# 2020 UPDATE

# LOWER YUBA RIVER VAKI RIVERWATCHER$^{TM}$ CHINOOK SALMON PASSAGE AND RUN DIFFERENTIATION ANALYSES

December 2020

**Prepared for:**

YUBA WATER AGENCY

1220 F Street
Marysville, CA 95901

**Prepared by:**

**Brian Poxon and Paul Bratovich**

HDR
**2365 Iron Point Road, Suite 300**
**Folsom, CA 95630**

# 2020 UPDATE

# LOWER YUBA RIVER VAKI RIVERWATCHER™ CHINOOK SALMON PASSAGE AND RUN DIFFERENTIATION ANALYSES

## Table of Contents

Page

## List of Figures

## List of Tables

Page



## List of Attachments

# 2020 UPDATE

# LOWER YUBA RIVER VAKI RIVERWATCHER™ CHINOOK SALMON PASSAGE AND RUN DIFFERENTIATION ANALYSES

## 1.0 INTRODUCTION

The Lower Yuba River Accord (Yuba Accord) River Management Team (RMT) presented annual escapement estimates for spring-run and fall-run Chinook salmon (*Oncorhynchus tshawytscha*) in the lower Yuba River upstream of Daguerre Point Dam (DPD), as well as the methodology by which the estimates were produced, in the Yuba Accord Monitoring and Evaluation Program Draft Interim Report (RMT 2013). This document describes a revised approach for analyzing upstream passage of Chinook salmon in the lower Yuba River at DPD for the purposes of calculating annual escapement estimates of spring-run and fall-run Chinook salmon in the lower Yuba River upstream of DPD. The revised approach builds upon the conceptual and statistical underpinnings of the original methodology to incorporate relevant biological considerations evident in available datasets that were not previously accounted for by the previous methodology, nor reflected in the escapement estimates reported in RMT (2013).

### 1.1 BACKGROUND

#### 1.1.1 Description of the Data

Located at river mile (RM) 11.5 in the lower Yuba River, DPD (**Figure 1**) provides fish passage *via* two fish ladders located at both the northern and southern most extremes of the dam crest (**Figure 2**). In 2003, the U.S. Army Corps of Engineers (USACE) granted the California Department of Fish and Wildlife (CDFW) a license to install and operate electronic fish counting devices, referred to as a VAKI Riverwatcher™ infrared and photogrammetric system (VAKI Aquaculture Systems Ltd., www.VAKIIceland.is), in the fish ladders at DPD (**Figure 3**).

Fish passage at DPD has been monitored since 2004 using VAKI Riverwatcher™ systems installed in both ladders. The technology allows for image capture in turbid conditions and is capable of detecting objects with a height (i.e., equivalent to body depth in fish) of 40mm or greater. Both electronic images and silhouettes are captured and recorded with a timestamp for each detected fish passage event. Winari software (VAKI Aquaculture Systems Ltd.) is used to further analyze data and document relevant metrics such as fish length, body depth, and speed of passage through the recordable field of view. Images are used to identify recorded fish to species, and they are frequently of sufficient detail to enable identification of additional physical features (e.g., adipose fin presence or absence for salmonids, physical injuries or deformities). Daily passage data are grouped into annual time series according to biological years, which extend from March 1 of any given year to February 28 (or February 29 in leap years) of the following year (RMT 2013).

**Figure 1. Daguerre Point Dam in the lower Yuba River, California.**

**Figure 2. North fish ladder at Daguerre Point Dam.** ***Source: USACE 2012***

**Figure 3. VAKI Riverwatcher$^{TM}$ installation at DPD.**

The VAKI Riverwatcher$^{TM}$ systems located at both the north and south ladder of DPD were able to record and identify the timing and magnitude of passage for Chinook salmon at DPD during most temporal periods of a given year, although system failures have reduced the ability of the equipment to document ladder use during some months. VAKI system non-operation events have occurred due to: (1) low-voltage disconnections (LVD); (2) system maintenance; (3) power system failures; or (4) unknown malfunctions. Due to these system outages, a complete census of the adult Chinook salmon migrating daily upstream of DPD was not possible for each year. Power system failures have occurred due to operationally detrimental environmental conditions (heavy overcast and foggy conditions resulting in lack of photovoltaic charging of the system).

Since June 2003, numerous improvements have been implemented to improve the reliability of the VAKI Riverwatcher$^{TM}$ systems. During 2004, LED strobes were added to provide nighttime photographic capability. During 2006, a tunnel weir configuration on the north ladder was implemented that housed a digital camera in addition to LED strobes to improve photographic quality. The north VAKI Riverwatcher$^{TM}$ system was replaced during 2010, and a suite of other improvements were implemented, including:

- Pre-cast buildings were erected at both north and south ladders, which allowed for better climate control and removal of sensitive equipment from the flood zone.

- New solar panels were installed (3 total per side) at both the north and south ladders.
- A new larger battery bank allowing for 5-day runtime without photovoltaic generation was added to both ladders.
- A new touch screen PC controller with improved data storage and processing abilities was added to both ladders, which allowed for digital video to be captured for each passing fish.
- A new tunnel weir with improved higher resolution video camera was added to the north ladder.
- The previous north tunnel weir was moved to the south ladder.

During 2019, generator back-up systems were installed to augment the photovoltaic panel arrays. System operational status is monitored and recorded at a four-hour temporal resolution.

### 1.1.2 Previous Analytical Methodology

The methodology previously employed for analyzing the annual DPD Chinook salmon passage time series for any given biological year consisted of one data preparation component, and two analytical components. First, the raw passage data were parsed for Chinook salmon records and daily net upstream passage counts of Chinook salmon were calculated for each ladder, by subtracting downstream passage events. Second, daily net upstream Chinook salmon passage counts and daily system operational data for each ladder were analyzed with non-parametric modeling methods to estimate ladder-specific daily upstream passage counts during periods of VAKI Riverwatcher™ system outage. Finally, nonlinear regression techniques were utilized in an iterative process to analyze total daily upstream Chinook salmon passage counts and identify a day within the biological year that signified the temporal demarcation between spring-run Chinook salmon passage and fall-run Chinook salmon passage. The methodology is described in detail in RMT (2013) and relevant portions of that document and RMT (2010, Appendix - https://www.yubawater.org/207/Annual-Reports) are included as attachments to this document. Key concepts of the methods are described below.

