Appendix

VAKI Riverwatcher Data Estimation Procedures

Introduction

Infrared-imaging technology has been used to monitor fish passage at Daguerre Point Dam in the lower Yuba River since 2004 using Vaki Riverwatcher systems to document specific observations addressing VSP parameters of adult abundance and diversity. The Vaki Riverwatcher infrared systems produced by Vaki Aquaculture Systems Ltd., of Iceland, provide a tool for monitoring fish passage year-round without need for continuous video feeds. The Vaki Riverwatcher system records both silhouettes and electronic images of each fish passage event. By capturing silhouettes and images, fish passage can be accurately monitored even under turbid conditions.

The Vaki Riverwatcher systems were able to record and identify the timing and magnitude of passage for multiple species at Daguerre Point Dam during most temporal periods, although system failures reduced the ability of the equipment to document ladder use during some months. Vaki system non-operation events were classified by one of three categories; low-voltage disconnections (LVD), system maintenance or unknown malfunctions.

Most system failures were caused by LVD. LVD events occurred when the electrical demands of the Vaki Riverwatcher systems exceeded photovoltaic power generation and/or storage (e.g. system voltage dropped below 11.7 volts). The units were also occasionally disconnected for maintenance by fishery technicians (e.g. battery recharging, camera lens cleaning, etc.). Other malfunctions were observed in which no direct explanations for system disconnect could be diagnosed.

LVD often affected system operation during the winter months as a result of low photovoltaic power generation and a lack of capacity to store sufficient power to bridge periods of low photoperiod. In contrast, LVD were observed less frequently during the fall months (September – November), but other unidentified malfunctions resulted in system downtime during this period. Although the definitive causes of these unidentified system malfunctions were unknown, the periods of non-operation were suspected to be the result of data processing limitations. Multiple sustained passage events that coincided with peak fall-run Chinook salmon immigrations are thought to have exceeded the system’s data processing capabilities. These unknown malfunctions ultimately resulted in multiple lapses of data continuity.

Objective

The objective of this technical appendix is to develop and apply a process to estimate counts of anadromous salmonids passing through the Vaki Riverwatcher systems located in the North and South ladders at DPD when the Vaki Riverwatcher systems were not fully operational. This objective incorporates the following considerations.

- For each biological year (March 1 through February 28), series of daily counts of Chinook salmon and steelhead moving upstream of Daguerre Point Dam are obtained for each of the ladders as the daily difference, whenever positive, of the observations of fish moving up the ladder minus those moving down the ladders.

- Independent measures of the operational status of the Vaki Riverwatcher systems exist as series of the proportion of hours the Vaki Riverwatcher systems operated at each ladder during each day of each biological year.

- The annual series of the operational status of the Vaki Riverwatcher systems indicate that these systems were not always operating 24 hours each day, and sometimes did not operate during various consecutive days, and that the daily operational status of the Vaki Riverwatcher systems was often different between ladders.

- Given the above conditions, a method was developed to correct the daily counts of Chinook salmon and steelhead moving upstream of Daguerre Point Dam for those days when the Vaki Riverwatcher systems were not fully operational.

Daily Count Estimation Approach

The method to estimate the daily counts of Chinook salmon and steelhead moving upstream of Daguerre Point Dam for those days when the Vaki Riverwatcher systems was not fully operational consists of the following steps.

- For each biological year (March 1 through February 28) and ladder, the series of observed daily counts of Chinook salmon or steelhead moving upstream of Daguerre Point Dam (response variable Y) are used to fit a smooth function that describes the expected number of fish moving upstream of Daguerre Point Dam each day of the biological year.

- The days within the biological year are the only predictor (d), with values ranging from d = 1 for March 1 to d = 365 for February 28, or d = 366 for February 29.

- The series of proportion of hours the Vaki Riverwatcher systems operated at each ladder during each day of each biological year are used as weights in the minimization process used to obtain the smooth function.

- Once the smooth function that best fit the data is obtained, the values predicted by the smooth function on days with proportion of hours of operating Vaki Riverwatcher systems different from one are rounded to the nearest integer and compared with the corresponding observed count.

- If the prediction is higher than the observed count, the prediction is accepted as corrected count.

- If the prediction is lower than the observed count the prediction is discarded and the observed count is considered the corrected count.

- Once that this process is finished for the series of daily counts of the north ladder and that of the south ladder, the resulting corrected series are combined in one series of corrected daily counts of Chinook salmon or steelhead moving upstream of Daguerre Point Dam by summing the corresponding daily values.

Prior to undertaking the data estimation procedures, the first steps are to: (1) conduct an autocorrelation analysis of the daily count data to determine whether the datasets are random series, or are characterized by short-term correlation; and (2) whether the datasets are non-randomly distributed and therefore appropriately addressed through the fitting of a smoothing function.

Autocorrelation Analysis

The autocorrelation analysis of the daily count data of Chinook salmon passing Daguerre Point Dam was conducted separately for each ladder (North or South) for each biological year (March through February) for the period extending from March 2004 through February 2011. The analysis was conducted by lagging the dates of the count data from 0 through 25 days. Results of the autocorrelation analysis are presented in **Attachment 1**.

