreted in terms of the ability of the model to correctly identify the true states of recovery (i.e., meet or exceed RPA expectations). Tables 6, 9, and 12 summarize the probabilities of making the correct decisions with the new two-phase regression tests. This new model improves the statistical power compare to the standard t-test and has a very high probability of correctly assessing the true status of the recovery by the years 2005 and 2008.

Recommendations

Two-phase regression statistical tests do not incorporate prior knowledge into the decision process. Therefore, we propose to immediately examine Bayesian methods that complement classical statistics in situations where uncertainty must be taken into account. In addition, Bayesian approaches could be the models of choice in helping convey the conclusions on smolt survival recovery to the public.

The next phase of this project is to examine Bayesian decision rules. The Bayesian analysis will incorporate scientific/biological knowledge/expertise to thoroughly assess the RPA compliance in 2005 and 2008.

ACKNOWLEDGMENTS

This report was funded by the Bonneville Power Administration (BPA), US Department of Energy, under Contract No. 96BI91572, as part of the BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries.

We wish to express thanks to the many federal agencies that have expended considerable resources in completing the Federal Columbia River Power System (FCRPS) 2000 Biological Opinion.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 Introduction

The Federal Columbia River Power System (FCRPS) 2000 Biological Opinion (BO) collected an extensive amount of data on smolt survival rates, designed a Reasonable and Prudent Alternative (RPA), and simulated estimates of the effects of the RPA in action areas by fish stock. In addition, the BO suggested the following compliance testing model for the RPA:

> "The progress check might consist of series of two-samples statistical tests on one-sided hypotheses about juvenile survival levels. The tests would take into account uncertainty in both the 1994-to-1999 and the more recent average. A first test could check whether the post-2000 estimate of survival was significantly lower than the 1994-to-1999 average, plus RPA improvements. The second test could check whether post-2000 survival was significantly higher than the 1994-to-1999 average."

The apparent motivation of this decision rule is to provide equal opportunity to conclude or reject the premise of recovery. In the primarily report (Skalski and Ngouenet 2001), we developed a compliance framework based on classical t-tests and used Monte-Carlo simulations to evaluate this performance under the assumption of a gradual improvement in smolt survivals. As anticipated, the two sample t-test had low statistical power to correctly identify the correct state of recovery or non-recovery.

In this report, we have developed a two-phase regression model for use with smolt survival data that has a higher power to detect RPA improvement with the current historical survival database for all the three fish stocks: yearling chinook salmon, subyearling chinook salmon and steelhead. Two-phase regression is a useful statistical method to detect and study shifts in survival rate over time. The method consists of two elements; identify a step-point where a trend in survival changes over time followed by a compression of mean survival on either side of the step-point.

2.0 Description of Data

Survival probabilities at each FCRPS project were estimated by NMFS with the Simulated Passage (SIMPAS) spreadsheet model. NMFS used the available empirical data on reach survival from passive integrated transponder (PIT)-tag release-recapture studies collected from 1994 through 1999 to estimate survival probabilities between successive dams (i.e., detection sites) for yearling and subyearling chinook salmon and steelhead. The data used in this report concentrate on the overall inriver survival rates of juvenile chinook salmon and steelhead throughout the system (i.e., between Lower Granite and Bonneville dams). Given the inriver survival rates for each dam of the FCRPS network, the overall reach survival through the FCRPS projects for a specified year is the product of the estimates for each of the shorter reaches.

The data used in this analysis came from tables showing project survival rates of juvenile salmonids in Appendix D of the 2000 FCRPS biological opinion. These tables recorded three types of survival rates for a given year: reach, pool and dam. One table was presented for each of the three ESUs: yearling chinook, subyearling chinook, and steelhead. Tables 1-3 summarize the survival rates of juvenile salmonids used to investigate the RPA decision rules. For the yearling chinook and steelhead salmonids, data are available from 1994 to 1999. The subyearling chinook salmon data are available from 1995 to 1999. The parameters of interest in our statistical evaluation are the number of years baseline data and the mean and variance in annual survival estimates (Tables 1-3). For example, from Table 1, the test of RPA compliance for the yearling chinook salmon survival from Lower Granite to Bonneville will use six years of baseline estimates, with mean survival 0.409, and variance 0.006.

Table 1. Reach survival rates of juvenile yearling (spring/summer) chinook salmon through FCRPS: 1994-1999, Lower Granite (LGR) to Bonneville (BON), McNary (MCN) to Bonneville, and LGR to MCN.

Year	LGR to BON	LGR to MCN	MCN to BON
1994	0.272	0.586	0.465
1995	0.418	0.692	0.604
1996	0.407	0.733	0.555
1997	0.385	0.687	0.560
1998	0.451	0.743	0.607
1999	0.518	0.786	0.660
Mean	0.409	0.704	0.575
Variance	0.006	0.005	0.004

Table 2. Reach survival rates of juvenile subyearling (fall) chinook salmon through FCRPS: 1995-1999, Lower Granite (LGR) to Bonneville (BON), McNary (MCN) to Bonneville, LGR to MCN.

Year	LGR to BON	LGR to MCN	MCN to BON
1995	0.164	0.415	0.396
1996	0.113	0.294	0.386
1997	0.005	0.082	0.059
1998	0.139	0.348	0.399
1999	0.086	0.364	0.237
Mean	0.102	0.301	0.296
Variance	0.004	0.018	0.022

Table 3. Each survival rates of juvenile Steelhead through FCRPS: 1994-1999, Lower Granite (LGR) to Bonneville (BON), McNary (MCN) to Bonneville, LGR to MCN.

Year	LGR to BON	LGR to MCN	MCN to BON
1994	0.322	0.615	0.523
1995	0.478	0.747	0.640
1996	0.428	0.730	0.586
1997	0.456	0.766	0.595
1998	0.417	0.683	0.611
1999	0.402	0.702	0.573
Mean	0.417	0.707	0.588
Variance	0.003	0.003	0.001

The broad ranges of the survival rates reported in the RPA (Table 4) reflect variable hydraulic/environmental conditions between years and statistical sampling error.

Table 4. Summary of estimated effects of the RPA in the action areas by fish stock

Location / Stock	Survival Rates		
	Juvenile Standard*	Juvenile Current*	Δ**
Snake River Spring/Summer Chinook	35-62%	27-52%	~9%
Snake River Fall Chinook	1-22%	0.5-15%	~5%
Upper Columbia Spring Chinook	55-76%	46-66%	~9%
Lower Columbia Spring Chinook	87-95%	83-91%	~5%
Lower Columbia Fall Chinook	57-85%	50-80%	~5%
Snake River Steelhead	42-58%	32-46%	~9%
Upper Columbia Steelhead	61-74%	57-64%	~9%
Middle Columbia Steelhead	61-74%	57-64%	~9%
Lower Columbia Steelhead	86-96%	85-92%	~4%

* From Table 9.7-5 (2000 FCRPS BO)

** From Table 9.7-18 (2000 FCRPS BO)

3.0 Two-Phase Regression

The two-phase regression is a useful model to diagnose significant changes in straight-line relationships on data. This is achieved by estimating the best location for a step-point. The step-point also labeled knot, elbow, spline, or break is the location where the straight-line relationship makes the transition from one regression line to the other (i.e., the point where the slope of the line shifts). Computer simulation studies suggest two-phase regression models are very robust with regard to the location of the transition point. We used least-squares regression procedures to estimate the best location for a step-point. In this study we used the year 2000 as the step-point in our simulations.