#### Correcting Net Upstream Passage Counts for VAKI Outages

For each biological year and each ladder, to correct the daily observed upstream passage counts for periods of VAKI Riverwatcher™ system outages, Generalized Additive Model (GAM) fitting tools were utilized to fit smoothing splines to the observed counts. In the GAM fitting process, the series of daily proportional hours that the VAKI Riverwatcher™ system in each ladder operated were used as weights in the fitting minimization process. The smoothing spline parameter, $\lambda$, was iteratively varied to produce a set of models for each ladder with variable degrees of smoothing. Residual deviance for each fitted model in the set was plotted against its $\lambda$ value to identify the threshold at which increasing the $\lambda$ value no longer minimized residual deviance - the model associated with the threshold $\lambda$ value was selected as the best-fitting spline. The best-fitting spline for each ladder was used to predict daily upstream passage and the predicted values were rounded to the nearest integer. For each ladder, on days of partial VAKI Riverwatcher™ system operation, if the predicted upstream passage was higher than the observed count, the predicted count was selected as the corrected count; if the predicted count was less than the observed count, the observed count was retained as the corrected count. The corrected series of counts for each ladder was summed by day across ladders to produce the annual series of total daily Chinook salmon passage upstream of DPD. Fitting of the smoothing splines was carried out using the 'gam' function in the S-Plus statistical computing software. A detailed description of the correction method is provided in **Attachment A**.

### Differentiation of Spring-Run and Fall-Run Chinook Salmon

Because of considerable inter-annual variability in the pattern of Chinook salmon passage upstream of DPD, and because of the lack of data to differentiate spring-run from fall-run for individual Chinook salmon based on physical appearance, available data were used to generate assumptions regarding the phenotypic passage timing of spring-run versus fall-run Chinook salmon at DPD. A statistical method was developed to assist with the differentiation of spring-run from fall-run Chinook salmon passing DPD. Key assumptions included that in any given biological year:

- a single day (i.e., "cutting day") could be identified by which passage of spring-run Chinook salmon could be differentiated from passage of fall-run Chinook salmon;
- all daily observations occurring prior to the cutting day were spring-run Chinook salmon, and all daily observations occurring on and subsequent to the cutting day were fall-run Chinook salmon;
- based on Yuba River data and adult immigration timing reported for Central Valley spring-run Chinook salmon, the cutting day existed between June 15 and September 15, inclusive; and,
- the temporal distributions of passage at DPD for both spring-run and fall-run Chinook salmon were individually unimodal and potentially asymmetric in shape.

Based on these assumptions, a statistical method was developed whereby, for each cutting day within the range of possible cutting days, the annual time series of daily observations was divided into two separate series. The first series, which represented spring-run Chinook salmon, consisted of all observations occurring from March 1 to the day before the cutting day. The second series represented fall-run Chinook salmon and consisted of all observations occurring on and subsequent to the cutting day. A generalized logistic function was fit to each of the two series and used to predict daily counts for each series. The two series of predicted counts were combined to create a single annual time series of predicted daily counts, and the proportion of annual daily variability explained by the predicted series (i.e., the coefficient of determination, $r^2$) was calculated. This process was repeated for each cutting day within the range of possible cutting days, and the cutting day that resulted in the maximum value of $r^2$ was selected as the temporal demarcation date between spring-run and fall-run Chinook salmon passing DPD. Fitting of the generalized logistic function was accomplished using a Visual Basic macro in Microsoft Excel. A detailed description of the method excerpted from RMT (2013) is provided in **Attachment B**.

### 1.1.3 Need for Revised Approach

The approach to Chinook salmon passage data analysis described above was developed based on rigorous reviews of available analytical techniques, and review of available and relevant data. The results of initial statistical tests conducted on the annual time series of passage data supported the appropriateness of the statistical methodology. However, because of the need to review all available data in an effort to estimate abundance during years with extensive VAKI Riverwatcher™ system outages (see Sections 2.2 and 4.1), a recent review of data from the 2009 through 2011 adult Chinook salmon acoustic telemetry study in the lower Yuba River identified an aspect of phenotypic spring-run adult Chinook salmon immigration not previously considered in run differentiation. Specifically, a distinctly bimodal pattern was identified in the pattern of passage timing at DPD for 65 Chinook salmon that were tagged in the lower Yuba River downstream of DPD during May, June, and July in 2009, 2010, and 2011 (**Figure 4**). The data show that one group of acoustically tagged adult Chinook salmon generally passed DPD from late May through early July, and that a second group of acoustically-tagged Chinook salmon exhibited a prolonged period of holding downstream of DPD before passing upstream of DPD during August through September.