Results indicate that the time series of daily count data are not random because the correlation coefficients for all non-zero values for most lagging values do not approximate 0 (and are significantly different from 0). By contrast to a random distribution, the data are indicative of stationary series exhibiting short-term correlation, with the highest correlation coefficients associated with a lag value of one day.

Non-Parametric Test for Randomness

The non-parametric U test for runs above and below the median (Freund and Simon 1991) was used to test for randomness of the time series of the daily count data of Chinook salmon passing Daguerre Point Dam was conducted separately for each ladder (North or South) for each biological year (March through February) for the period extending from March 2004 through February 2011. This test is a statistical method that uses ranked values relative to a test value. The test statistic R is used to calculate a Z value that is then compared to a tabulated value associated with the appropriate degrees of freedom.

Results indicate that the time series of the daily count data of Chinook salmon passing Daguerre Point Dam are not random, for either ladder (North or South) for any of the biological years (March through February) extending from March 2004 through February 2011 (**Table 1**).

Table 1. Results of the non-parametric U test (Freund and Simon 1991) for the time series of the daily count data of Chinook salmon passing Daguerre Point Dam separately for each ladder (North or South) for each biological year (March through February) for the period extending from March 2004 through February 2011.

Ladder	North					South				
Biological Year	R (runs)	Rμ	Rσ	Z	P(Z)	R (runs)	Rμ	Rσ	Z	P(Z)
2004-2005	35	181.20	9.4187	-15.522	1.23E-54	39	103.94	5.3691	-12.095	5.64E-34
2005-2006	31	183.43	9.5358	-15.985	8.09E-58	65	172.65	8.9706	-12.000	1.77E-33
2006-2007	91	183.39	9.5335	-9.691	1.65E-22	72	120.18	6.2204	-7.745	4.77E-15
2007-2008	109	183.99	9.5522	-7.851	2.06E-15	79	132.58	6.8612	-7.810	2.86E-15
2008-2009	101	182.78	9.5014	-8.607	3.76E-18	107	133.53	6.9200	-3.833	6.32E-05
2009-2010	71	183.01	9.5134	-11.773	2.68E-32	83	117.80	6.0957	-5.709	5.68E-09
2010-2011	75	183.33	9.5307	-11.367	3.06E-30	55	171.14	8.8914	-13.062	2.72E-39

Smooth Function Estimation

The following components describe the estimation method used to obtain the fitted smooth function that describes the expected number of fish moving upstream of each ladder of Daguerre Point Dam each day of a biological year, in the above mentioned estimation approach.

- The function fitted to the data (the pairs of daily count and day within the biological year) is a spline. A spline is a special function defined piecewise by polynomials, often of small degree, that is continuous along the predictor range and n-1 times derivable (where n is the sample size).

- The fitted spline is obtained through a process termed smoothing spline. The smoothing spline is a method of smoothing (fitting a smooth curve to a set of "noisy" observations) where spline fitted or expected values are defined to be the minimizer (over the class of twice differentiable functions) of the following expression:

$$\sum_{i=1}^{n}(Y_i - \hat{\mu}(x_i))^2 + \lambda \int_{x_1}^{x_n} \hat{\mu}''(x)^2\, dx.$$

  where the summation term describes the fidelity of the fitted spline ($\hat{\mu}(x_i)$) to the data, and the second term is a penalization term controlled by the smoothing parameter λ. The parameter λ is a value greater or equal than 0 that controls the trade-off between fidelity to the data and roughness of the function estimate.

- The daily count data of Chinook salmon passing Daguerre Point Dam for each ladder (North or South) and each biological year (March through February) from March 2004 through February 2011 was used as the response variable in the fitting of individual splines, with the days within the biological year as the only predictor (d), and the corresponding series of proportion of hours the Vaki Riverwatcher systems operated at each ladder during each day of each biological years, used as weights in the minimization process.

- The minimization process was carried out using the S-Plus Generalized Additive Model (GAM) module. The GAM is a statistical model for blending properties of generalized linear models with additive models. In GAM, a distribution (e.g., a normal distribution, a

Poisson distribution or a binomial distribution) and a link function g needs to be specified to relate the expected value of the distribution of the response variable to the predictors.

- To fit different smoothed splines to particular data series through the S-Plus GAM module, the response variable (the daily counts at the North or South ladders) was assumed to have a Poisson distribution with the logarithm as link function. A total of 13 distinct smoothed splines varying in the roughness of the resulting splines were obtained by varying the smoothing parameter λ for values ranging from 40 to 105 in steps of 5.

- The residual deviance for reach of the 13 individual splines was plotted against the parameter λ to detect the threshold at which the residual deviance no longer minimized with increasing value of λ. The λ value associated with those thresholds was used to identify the best fitted spline for each of the time series (i.e., each biological year) separately for each ladder (North or South).

Results of the daily count estimation process incorporating all of the above-mentioned components and steps are presented in **Attachment 2**. On each individual page in Attachment 2, the top figure presents the fitted spline, and the bottom figure represents the resultant estimated daily count of Chinook salmon for the specified ladder (North or South) for each of the time series (i.e., each biological year).

References

Freund, J. E. and G. A. Simon. 1991. Statistics: A first course, 5th ed. Prentice Hall, Englewood Cliffs, New York.