Under the scenario that RPA improvements are occurring, a break along the trend might manifest itself as a substantial change in survival. Figure 1 illustrates two-phase regression relationships with a jump in 2000 and a sustained improvement throughout the post-2000 years. A step or jump in mean survival value along a trend suggests an improvement in survival. When such a jump appears in year 2000 or any post-2000 years, the improvement is attributed to the RPA. A reformulation of the decision criteria suggested in the BO in the context of a two-phase regression model is as follows:

Test #1:

H_o: A step-point (knot) occurred during the year-2000 and the post-knot estimate of survival is greater than the pre-knot average plus 9% RPA expected survival improvements. (1)

H_a: The post-knot estimate of survival is lower than the pre-knot average plus 9% RPA expected survival improvements.

Test #2:

H_o: The post-knot predicted survival is lower than the pre-knot survival rate in average. (2)

H_a: A step point (knot) occurred during the year-2000 and the post-knot predicted survival is greater than pre-knot survival rate in average.

Monte Carlo simulation techniques were used to estimate the statistical power of the test procedures. The baseline data were the historical observations. Post-2000 data were generated by establishing an upward linear trend with the improvement reaching the RPA target in a specified year. The trend is estimated with a linear regression model fit on pre-2000 smolt survival data and future survivals randomly generate by adding a normal noise to the baseline for *N* subsequent years (where N is equal to 5 or 8 and represents the time span set by the FCRPS 2000 BO for RPA compliance testing 2005 or 2008). In our assessment of two-step regression method, we studied three scenarios with the improvement reaching $\Delta = 0.09$ in the beginning, half way through or at the end of the monitoring period.

In the primary report (Skalski and Ngouenet 2001), the analysis used an optimistic scenario, where the target improvement in survival was suddenly attained in 2000 and was sustained through years 2005 and 2008. Figure 1 presents different scenarios of gradual improvement with the RPA target reached by the end of a given period . For example, the worst-case scenario for statistical power to detect improvement is on the far right of Figure 1, where the improvement in survival reaches the RPA target in the last year of the recovery phase. At the far left side of Figure 1, a dashed line illustrates that the RPA target was reached in 2003, while in the middle, a dotted line assumes that the RPA target was reached in 2005.

Figure 1. Gradual improvement in survival of size Δ = 0.09 by the end of year 2003, 2005, or 2007.

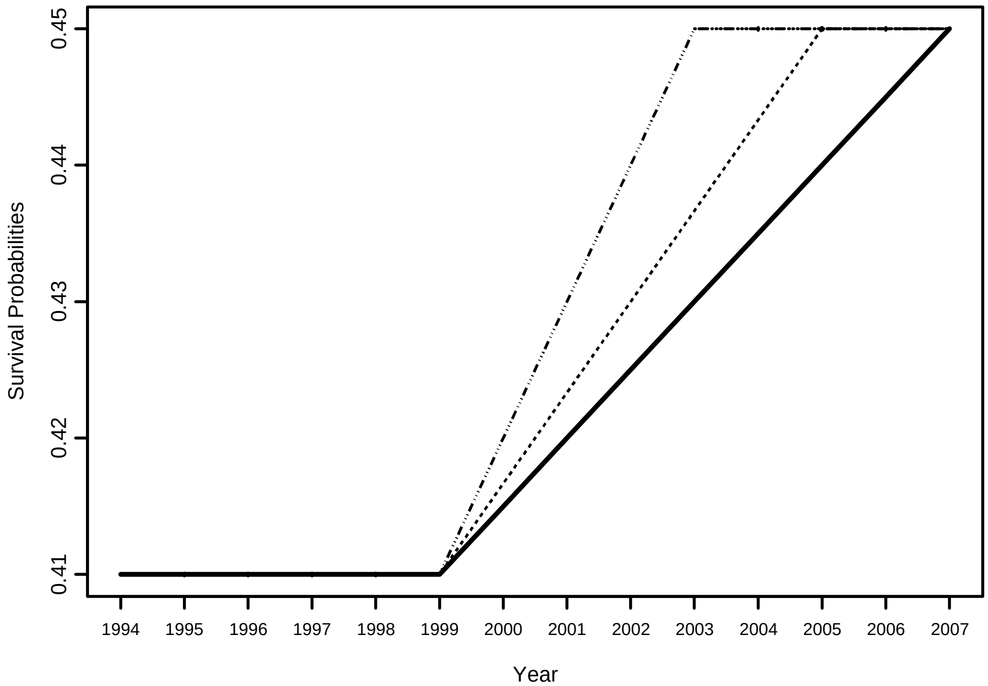
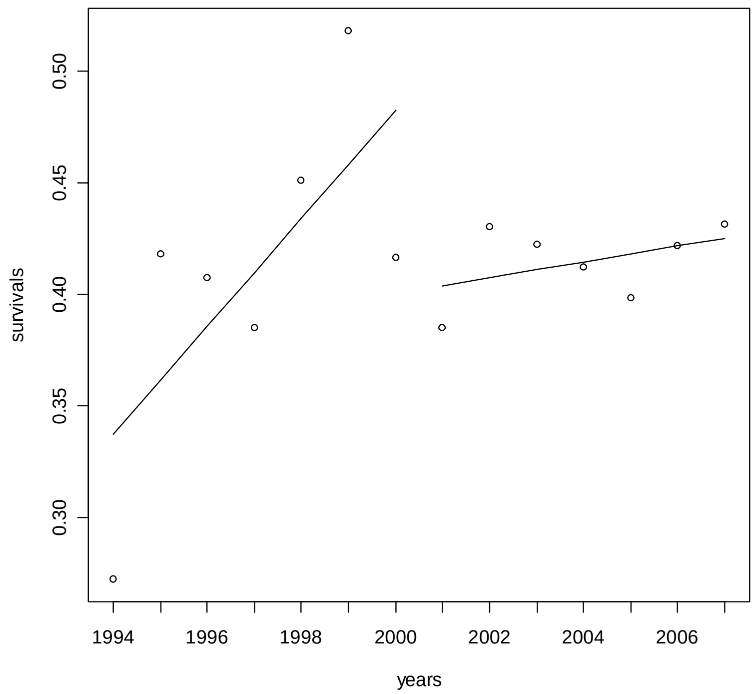


Figure 2. Two-phase regression for yearling chinook with survival improvement gradually reaching 9% in 2005.