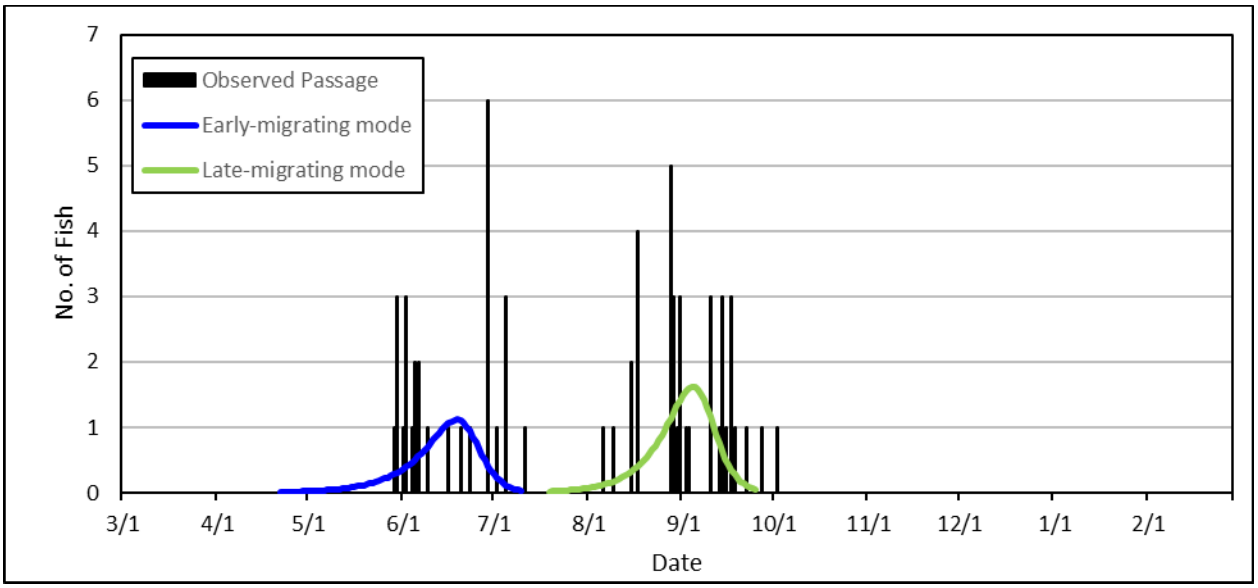



**Figure 4. Daily passage of acoustic-tagged Chinook salmon at DPD for 65 individuals that were tagged in May, June, and July in 2009, 2010, and 2011. The indicated modalities represent the middle 99% of the fitted generalized logistic functions.**

Eight adult Chinook salmon were caught and acoustically tagged during the month of October of 2011 to examine migratory and spawning behavior of phenotypic fall-run Chinook salmon. By contrast to the phenotypic spring-run, the eight adult Chinook salmon (presumably fall-run) tagged in the lower Yuba River during October of 2010 and 2011 only spent from 0 to 7 days from the date of tagging to date of passing DPD. May and June are months in which upstream migrating Chinook salmon would typically be characterized as phenotypic spring-run (Vogel and Marine 1991; Yoshiyama et al. 1996; YCWA et al. 2007; RMT 2013). July is a month during which available information indicates that both phenotypic spring-run and phenotypic fall-run Chinook salmon may be immigrating, although phenotypic fall-run immigration is typically recognized to peak in October or November, and July is typically recognized as the earliest month for phenotypic fall-run Chinook salmon adult immigration (Vogel and Marine 1991; Reclamation 2008; RMT 2013). For the individuals described above that were tagged during May, June, and July, 23 of the 65 individuals (approximately 35 percent) were tagged during July, and all of those individuals were tagged during 2011. VAKI Riverwatcher™ data for 2011 indicate that adult Chinook salmon passage at DPD occurred later than would be typically observed, with 21 observations of adult Chinook salmon upstream passage recorded at DPD prior to July 1, 2011. The year 2011 was considered a wet water year type, and spring-run Chinook salmon appeared to return to their natal rivers in the Central Valley later than usual during 2011. This infers a high likelihood that the individuals tagged in July 2011 were phenotypic spring-run Chinook salmon, although the true genotype of these fish is not known.

The data discussed above indicate that the distribution of phenotypic spring-run Chinook salmon immigration is temporally bimodal, with one group immigrating past DPD from late May through early July with a peak during late June, and another group immigrating past DPD from August through September with a peak during late August. This understanding, combined with the existing characterization that phenotypic fall-run Chinook salmon immigration occurs from July through December, presents a new picture of the annual distribution of Chinook salmon passage at DPD. Specifically, adult Chinook salmon passage at DPD occurring prior to some point in July would be expected to be composed entirely of phenotypic spring-run Chinook salmon, while passage occurring subsequent to some point in October would be expected to be composed entirely of phenotypic fall-run Chinook salmon. Passage occurring during the in-between period would be expected to be composed of a mix of phenotypic spring-run and phenotypic fall-run Chinook salmon.

Recent advances in the understanding of the underlying genetic basis for different adult immigration timing (i.e., different run types) in meta-populations of Pacific salmonid species, including Chinook salmon, support the temporally dynamic composition of passage described above. Prince et al. (2017) found that differential allelic expression at the GREB-1L gene on chromosome 28 was associated with phenotypic migration timing across populations of both steelhead (*Oncorhynchus mykiss*) and Chinook salmon; a homozygotic condition in one direction presented phenotypically as early migrating behavior, while the homozygotic condition in the other direction presented phenotypically as late migrating behavior. Further, Prince et al. (2017) re-analyzed data originally presented in Hess et al. (2016) and showed that the heterozygotic allelic condition at GREB-1L presented phenotypically as migratory behavior with an intermediate timing. The region of Chinook salmon chromosome 28 most strongly associated with migration timing has further been expanded to include the ROCK1 gene, which is adjacent to GREB-1L (Narum et al. 2018; Kock and Narum 2020; Meek et al. 2020; Thompson et al. 2020). Recent genetic analyses of Chinook salmon provide further support for the intermediate adult migration timing of heterozygous individuals (Thompson et al. 2019; Kock and Narum 2020; Thompson et al. 2020).

The previously utilized methodology for differentiating between phenotypic spring-run and phenotypic fall-run Chinook salmon passing DPD did not account for a temporally dynamic and mixed composition of spring-run and fall-run Chinook salmon in the daily passage at DPD. By design, and as described above, the method assumed that a single day existed within any given biological year that could be used to separate daily passage observations into either all phenotypic spring-run Chinook salmon or all phenotypic fall-run Chinook salmon. For this reason, an effort was undertaken to refine the assumptions of the run differentiation methodology based on the newly recognized temporal dynamism described above, and to adjust the method to allow for certain daily passage counts to be composed of both phenotypic spring-run and phenotypic fall-run Chinook salmon. The statistical concepts underlying the original method were largely retained in the revised approach, since the nature of the underlying stochastic processes by which the data are generated have not changed. However, certain aspects of the model selection process have been updated to reflect modern model selection paradigms, and the analyses have been migrated from the S-Plus and Visual Basic/Microsoft Excel environments to the open-source R Statistical Computing environment (R Core Team 2020), a powerful statistical computing environment that is commonly used by data analysts, researchers, and academicians. The revised analyses, including updated assumptions and methodologies, are detailed below.