To create the post-2000 data used in our simulations studies, we computed the expected values of the annual survival estimates as follows:

$$E_b\left(S_{\text{post}-2000,i}\right)=\left(1+\frac{0.09\times i}{n}\right)\times\bar{S}_{\text{pre}-2000}\ \text{ for } i=1,\dots,n \tag{3}$$

where

n = number of post-2000 years involved in the study (8 or 5),

$\bar{S}_{\text{pre-2000}}$ = mean survival estimate for the years 1994 or 1995 to 1999,

0.09 = anticipated 9% RPA improvement by the end of year 2005 (2008).

The annual survival probabilities are then generated using the expected values in Equation (3) plus a random error term ε where

$$S_{\text{post}-2000,i}=E_b\left(S_{\text{post}-2000,i}\right)+\varepsilon_i \tag{4}$$

where $\varepsilon \sim N\left(0, s^2\right)$ is a normal random variable with the variance s^2 equal to the pre-2000 survival inter-annual variance $Var\left(x_{\text{pre}-2000}\right)$.

The two-phase regression at year 2000 using a least-square model gives two sets of predicted survivals with the following means

$\bar{y}_{\text{pre-2000}}$ = mean predicted survivals for the years 1994 or 1995 to 1999,

$\bar{y}_{\text{post-2000}}$ = mean predicted survivals for the years 2000 to 2005 or 2008.

From the perspective of the BO, the null hypothesis of the first test assumes the RPA has been satisfied, unless there is evidence to the contrary. The null hypothesis of the second test assumes no improvement whatsoever, unless there is evidence to the contrary. As such, the two proposed tests of hypotheses juxtapose the nature of the statistical test. The apparent motivation of the two tests is to provide equal opportunity to conclude or reject the premise of recovery.

The tests of hypotheses can be based on two-sample t-tests of the form

$$t = \frac{\left(\bar{y}_{\text{post-2000}} - \bar{y}_{\text{pre-2000}}\right) - \left(\mu_{\text{post-2000}} - \mu_{\text{pre-2000}}\right)}{\sqrt{\widehat{Var}\left(\bar{y}_{\text{post-2000}} - \bar{y}_{\text{pre-2000}}\right)}}$$

that follows a t-distribution with $n_1 + n_2 - 2$ degrees of freedom.

Because the sample $\bar{y}_{\text{post-2000}}$ has not yet been collected, we shall assume for this analysis an equal inter-annual variance (i.e. s^2) during pre- and post-2000 years. Therefore,

$$\widehat{Var}\left(\bar{y}_{\text{post-2000}} - \bar{y}_{\text{pre-2000}}\right) = s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

for sample sizes n_1 and n_2 for the pre- and post-2000 periods, respectively.

Hence, the tests of hypotheses will be based on the t-statistic

$$t = \frac{\left(\bar{y}_{\text{post-2000}} - \bar{y}_{\text{pre-2000}}\right) - \Delta}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{5}$$

where $\Delta = \mu_{\text{post-2000}} - \mu_{\text{pre-2000}}$ under the null hypotheses. Test #1 specifies Δ = 0.09 while Test #2 specifies Δ = 0 for yearling chinook salmon between Lower Granite and Bonneville dams. The power of Test #1 is given by the probability that the test statistic falls in the rejection region:

$$P\left[t \le \frac{\Delta}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} - t_{\alpha,\mu}\right]$$

The power of Test #2 is given by:

$$P\left[t \le -\frac{\Delta - 0.09}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} - t_{\alpha,\mu}\right].$$

Power calculations were performed for the one-tailed hypotheses (1) and (2) at α = 0.05 and α = 0.10 using test statistic (5) for different values of Δ.

4.0 Results

In this section, plots of power curves are presented and a table of joint probabilities for Test #1 and Test #2 are given. The results of the power calculations are then interpreted in terms of the ability of the statistical tests to correctly identify the true state of recovery (i.e., $\Delta \le 0$ or $0 < \Delta > 0.09$, or $\Delta \ge 0.09$).

4.1 Yearling Chinook Salmon

Figure 3 presents the power of the two-phase regression test to reject the null hypothesis (1) of 0.09 improvement or greater in survival between Lower Granite and Bonneville dams. When $\Delta = 0$, the Test #1 has 78% chance of rejecting (1) at α = 0.05.

Figure 4 presents power of the two-phase regression to reject the null hypothesis (2) of no improvement in survival between Lower Granite and Bonneville dams for yearling chinook salmon. At $\Delta \approx 0.09$, the test has approximately a 73% chance of rejecting (2) at α = 0.05. A 0.15 improvement in survival between periods is needed before the t-test is almost certain to reject the null hypothesis of no improvement (2). These results show a moderate improvement in statistical power compared to the basic two-sample t-test in (Skalski and Ngouenet 2001)

The results of the power curves for the two-phase regression in Figures 3-4 are summarized in Tables 5 and 6, respectively. By design, Test #1 will make an incorrect decision $\alpha \cdot 100\%$ of the time and conclude $\Delta < 0.09$ when impact recovery has occurred with $\Delta \ge 0.09$. However, Test #1 will make an incorrect decision between 27%-95% of the time and conclude

$\Delta > 0.09$ when in fact $0 \le \Delta < 0.09$. The mean survival during post-2000 years can be less than pre-2000 years, and Test #1 has a $\le 27\%$ chance of concluding $\Delta > 0.09$ (Table 5). By design, Test #2 will make an incorrect decision $\alpha \cdot 100\%$ of the time and conclude $\Delta > 0$ when, in fact, $\Delta \le 0$. However, Test #2 will make an incorrect decision between 27%-95% of the time and conclude $\Delta < 0$ when, in fact, $0 \le \Delta < 0.09$. Test #2 has a $\ge 73\%$ chance of making the correct decision when the improvement in survival ≥ 0.09.

The results of the power calculations for Tests 1 and 2 are summarized under alternative scenarios for recovery by the year 2008 in Tables 5 and 6. The ideal results for Tests #1 and #2 would be to have probabilities of correct decisions near 1 in the shaded boxes and probabilities of incorrect decision near 0 in the unshaded boxes in Tables 5 and 6, respectively. Deviations from these expectations are a measure of the lack of performance of the proposed tests of compliance. Table 7 summarizes the probabilities of making the correct decisions with Tests 1-2 under alternative states of nature. The chance of both Tests 1 and 2 making the correct decision for the yearling chinook stock when $\Delta < 0$ is $\ge 69\%$ of the time. The chance is $\ge 69\%$ that Tests 1 and 2 will both make the correct decision when $\Delta > 0.09$. There is 0.25%-53% chance of the correct decision for both tests when $0 < \Delta < 0.09$ by the year 2008 (Table 7). That is a 53% probability of making a correct decision for a 9% RPA improvement.

Figure 3. Power of Test #1 using two-phase regression for yearling chinook salmon based on inriver smolt survival between Lower Granite and Bonneville dams for 8 years of RPA (2001-2008). The simulation is based on a progressive improvement reaching 9% in 2006.

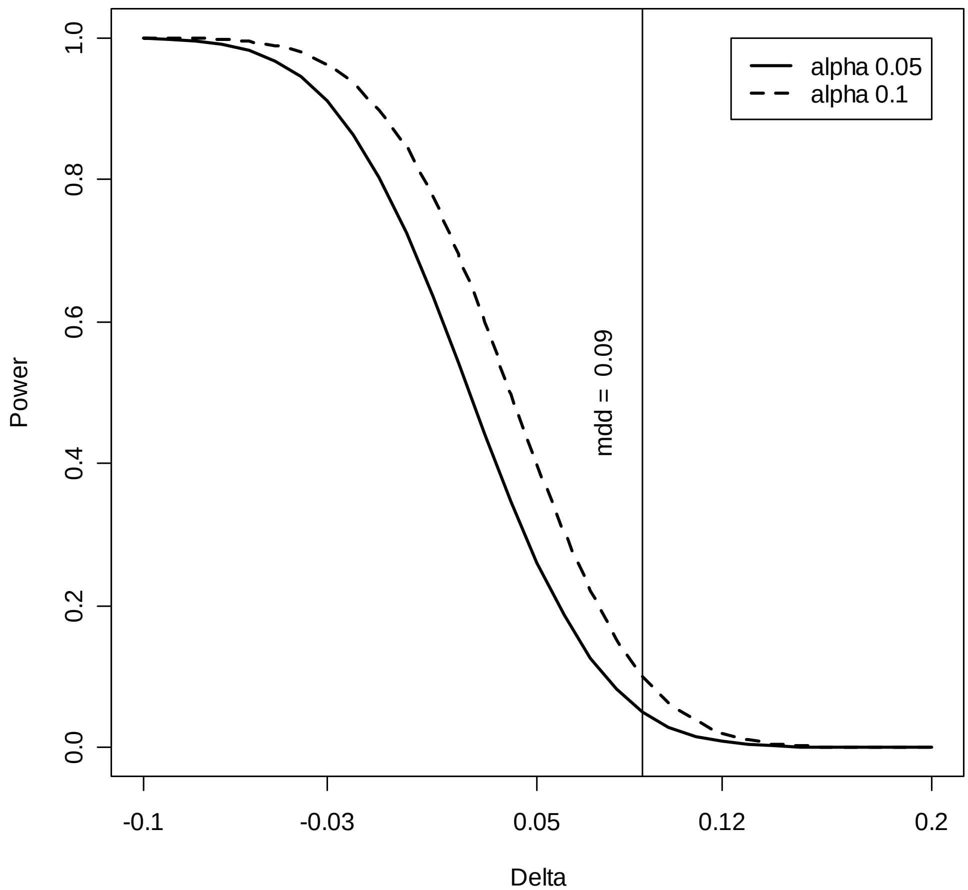


Figure 4. Power of Test #2 using two-phase regression for yearling chinook salmon based on inriver smolt survival between Lower Granite and Bonneville dams for 8 years of RPA (2001-2008). The simulation is based on a progressive improvement reaching 9% in 2006.

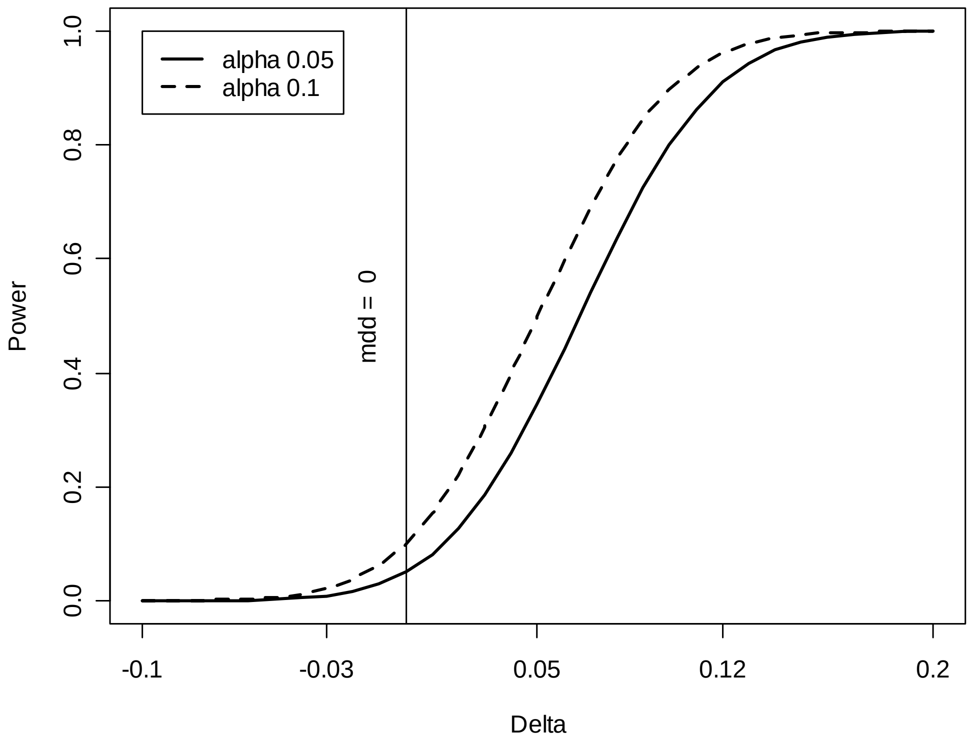


Table 5. Probabilities of marking correct (shaded) and incorrect (unshaded) decisions using two-phase regression and Test #1 at α = 0.05 for yearling chinook salmon for the survival estimates from Lower Granite to Bonneville dams under different states of nature by 2008.

	Alternative States of Nature		
	No Improvement $\Delta \leq 0$	Some Improvement $0 < \Delta < 0.09$	Recovery $\Delta \geq 0.09$
Conclude $\Delta \geq 0.09$	$0 < \beta < 0.27$	$0.27 < \beta < 0.95$	$1 - \alpha = 95$
Conclude $\Delta < 0.09$	$0.73 < 1 - \beta < 1.0$	$0.05 < 1 - \beta < 0.73$	$\alpha = 0.05$

Table 6. Probabilities of making correct (shaded) and incorrect (unshaded) decisions using two-phase regression and Test #2 at α = 0.05 for yearling chinook salmon for the survival estimates from Lower Granite to Bonneville dams under different states of nature by 2008.

	Alternative States of Nature		
	No Improvement $\Delta \leq 0$	Some Improvement $0 < \Delta < 0.09$	Recovery $\Delta \geq 0.09$
Conclude $\Delta \leq 0$	$1 - \alpha = 0.95$	$0.27 < \beta < 0.95$	$\beta < 0.27$
Conclude $\Delta > 0$	$\alpha = 0.05$	$0.05 < 1 - \beta < 0.73$	$1 - \beta \geq 0.73$

Table 7. Probabilities Tests #1 and #2 using two-phase regression will make the correct decisions, individually and jointly, under alternative states of nature at α = 0.05 for the yearling chinook salmon by 2008.