# 2.0 METHODS

## 2.1 DATA PREPARATION

Passage observations are recorded by the VAKI Riverwatcher™ system in real time, while system operational status is recorded in four-hour increments. For any given biological year, the raw data from the system are summarized by ladder to generate an annual time series of daily net upstream passage counts for Chinook salmon, as well as percentage of time each day that each VAKI Riverwatcher™ system was operational.

## 2.2 CORRECTING NET UPSTREAM PASSAGE COUNTS FOR VAKI OUTAGES

Similar to the previous methodology, fitting smoothing functions under a GAM framework using VAKI Riverwatcher™ system daily operational proportion as a weighting component for model fitting was used to correct counts for periods when the VAKI Riverwatcher™ system was non-operational. However, certain aspects of the method have been updated to make use of modern analytical and inferential tools to make the analysis more statistically robust. The following discussion describes the key changes to the method.

### 2.2.1 Computational Environment and Utilities

The S-PLUS software that was previously used to fit GAMs to the data is no longer available and has largely been replaced by the R programming language and statistical computing environment. The R programming language is based on the S language native to S-PLUS (Morandat et al. 2012), and the R environment is widely used and freely available under a GNU General Public License (www.gnu.org/licenses/licenses.html). For this reason, the decision was made to transition the count correction analysis into R. The S-PLUS function for fitting GAMs that was previously used for the count correction analysis has been implemented in the 'gam' package (Hastie 2020) for R, but other packages implementing different approaches to fitting GAMs have been developed for R. One such package is the 'mgcv' package (Wood 2017, 2020) which, unlike the GAM fitting function in the package 'gam', uses generalized cross-validation in the fitting process, provides facilities for automatic smoothness selection, and accommodates data from a broad range of statistical distributions. The increased utility and flexibility provided by these features lead to the selection of the 'mgcv' package for fitting smoothing functions to the daily DPD passage data in this revised methodology. The parameter controlling the degree of smoothing in package 'mgcv' is denoted $k$, as opposed to $\lambda$ in the S-PLUS function (and in the R package 'gam') and as described in RMT (2013).

### 2.2.2 Assumption Verification and Model Selection

#### Assumption Verification

The original method for the count correction method assumed that the annual time series of daily passage counts were Poisson-distributed, although it is unclear from available descriptions of the method how or if that assumption was tested. The Poisson distribution is appropriate for modeling random count data where the sample mean and variance are equal (Zar 1999; Grafen and Hails 2002). When variance is less than the sample mean, the data are considered to be under-dispersed. Conversely, when variance is greater than the sample mean, the data are considered to be over-dispersed (Grafen and Hails 2002). Both conditions violate the mean-equal-to-variance assumption of the Poisson distribution, but over-dispersion is more commonly encountered in count data (Ver Hoef and Boveng 2007).

Alternative approaches available for over-dispersed count data include quasi-Poisson estimation, and the zero-inflated Poisson and negative binomial distributions, each of which relax the Poisson assumption of equal mean and variance (Wood et al. 2016). The quasi-Poisson approach uses a quasi-likelihood function (i.e., a function that is similar to, but does not equate to the log-likelihood for a given probability distribution) with Poisson-like assumptions (Ver Hoef and Boveng 2007). The zero-inflated Poisson distribution assumes that counts equal to zero in a dataset are generated by both a Poisson-distributed process (which generates both zero and non-zero counts) as well as a separate non-Poisson process (which generates only counts equal to zero; Lambert 1992). The negative binomial distribution can be parameterized as a generalized form of the Poisson distribution that includes an extra parameter which scales the variance based on the mean (Allison and Waterman 2002). Each of these three alternative approaches are potentially appropriate for modeling over-dispersed count data and tend to yield similar results (Ver Hoef and Boveng 2007), although quasi-Poisson regression has been noted to produce results that are difficult to interpret and draw inference from, due to the approach's lack of a defined probability distribution (Ver Hoef and Boveng 2007). Further, zero-inflation has been described as a special case of over-dispersion that may only be appropriate when non-zero counts are clustered and infrequent (Potts and Elliot 2003 *as cited in* Ver Hoef and Boveng 2007). For these reasons, this revised methodology incorporates an initial step to evaluate whether the data are Poisson-distributed, or if the negative binomial distribution is more appropriate.

### Model Selection

With the original method consisting of a process where models of varying degrees of smoothness are iteratively fit to the data, a means of assessing the goodness-of-fit across the set of fitted models and identifying a best model is necessary. The original method relied on model (i.e., residual) deviance as the assessment criterion. While residual deviance does provide a measure of how well a model fits a set of data, its use as a model selection criterion can lead to selection of models that are overfit to the data. The Akaike Information Criterion (AIC) is an alternative estimator of goodness-of-fit that incorporates a penalty term for the number of estimated parameters, thereby rewarding a model's parsimony in addition to goodness-of-fit and balancing the tradeoff between underfitting and overfitting (Burnham and Anderson 2002). This revised count correction methodology utilizes AIC scores as the basis for model selection.

### 2.2.3 Revised Methodology

Daily passage data are grouped into annual time series according to biological years, which extend from March 1 of any given year to February 28 (or February 29 in leap years) of the following year. The dataset for each biological year consists of Day of Biological Year[1] (integer 1 through 365 or 366 in leap years), North Ladder daily net upstream counts, North Ladder daily VAKI operational proportion, South Ladder daily net upstream counts, and South Ladder daily VAKI operational proportion. For each ladder, Day of Biological Year is the only variable used for predicting daily net upstream passage counts, and daily VAKI operational proportion is used as a weight in the model minimization process. The following steps describe the revised method for correcting the net daily upstream passage data for each biological year.

- Step 1: Use the 'gam' function in package 'mgcv' to fit two initial smoothing models for each ladder assuming Poisson and negative binomial distributions, and record dispersion coefficient and AIC score for each initial model. Select the statistical distribution associated with the initial model having the lowest AIC score and dispersion coefficient closest to 1.0 as the statistical distribution most appropriate for the data.
- Step 2: Iterating across a range of *k* values (*k* ranging from 5 to 100 by 5), fit smoothing models to the data assuming the selected distribution. For each model at each value of *k*, record AIC score and model deviance.
- Step 3: Select the best model for each ladder based on AIC values. Confirm appropriateness of selected best model by comparing the best model AIC score and dispersion coefficient against those recorded for the initial models fit in Step 1.
- Step 4: For each ladder, use the selected best model to predict daily counts, which are rounded to the nearest integer.
- Step 5: Compare predicted daily counts to observed daily counts for each ladder. For each day, if predicted count is greater than observed count and VAKI operational proportion is <1, replace the observed count with the predicted count. If daily VAKI operational

[1] Day 1 of a biological year corresponds to March 1.

proportion is equal to 1, or if the predicted daily count is less than the observed daily count, reject the predicted count and select the observed count as the corrected count.

### 2.3 RUN DIFFERENTIATION

Acknowledging that phenotypic spring-run Chinook salmon annual passage at DPD is bimodal, and that the composition of passage is, for some period during the year, a mixture of phenotypic spring-run Chinook salmon and phenotypic fall-run Chinook salmon, a review of existing statistical techniques appropriate for identifying and discerning between subpopulation components within a larger population was conducted. Without information regarding the genetic characteristics (i.e., GREB-1L/ROCK1 genotype) of Chinook salmon passing DPD within each biological year, the primary predictor of run-type available for differentiating between the spring-run and fall-run Chinook salmon populations within each biological year is date of passage. Statistical approaches available for classifying data into groups based on predictors include discriminant analysis, logistic regression, cluster analysis, and mixture modeling. Discriminant analysis and logistic regression are both statistical approaches that can be used for classifying individual data points to different source populations (Afifi et al. 2004). However, both of these approaches require that the initial dataset used to fit the model include information regarding the source population to which each data point belongs (Afifi et al. 2004; Crawley 2007), precluding their application to the DPD Chinook salmon passage data.

Cluster analysis and mixture models comprise a set of techniques which can be used to classify data points to source populations when no information regarding source population is available (Afifi et al. 2004; Crawley 2017), and therefore are potentially appropriate for the DPD Chinook salmon passage data. While both approaches have been formulated to accommodate data from a variety of statistical distributions, neither approach provides the means by which an individual data point could be partitioned among multiple groups, as is needed for the DPD Chinook salmon passage data. Therefore, cluster analysis and mixture models were not selected for differentiating run types from daily DPD passage counts, given that a daily count could be comprised of both spring-run and fall-run Chinook salmon.

Given a lack of existing statistical methods that are appropriate for the DPD Chinook salmon daily passage data, the original methodology for differentiating Chinook salmon run types from DPD passage data was retained as the basis for updating assumptions and analytical procedures to account for considerations derived from the results of the 2009-2011 acoustic study. The updated assumptions and the rationale for each assumption are provided in Section 2.3.2 below, while Section 2.3.3 provides the revised methodology. Key changes to the assumptions include defining the temporal periods during which early-migrating spring-run, late-migrating spring-run, and fall-run Chinook salmon are expected to be passing DPD, defining a period during which late-migrating spring-run and fall-run Chinook salmon are expected to be passing DPD concurrently, and adjusting the range of days in which the cutting day is assumed to exist. One of the key changes to the methodology is that, in the revised methodology, the cutting day has been renamed to "differentiating day" and repurposed. Whereas in the original methodology the cutting day served to identify the temporal demarcation between passage of spring-run Chinook salmon and fall-run Chinook salmon, the differentiating day in the revised methodology serves only to define the number of daily observations included in both the late-migrating spring-run and fall-run Chinook salmon datasets for the model fitting process. However, a byproduct of this function is that, in effect, the differentiating day directs the temporal location of the point at which the late-migrating spring-run and fall-run Chinook salmon passage distributions intersect each other. Other key changes to the methodology resulting from the updated assumptions include truncating the annual passage time series (see Section 2.3.3, below) and using model predictions to allocate portions of some daily passage counts to both late-migrating spring-run and fall-run Chinook salmon.

Similar to the revised daily count correction analysis, the potential to update the model selection criterion from the coefficient of determination to AIC score was evaluated. However, the methodology consists of fitting individual models to separate portions of the annual dataset and then assessing both models, in combination, respective to the total annual time series of observed passage counts. This aspect of the analysis negates the use of AIC scores for model selection for run differentiation (by contrast to use of AIC scores for model selection in the count correction analysis), as no method exists for calculating an overall AIC score from two distinct models generated from separate datasets. Therefore, the original method of selecting best models based on maximizing the coefficient of determination is retained in this revised run differentiation methodology. Further, all model fitting and analysis for the revised methodology has been migrated from Visual Basic and Microsoft Excel to the R Statistical Computing environment. Non-linear models were fit to the data using a nonlinear least squares procedure with a grid search optimization method for parameter estimation (function 'nls2' in R package 'nls2'; Grothendieck 2013).

### 2.3.1 Run Component Terminology

As described in Section 1.1.3, available data from acoustic tagged phenotypic spring-run Chinook demonstrate that adult Chinook salmon migrated upstream of DPD in two temporally distinct groups. Section 1.1.3 also provides a discussion of recent genetic research indicating that differences in adult Chinook salmon migration timing may be explained by differential allelic expression in the region of chromosome 28 occupied by the GREB-1L and ROCK-1 genes, and that a heterozygous condition there has been found to result in adult migration timing intermediate to either homozygous condition. Taken in combination, the bimodal pattern of acoustic tagged phenotypic spring-run Chinook salmon passage may be that the second temporal mode of passage represents individuals of the heterozygous condition. However, the acoustic study occurred several years prior to the publication of reports regarding the GREB-1L/ROCK1 connection and, as such, no specific analyses to determine the allelic identity of the acoustic tagged Chinook salmon were conducted. Without the results of such genetic analyses to support the grouping of the acoustic tagged Chinook salmon by allelic condition (i.e., homozygous early-migrating, heterozygous, homozygous late-migrating), this report utilizes the following terminology to describe the different temporal components of adult Chinook salmon passage at DPD:

- 'Early-migrating phenotypic spring-run Chinook salmon' is used to describe Chinook salmon associated with the first temporal mode of acoustic tagged adult Chinook salmon passage.