	Alternative States of Nature		
	$\Delta \le 0$	$0 < \Delta < 0.09$	$\Delta \ge 0.09$
Test #1	Reject H_0 $0.73 \le 1-\beta \le 1.0$	Reject H_o $0.05 < 1-\beta < 0.73$	Do not reject H_o $1-\alpha = 0.95$
Test #2	Do not reject H_o $1-\alpha = 0.95$	Reject H_o $0.05 < 1-\beta < 0.73$	Reject H_o $0.73 \le 1-\beta \le 1.0$
Joint Tests #1 and #2	0.69 – 0.95	0.0025 – 0.53	0.69 – 0.95

4.2 Subyearling Chinook Salmon

Figure 5 presents the power of the two-phase regression test to reject the null hypothesis of no improvement in survival (1) between Lower Granite and Bonneville dams. When Δ = 0, Test #1 has 100% chance of rejecting (1) at α = 0.05.

Figure 6 presents the power of the two-phase regression test to reject the null hypothesis of no improvement for survival (2) between Lower Granite and Bonneville dams for subyearling chinook salmon. At $\Delta \approx 0$, the two-phase regression test has 100% chance of rejecting (2) at α = 0.05. At Δ = 0.05, improvement in survival between periods had 85% chance for the two-phase regression test to reject the null hypothesis of no improvement (2).

The results of the power curves in Figures 5-6 are summarized in Tables 8 and 9, respectively. By design, Test #1 will make an incorrect decision $\alpha \cdot 100\%$ of the time and conclude $\Delta < 0.09$ when impact recovery has occurred with $\Delta \ge 0.09$. However, Test #1 will make an incorrect decision less than 95% of the time and conclude $\Delta > 0.09$ when in fact $0 \le \Delta < 0.09$. The mean survival during post-2000 years can be less than pre-2000 years, and Test #1 has virtually no chance of concluding $\Delta > 0.09$ (Table 8). By design, Test #2 will make an incorrect decision $\alpha \cdot 100\%$ of the time and conclude $\Delta > 0$ when, in fact, $\Delta \le 0$. However, Test #2 will make an incorrect decision less than 95% of the time and conclude $\Delta < 0$ when, in fact, $0 < \Delta < 0.09$. Test #2 had 100% chance of making the correct decision when the improvement in survival ≥ 0.09.

The result of the power calculations for Tests #1 and #2 are summarized under alternative scenarios for recovery by the year 2008 for subyearling chinook salmon in Tables 8 and 9. Table 10 summarizes the probabilities of making the correct decisions with Test 1-2 under alternative states of nature for improvement in survival of the subyearling chinook salmon. The chance of both Tests 1 and 2 making the correct decision for the subyearling chinook stock when $\Delta < 0$ is $\geq$95% of the time. The chance is $\geq$95% that Tests 1 and 2 will both make the correct decision when $\Delta > 0.09$. There is 0.25%-100% chance of the correct decision for both tests when $0 < \Delta < 0.09$ by the year 2008 (Table 10). There is quasi-certainty of making a correct decision for a 9% RPA improvement or greater. These results show some increase in statistical power with the two-phase regression testing model over the classical two-sample t-test reviewed in Skalski and Ngouenet (2001) and presented in the BO.

Figure 5. Power of Test #1 using the two-phase regression for subyearling chinook salmon based on inriver smolt survival between Granite and Bonneville dams for 8 years of RPA (2001-2008). The simulation is based on a progressive improvement reaching 9% in 2006.

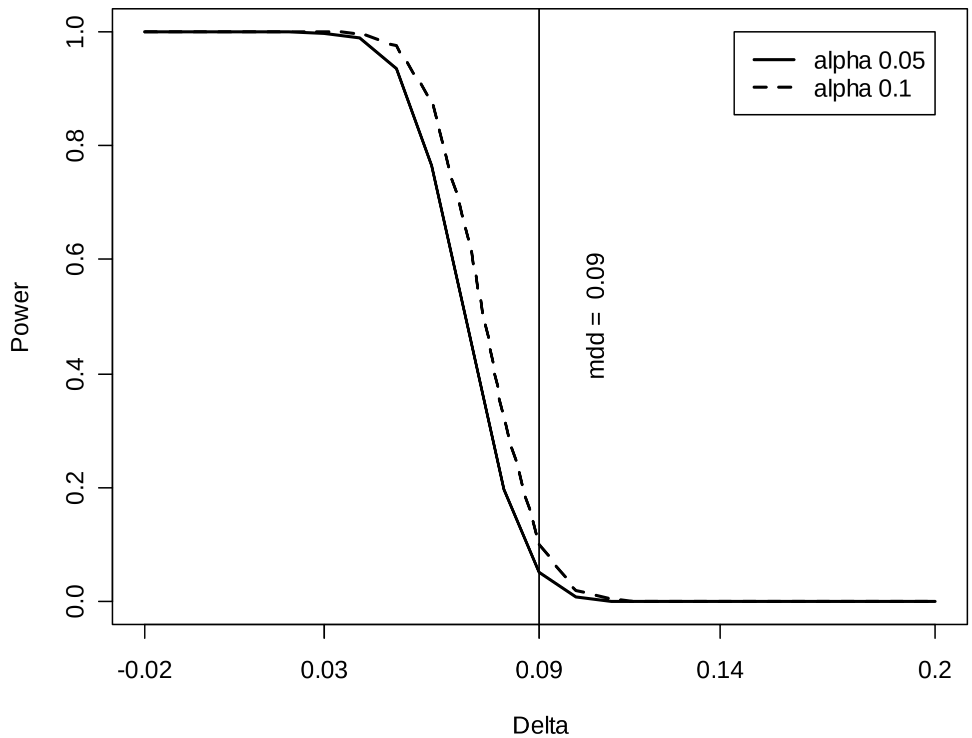


Figure 6. Power of Test #2 using two-phase regression for subyearling chinook salmon based on inriver smolt survival between Lower Granite and Bonneville dams for 8 years of RPA (2001-2008). The simulation is based on a progressive improvement reaching 9% in 2006.