- 'Late-migrating phenotypic spring-run Chinook salmon' is used to describe Chinook salmon associated with the second temporal mode of acoustic tagged adult Chinook salmon passage.
- 'Fall-run Chinook salmon' is used to describe Chinook salmon associated with the typical phenotypic fall-run adult Chinook salmon passage timing.

The term 'late-migrating phenotypic spring-run Chinook salmon' was selected, as opposed to 'early-migrating fall-run Chinook salmon', for the following reasons. First, as described in Section 1.1.3, the date of capture for the acoustic tagged Chinook salmon exhibiting the later mode of passage timing aligns with the migration timing recognized for phenotypic spring-run Chinook salmon in the lower Yuba River (RMT 2013). Second, prior to passing upstream of DPD, the acoustic tagged Chinook salmon exhibiting the later mode of passage timing also exhibited a period of prolonged holding downstream of DPD, consistent with the extended holding behavior typically recognized for phenotypic spring-run Chinook salmon and not for phenotypic fall-run Chinook salmon.

### 2.3.2 Assumptions

The underlying assumptions of this revised run differentiation analysis are as follows:

Assumption 1: All passage prior to July 15 (day $D = 137$) is assumed to be entirely composed of early-migrating spring-run Chinook salmon.

*Rationale for Assumption 1*: As described in Section 1.1.3 and illustrated in Figure 4, acoustic tagged phenotypic spring-run Chinook salmon passed upstream of DPD in two temporal groups during 2009, 2010, and 2011. The first group passed DPD from May 29 through July 11. The second group of late-migrating phenotypic spring-run Chinook salmon passed DPD from August 6 through October 2. No passage at DPD of acoustic tagged phenotypic spring-run Chinook salmon occurred from July 12 through August 5. July 15 was selected as the Day of the Biological Year before which passage at DPD is assumed to be composed entirely of phenotypic early-migrating spring-run Chinook salmon because that date encapsulates the entire temporal range of upstream passage of acoustic tagged early-migrating spring-run Chinook salmon, while acknowledging that range may extend slightly later in some years. This assumption is necessary to define the beginning of the period during which the composition of DPD passage is a mixture of late-migrating spring-run and fall-run Chinook salmon.

Assumption 2: Passage occurring on and subsequent to October 1 (day $D = 215$) is assumed to be entirely composed of fall-run Chinook salmon.

*Rationale for Assumption 2*: As discussed in the rationale for Assumption 1, the latest date of passage of acoustic tagged phenotypic spring-run Chinook salmon at DPD during 2009, 2010, and 2011 was October 2. Selection of October 1 as the date on and after which upstream passage of Chinook salmon at DPD is assumed to be composed entirely of phenotypic fall-run Chinook salmon. October 1 was selected because it predominantly excludes observations of phenotypic acoustic tagged spring-run passage upstream of DPD, avoids undue veracity of date selection, avoids over inflation of spring-run Chinook salmon abundance estimation relative to fall-run Chinook salmon, and conveniently is consistent with water year commencement.

Assumption 3: Passage occurring from July 15 through September 30 (78 days, days *D* = 137 through 214), inclusive, is assumed to be composed of a mixture of late-migrating spring-run Chinook salmon and fall-run Chinook salmon, and the daily composition of the mixture is assumed to be annually variable.

*Rationale for Assumption 3*: Given Assumptions 1 and 2, Assumption 3 is justified by the expectation that the distribution of passage of the late-migrating group of phenotypic spring-run Chinook salmon is entirely contained within the period from July 15 through September 30. The distribution of phenotypic fall-run Chinook salmon passage is typically expected to begin as early as July, thereby overlapping with the distribution of late-migrating spring-run Chinook salmon passage. This results in a mixed composition of daily passage during the period where the two distributions overlap. The window of July 15 through September 30 acknowledges the uncertainty regarding the run-type composition of DPD passage during the period when run-specific passage overlaps.

Assumption 4: The period from September 1 through September 30 (30 days, days *D* = 185 through 214) is assumed to be the most appropriate period to search for the point where the late-migrating spring-run Chinook salmon and fall-run Chinook salmon passage distributions intersect each other – this range of dates is termed the "search window".

*Rationale for Assumption 4*: The search window contains the set of possible differentiating days, and the differentiating day directs the temporal location of the point at which the distributions of late-migrating spring-run and fall-run Chinook salmon passage intersect each other. Therefore, the search window should be restricted to the right half of the distribution of late-migrating spring-run passage. Descriptive statistics for the date of DPD passage of the late-migrating group of acoustic tagged phenotypic spring-run Chinook salmon in 2009, 2010, and 2011 were evaluated to inform the location and duration of the search window. The median date of passage at DPD for those individuals was August 31, and the last date of passage was October 2. September 1 was selected as the start of the search window because it is the date nearest to and subsequent to the median date. September 30 was selected because it coincides with the end of the period identified in Assumption 3.

Assumption 5: A generalized logistic function (Richards 1959) is assumed to be the most appropriate function to model the distributions of both late-migrating spring-run and fall-run Chinook salmon daily passage counts.

*Rationale for Assumption 5*: The distributions of passage for late-migrating spring-run, as well as for fall-run Chinook salmon, are each expected to exhibit a single mode and potentially be asymmetrical in shape. A generalized logistic function, parameterized as described in the methodology below, is capable of generating unimodal smooth curves while providing flexibility with respect to the curve's skewness.

### 2.3.3 Revised Methodology

The initial dataset for each biological year consists of Day of Biological Year, and total daily net upstream corrected passage count (daily sum of corrected net upstream counts for both ladders). With the exception of truncating the adult Chinook salmon passage annual time series to only include July 15 through the end of the biological year, and changing the search window to extend from September 1 through September 30, the following methodology represented by steps (1 through 9) is the same as that presented in RMT (2013) and is presented here for completeness.

- Step 1: Separate the annual dataset into two parts: counts occurring from March 1 to July 15, and counts occurring on July 15 through the end of the biological year.
  - The dataset containing all daily counts from March 1 through July 15 will describe the early-migrating component of spring-run Chinook salmon and will sum to a total $N_{SEM}$.
  - The dataset containing counts from July 15 through the end of the biological year is termed the "working dataset".

- Step 2. For each differentiating day $D$ in the search window, beginning with $D$ = 185, separate the set of daily counts $n_d$ in the working dataset into two subsets.
  - One subset that includes all daily counts occurring prior to day $D$ will describe the potential late-migrating spring-run Chinook salmon counts and will sum to a total $N_{SLM}$.
  - The other subset that includes all observations that occur from day $D$ through day 365 or 366 (leap year) will describe potential fall-run Chinook salmon counts and will sum to a total $N_F$.

- Step 3. Calculate the daily relative cumulative distributions for the two data subsets of the previous step.
  - The cumulative distribution of potential late-migrating spring-run Chinook salmon daily observations to day $d = X$ will be calculated as $Y_{SLM_X} = \sum_{d=1}^{X<D} n_d / N_{SLM}$.
  - The cumulative distribution of potential fall-run Chinook salmon daily observations through day $d = X$ will be calculated as $Y_{F_X} = \sum_{d \geq D}^{X \leq 365} n_d / N_F$.

- Step 4. Fit a generalized logistic function (Richards 1959) to each of the two sets of daily relative cumulative distributions calculated in the previous step using a nonlinear least squares procedure.
  - The expected cumulative distribution of potential late-migrating spring-run Chinook salmon daily counts to day $d = X$ as represented by the corresponding fitted asymmetric logistic function has the formula:

$$\hat{Y}_{SLM_X} = 1 \Big/ [1 + exp(\alpha_{SLM} - \beta_{SLM} * X)]^{\left(1/\delta_{SLM}\right)}.$$

  - Similarly, the expected cumulative distribution of potential fall-run Chinook salmon daily counts through day $d = X$ as represented by the corresponding fitted asymmetric logistic function has the formula:

$$\hat{Y}_{F_X} = 1 \Big/ [1 + exp(\alpha_F - \beta_F * X)]^{\left(1/\delta_F\right)}.$$

- The $\alpha_S$, $\alpha_F$, $\beta_S$, $\beta_F$, $\delta_S$ and $\delta_F$ are the fitted parameter values that describe the shapes of the resulting distribution functions. In particular, the parameter $\delta$, whose value is constrained to be greater than or equal to 0.1 and less than or equal to 10, determines the asymmetry of the resulting functions.

- Step 5. Using the fitted asymmetric logistic functions from the previous step, calculate $\hat{n}_d$ with the formula:

$$\hat{n}_d = N_{SLM} * \left(\hat{Y}_{SLM_{d+1}} - \hat{Y}_{SLM_d}\right) + N_F * \left(\hat{Y}_{F_{d+1}} - \hat{Y}_{F_d}\right).$$

where $\hat{n}_d$ is the expected number of Chinook salmon on day $d$ and $N_{SLM}$ and $N_F$ are as defined in Step 2.

- Step 6. Calculate and record the proportion of the annual daily variability explained by the expected daily observations ($\hat{n}_d$) for both asymmetric functions combined using the following formula:

$$\varphi_D^2 = 1 - \frac{\sum_{d=1}^{365}(n_d - \hat{n}_d)^2}{\sum_{d=1}^{365}(n_d - \bar{n})^2}$$

where $\bar{n}$ is the annual average of daily Chinook salmon observations $n_d$.
  - Note that the calculation method for $\varphi_D^2$ is essentially equivalent to that of the coefficient of determination, commonly denoted $r^2$, except that this expression incorporates the combined variability of the two separate fitted logistic functions.

- Step 7. Select the next differentiating day $D$ in the search window and repeat Steps 2 through 6.

- Step 8. Repeat Step 7 for each of the remaining search window dates $D$.

- Step 9. Once $\varphi_D^2$ has been obtained for each differentiating day $D$ from the search window range 185 through 214 for the biological year being analyzed, select the $D$ corresponding to the maximum $\varphi_D^2$ in the set as $D_{max}$.

- Step 10. Using the fitted logistic function associated with $D_{max}$:
  - Calculate the daily predicted cumulative distribution for late-migrating spring-run Chinook salmon ($\hat{Y}_{SLM_d}$) from day $d$ = 137 through $D_{max} - 1$, and then calculate the predicted daily relative proportion $\hat{p}_{SLM_d}$ of late-migrating spring-run Chinook salmon counts as:

$$\hat{p}_{SLM_d} = \hat{Y}_{SLM_{d+1}} - \hat{Y}_{SLM_d}.$$

  - Calculate the daily predicted cumulative distribution for fall-run Chinook salmon $(\widehat{Y_{F_d}})$ from day $d$ = $D_{max}$ through 365 or 366 (leap year), and then calculate the predicted daily relative proportion $\hat{p}_{F_d}$ of fall-run Chinook salmon counts as:

$$\hat{p}_{F_d} = \hat{Y}_{F_{d+1}} - \hat{Y}_{F_d}.$$

- Step 11. For each day $d$ = 137 through 365 or 366 (leap year), calculate the daily sum of predicted daily relative proportions as $\hat{p}_{Total_d} = \hat{p}_{SLM_d} + \hat{p}_{F_d}$ and then calculate the daily ratio ($\rho_{SLM:Total_d}$) of predicted relative daily proportion of late-migrating spring-run Chinook salmon as:

$$\rho_{SLM:Total_d} = \hat{p}_{SLM_d} \Big/ \hat{p}_{Total_d}.$$

- Step 12: For each day $d$ = 137 through 365 or 366 (leap year), and using $\rho_{SLM:Total_d}$ and the daily observed counts of Chinook salmon $n_d$, calculate the daily estimated number of late-migrating spring-run Chinook salmon as $\hat{n}_{SLM_d} = \rho_{SLM:Total_d} * n_d$, and the daily estimated number of fall-run Chinook salmon as $\hat{n}_{F_d} = n_d - \hat{n}_{SLM_d}$.
- Step 13: The annual escapement estimates for early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon are the expected total sums $\widehat{N}$ for each run component, and are calculated as:
  - Early-migrating spring-run Chinook salmon annual escapement estimate,
    $$\widehat{N}_{SEM} = \sum_{d=1}^{136} n_d.$$
  - Late-migrating spring-run Chinook salmon annual escapement estimate,
    $$\widehat{N}_{SLM} = \sum_{d=137}^{214} \hat{n}_{SLM_d},$$
    where $\hat{n}_{SLM_d}$ is rounded to the nearest integer.
  - Fall-run Chinook salmon annual escapement estimate,
    $$\widehat{N}_F = \sum_{d=137}^{214} \hat{n}_{F_d} + \sum_{d=215}^{365} n_d,$$
    where $\hat{n}_{F_d}$ is rounded to the nearest integer.