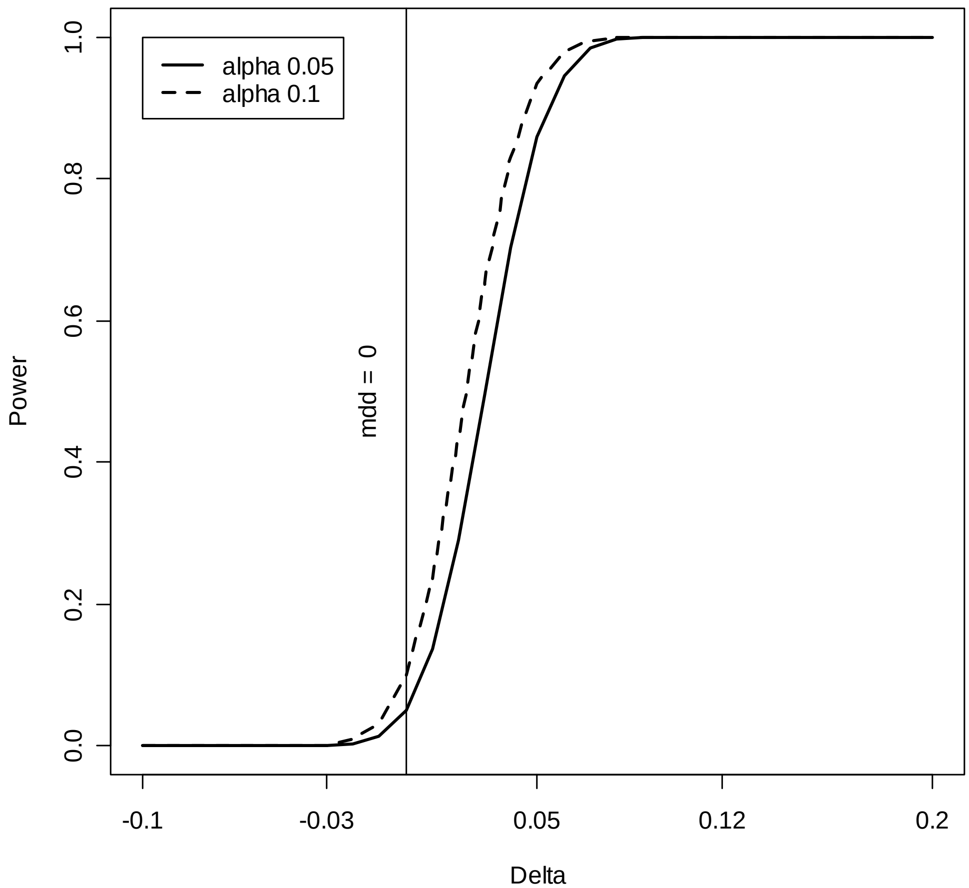


Table 8. Probabilities of marking correct (shaded) and incorrect (unshaded) decisions using two-phase regression and Test #1 at α = 0.05 for subyearling chinook salmon for the survival estimates from Lower Granite to Bonneville dams under different states of nature by 2008.

	Alternative States of Nature		
	No Improvement $\Delta \leq 0$	Some Improvement $0 < \Delta < 0.09$	Recovery $\Delta \geq 0.09$
Conclude $\Delta \geq 0.09$	$\beta = 0.0$	$0.0 < \beta < 0.95$	$1-\alpha = 95$
Conclude $\Delta < 0.09$	$1-\beta = 1.0$	$0.05 < 1-\beta < 1.0$	$\alpha = 0.05$

Table 9. Probabilities of making correct (shaded) and incorrect (unshaded) decisions using two-phase regression and Test #2 at α = 0.05 for subyearling chinook salmon for the survival estimates from Lower Granite to Bonneville dams under different states of nature by 2008.

	Alternative States of Nature		
	No Improvement $\Delta \leq 0$	Some Improvement $0 < \Delta < 0.09$	Recovery $\Delta \geq 0.09$
Conclude $\Delta \leq 0$	$1-\alpha = 0.95$	$0.0 < \beta < 0.95$	$\beta = 0.0$
Conclude $\Delta > 0$	$\alpha = 0.05$	$0.05 < 1-\beta < 1.0$	$1-\beta = 1.0$

Table 10. Probabilities Tests #1 and #2 using two-phase regression will make the correct decisions, individually and jointly, under alternative states of nature at α = 0.05 for the subyearling chinook salmon for 2008.

	Alternative States of Nature		
	$\Delta \le 0$	$0 < \Delta < 0.09$	$\Delta \ge 0.09$
Test #1	Reject H_0 $1-\beta = 1.0$	Reject H_o $0.05 < 1-\beta < 1.0$	Do not reject H_o $1-\alpha = 0.95$
Test #2	Do not reject H_o $1-\alpha = 0.95$	Reject H_o $0.05 < 1-\beta < 1.0$	Reject H_o $1-\beta = 1.0$
Joint Tests #1 and #2	0.95	0.0025 – 1.0	0.95

4.3 Steelhead

Figure 7 presents the power of the two-phase regression test to reject the null hypothesis of 0.09 improvement or greater in survival (1) for steelhead between Lower Granite and Bonneville dams. When $\Delta \approx 0$, Test #1 had 100% chance to reject (1) at α = 0.05.

Figure 8 presents the power of the two-phase regression test to reject the null hypothesis of no improvement in survival (2) between Lower Granite and Bonneville dams for steelhead. At $\Delta \approx 0.10$, the t-test is almost certain to reject (2) at α = 0.05. A Δ = 0.05 improvement in survival between periods also has a 100% chance for the two-phase regression test to reject the null hypothesis of no improvement (2).

The results of the power curves in Figures 7-8 are summarized in Tables 11 and 12, respectively. By design, Test #1 will make an incorrect decision $\alpha \cdot 100\%$ of the time and conclude $\Delta < 0.09$ when impact recovery has occurred with $\Delta \ge 0.09$. The mean survival during post-2000 years can be less than pre-2000 years, and Test #1 has no power of concluding $\Delta > 0.09$ (Table 11). By design, Test #2 will make an incorrect decision $\alpha \cdot 100\%$ of the time and conclude $\Delta > 0$ when, in fact, $\Delta \le 0$. Test #2 is almost certain to make the correct decision when the improvement in survival ≥ 0.09.

The results of the power calculations for Tests #1 and #2 are summarized under alternative scenarios for recovery of steelhead by the year 2005 and 2008 in Tables 11 and 12. Table 13 summarizes the probabilities of making the correct decisions with Test 1-2 under alternative states of nature for improvement in survival of the juvenile steelhead. The chance of both Tests 1 and 2 making the correct decision for the steelhead stock when $\Delta < 0$ is $\geq 95\%$ of the time. The chance is $\geq 95\%$ that Tests 1 and 2 will both make the correct decision when $\Delta > 0.09$. There is 0.25%-100% chance of the correct decision for both tests when $0 < \Delta < 0.09$ for the year 2008 (Table 13). These results show a good increase in statistical power with the two-phase regression over the classical two-sample t-test proposed in the BO.

Figure 7. Power of Test #1 for steelhead based on inriver smolt survival between Lower Granite and Bonneville dams for 8 years of RPA (2001-2008). The simulation is based on a progressive improvement reaching 9% in 2006.

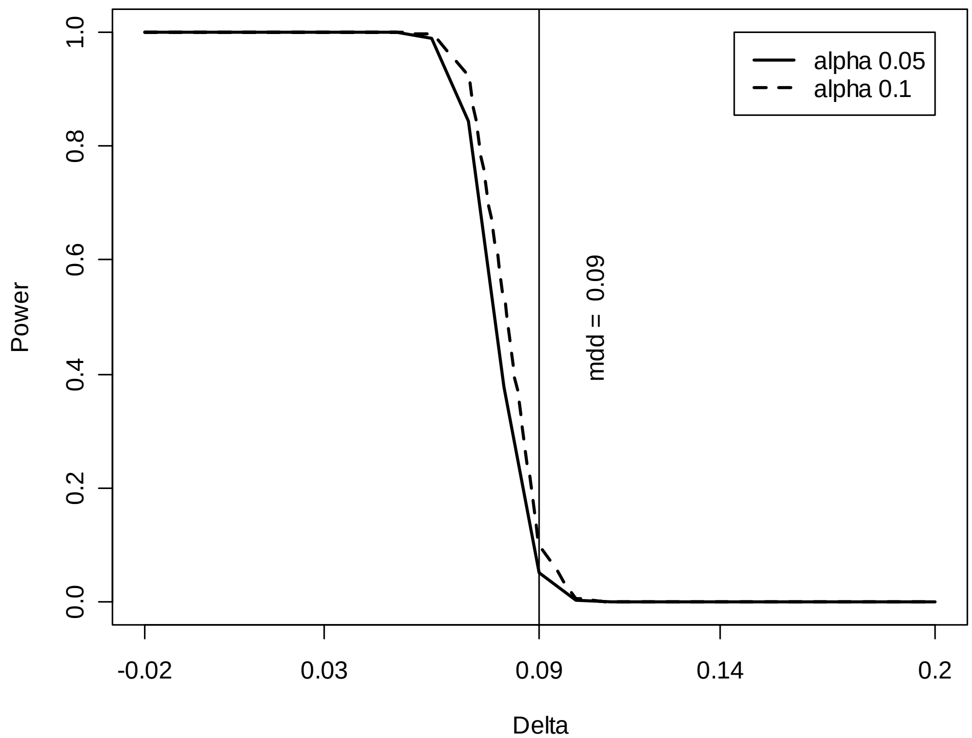


Figure 8. Power of Test #2 for steelhead based on inriver smolt survival between Granite and Bonneville dams for 8 years of RPA (2001-2008). The simulation is based on a progressive improvement reaching 9% in 2006.

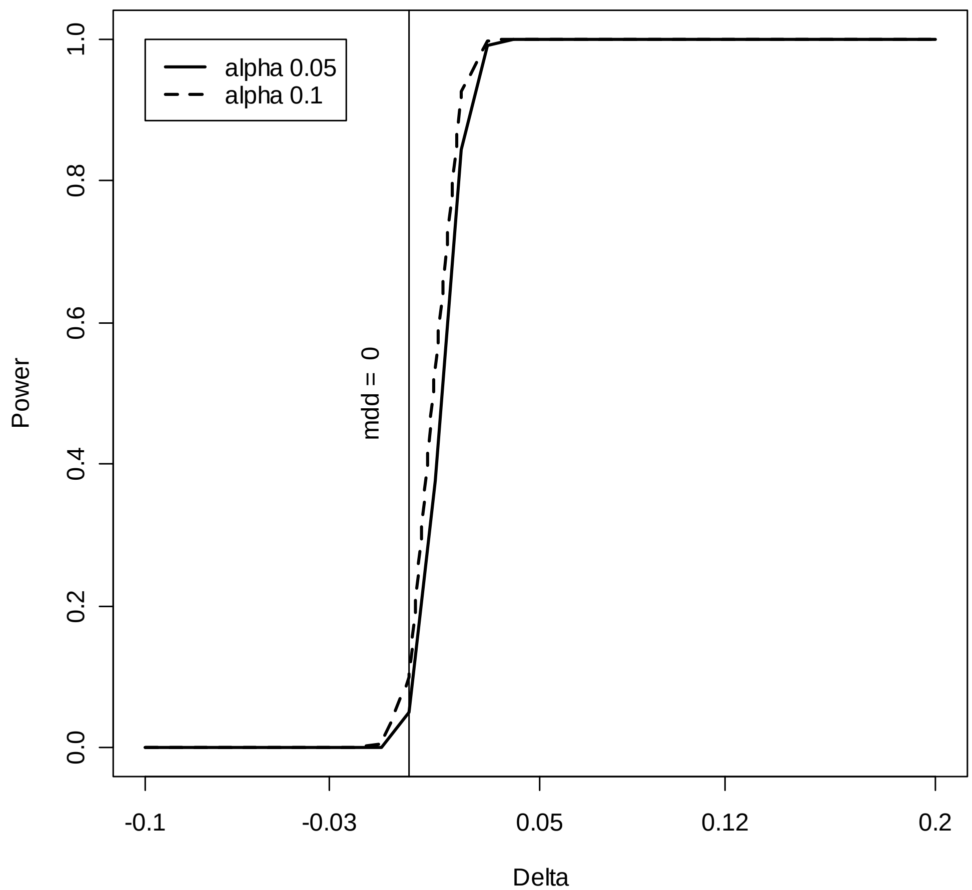


Table 11. Probabilities of Probabilities of marking correct (shaded) and incorrect (unshaded) decisions using Test #1 at α = 0.05 for steelhead for the survival estimates from Lower Granite to Bonneville dams under different states of nature by 2008.

	Alternative States of Nature: No Improvement $\Delta \le 0$	Some Improvement $0 < \Delta < 0.09$	Recovery $\Delta \ge 0.09$
Conclude $\Delta \ge 0.09$	$\beta = 0$	$0.0 < \beta < 0.95$	$1 - \alpha = 95$
Conclude $\Delta < 0.09$	$1 - \beta = 1.0$	$0.05 < 1 - \beta < 1.0$	$\alpha = 0.05$

Table 12. Probabilities of making correct (shaded) and incorrect (unshaded) decisions using Test #2 at α = 0.05 for steelhead for the survival estimates from Lower Granite to Bonneville dams under different states of nature by 2008.

	Alternative States of Nature: No Improvement $\Delta \le 0$	Some Improvement $0 < \Delta < 0.09$	Recovery $\Delta \ge 0.09$
Conclude $\Delta \le 0$	$1 - \alpha = 0.95$	$0.0 < \beta < 0.95$	$\beta = 0$
Conclude $\Delta > 0$	$\alpha = 0.05$	$0.05 < 1 - \beta < 1.0$	$1 - \beta = 1.0$

Table 13. Probabilities Tests #1 and #2 will make the correct decisions, individually and jointly, under alternative states of nature at α = 0.05 for the steelhead by 2008.

	Alternative States of Nature		
	$\Delta \le 0$	$0 < \Delta < 0.09$	$\Delta \ge 0.09$
Test #1	Reject H_0 $1-\beta = 1.0$	Reject H_o $0.05 < 1-\beta < 1.0$	Do not reject H_o $1-\alpha = 0.95$
Test #2	Do not reject H_o $1-\alpha = 0.95$	Reject H_o $0.05 < 1-\beta < 1.0$	Reject H_o $1-\beta = 1.0$
Joint Tests #1 and #2	0.95	0.0025 – 1.0	0.95

5.0 Discussion

The two-phase regression model improves substantially the statistical power over the standard two-sample t-test suggested in the Federal Columbia River Power System 2000 Biological Opinion. In addition, the two-phase regression model has a very high probability of correctly assessing the true status of the recovery.

The statistical power calculation results were interpreted in terms of the ability of the two-phase regression tests to correctly identify the true states of recovery (i.e., fail or succeed in fulfilling RPA expectations). For these selected stocks, Table 14 summarizes the probabilities of jointly making the correct decisions with two-phase regression Tests #1 and #2 versus the two-sample t-tests #1 and #2. The chance of both Tests 1 and 2 to both identify recovery for the yearling chinook is >69% for a 9% RPA improvement. For the subyearling chinook salmon and steelhead, the step-regression approach has a high certainty of correctly identifying recovery, $1-\beta = 0.95$.

The comparison in statistical power is much more dramatic between the alternative test procedures than indicated in Table 14. For the two-sample t-testing, the Monte Carlo simulations modeled recovery as occurring immediately post-2000 and sustaining itself throughout the recovery period (2001-2008). In contrast, the Monte Carlo simulations used to model recovery for the step-regression was a gradual recovery reaching a 9% improvement beginning in 2006 (Figure 9). We believe the scenario of instant recovery is unrealistic. In addition, the two-sample t-test would have much lower statistical power to detect recovery should it occur gradually over time as assumed in the Monte Carlo simulations for the step-regression. Hence, the differences in statistical power of the two test procedures are much grater than reported in Table 14. Decision rules to assess recovery need to consider both the nature of the proposed recovery as well as the statistical behavior of the data and the test statistics. The two-sample t-tests proposed in the BO are among the least informative of the options available. The step-regression methods allow for a more dynamic and realistic expectation for the recovery process. However, they still do not incorporate ancillary information or informed priors to improve the decision process. The nature of the opposing tests of hypotheses (1 and 2), as stated in the BO, has the unfortunate effect of producing relatively low power to correctly identify the state of incomplete recovery (i.e., $0 < \Delta < 0.09$). Only reformulating the decision rules can help improve the statistical power in this intermediate state of nature of some but not full recovery.

Table 14. Comparison of the statistical power of the BO proposed two-sample t-test versus step-regression in identifying the correct state of nature concerning the change in survival (Δ) pre- and post-2000 between Lower Granite and Bonneville dams for (a) yearling chinook salmon, (b) subyearling chinook salmon, and (c) steelhead.

a. Yearling chinook salmon

	$\Delta \leq 0$	$0 < \Delta < 0.09$	$\Delta \geq 0.09$
BO two-sample t-test	0.59-0.95	0.0025-0.38	0.59-0.95
Step-regression	0.69-0.95	0.0025-0.53	0.69-0.95

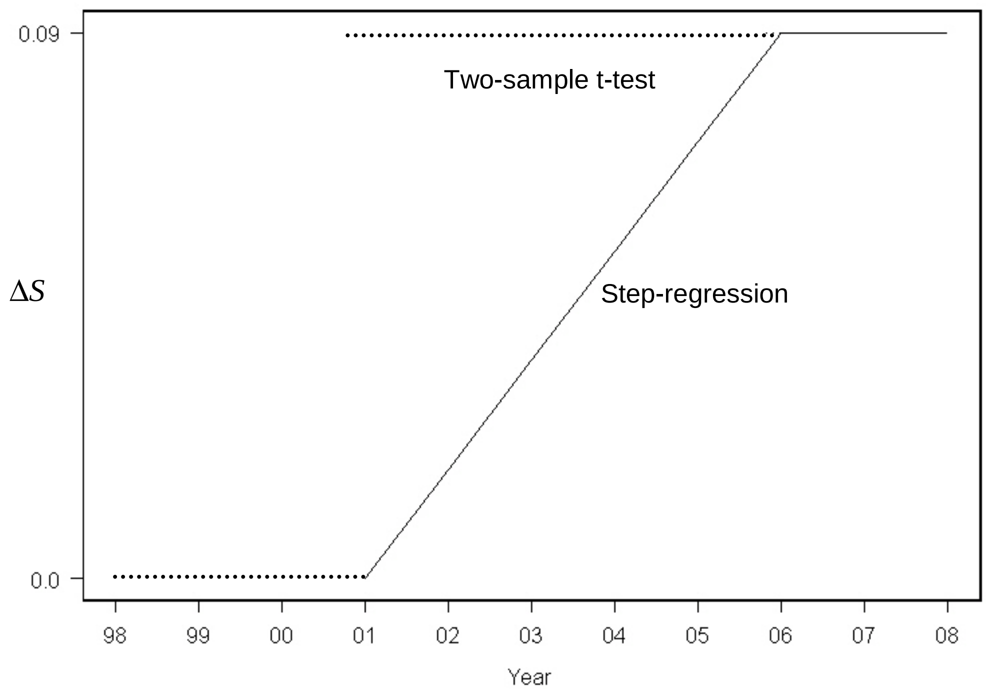
b. Subyearling chinook salmon

	$\Delta \leq 0$	$0 < \Delta < 0.09$	$\Delta \geq 0.09$
BO two-sample t-test	0.94-0.95	0.0025-0.61	0.94-0.95
Step-regression	0.95	0.0025-1.0	0.95

c. Steelhead

	$\Delta \leq 0$	$0 < \Delta < 0.09$	$\Delta \geq 0.09$
BO two-sample t-test	0.84-0.95	0.0025-0.79	0.84-0.95
Step-regression	0.95	0.0025-1.0	0.95

Figure 9. Schematic of the two recovery scenarios used to simulate recovery for the two-sample t-tests (instant recovery) and the step-regression (gradual recovery).



In the next phase of this project, Bayesian methods will be addressed. Bayesian analysis (Bernardo and Smith 1994) can complement classical statistics in situations where uncertainty must be taken into account. These models offer a useful framework for decision analysis by incorporating prior knowledge into the decision process and to help achieve a precautionary approach to compliance evaluation. The analysis incorporates scientific/biological knowledge/expertise by imposing a data-independent distribution on the parameters of the selected model; the analysis thus consists of formally combining both the prior distribution on the parameters and the collected data to jointly make inferences and/or test assumptions about the model parameters. The goal of this further investigation is to enhance the chances of correctly identifying the true state of recovery. Neither the two-sample t-test nor step-regression procedure has an excellent chance of correctly identifying improvement but not full recovery. Only by reformulating the decision rules can there be reasonable expectation of better decision making.

6.0 Literature Cited

Beckman, R. J., and R. D. Cook. 1979. Testing for two-phase regression. *Technometrics* 21: 65-69.

Bernardo, J. M., and A. F. Smith. 1994. Bayesian theory. John Wiley & Sons. New York, NY.

Hinkley, D. V. 1971. Inference in two-phase regression. *Journal of the American Statistical Association* 66:736-743.

National Marine Fisheries Service. 2000. 2000 Federal Columbia River Power System (FCRPS) Biological Opinion. Reinitiation of consultation on operation of the Federal Columbia River Power System, including the juvenile fish transportation program, and 19 Bureau of Reclamation projects in the Columbia Basin. Consultation conducted by the National Marine Fisheries Service, Northwest Region.

Skalski, J. R., and R. F. Ngouenet. 2001. Evaluation of the compliance testing framework for RPA improvement as stated in the 2000 Federal Columbia River Power System (FCRPS) Biological Opinion. Volume VII in the Monitoring and Evaluation of Smolt Migration in the Columbia Basin. US Department of Energy, Bonneville Power Administration, Portland, Oregon.

Searle, S. R. 1971. Linear models. John Wiley & Sons Ltd. New York, NY.