# 3.0 RESULTS

## 3.1 CORRECTING PASSAGE COUNTS FOR VAKI RIVERWATCHER™ SYSTEM OUTAGES

Biological years 2016 and 2017 were not used for count correction using GAM estimation due to long periods of VAKI Riverwatcher™ system outages during key periods that resulted in incomplete datasets for both years. Thus, for 2016 and 2017 the counts were reported as uncorrected values and were not used for run differentiation. For the remaining years, the initial model assuming the negative binomial distribution resulted in the lowest initial model AIC score and the dispersion coefficient closest to 1, compared to the initial model assuming the Poisson distribution. Therefore, the negative binomial distribution was assumed for all subsequent GAM count correction model fitting. Relative to AIC scores, utilizing model deviance as a model selection criterion selected models with higher values of $k$ for both ladders nearly every year, which resulted in models being selected that were overfit to the data. Thus, AIC scores were used as the criterion for GAM count correction model selection. For each biological year, corrected daily Chinook salmon counts and predicted values from the top model were combined for both ladders. The resulting daily total corrected counts and predicted values are provided in **Figures 5 through 8**. Also indicated in the figures are the daily percent of the time that the VAKI Riverwatcher™ systems were operational. Daily total corrected counts for all biological years except 2016 and 2017 were used as the basis for the run differentiation analysis.

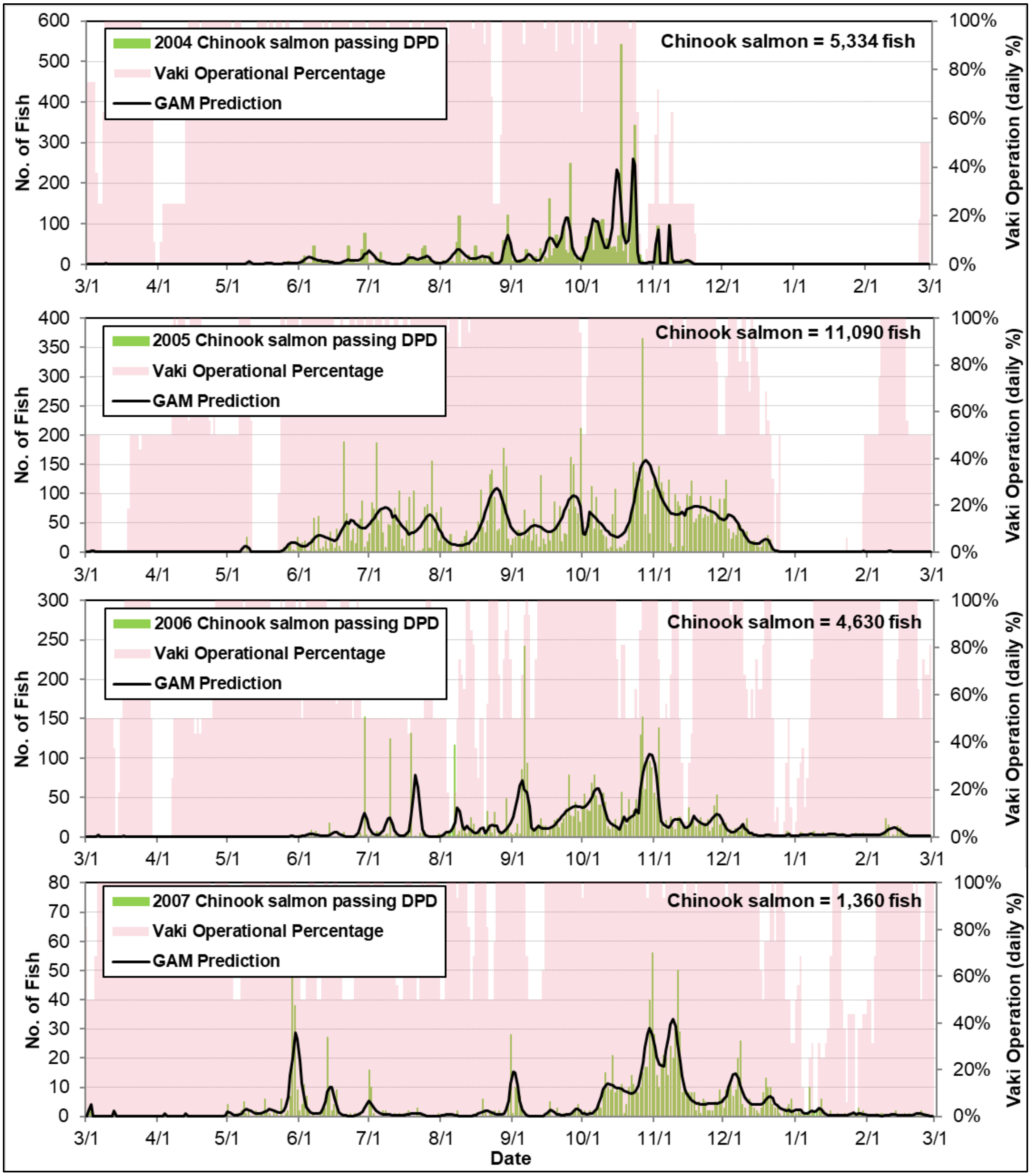



**Figure 5. Daily number of Chinook salmon passing upstream of Daguerre Point Dam during the 2004, 2005, 2006, and 2007 biological years (March 1 through February 28). Green bars indicate VAKI Riverwatcher™ daily counts, black lines indicate the predicted count resulting from the best fitting GAM, and pink shading indicates the daily percent of VAKI Riverwatcher™ system operation.**

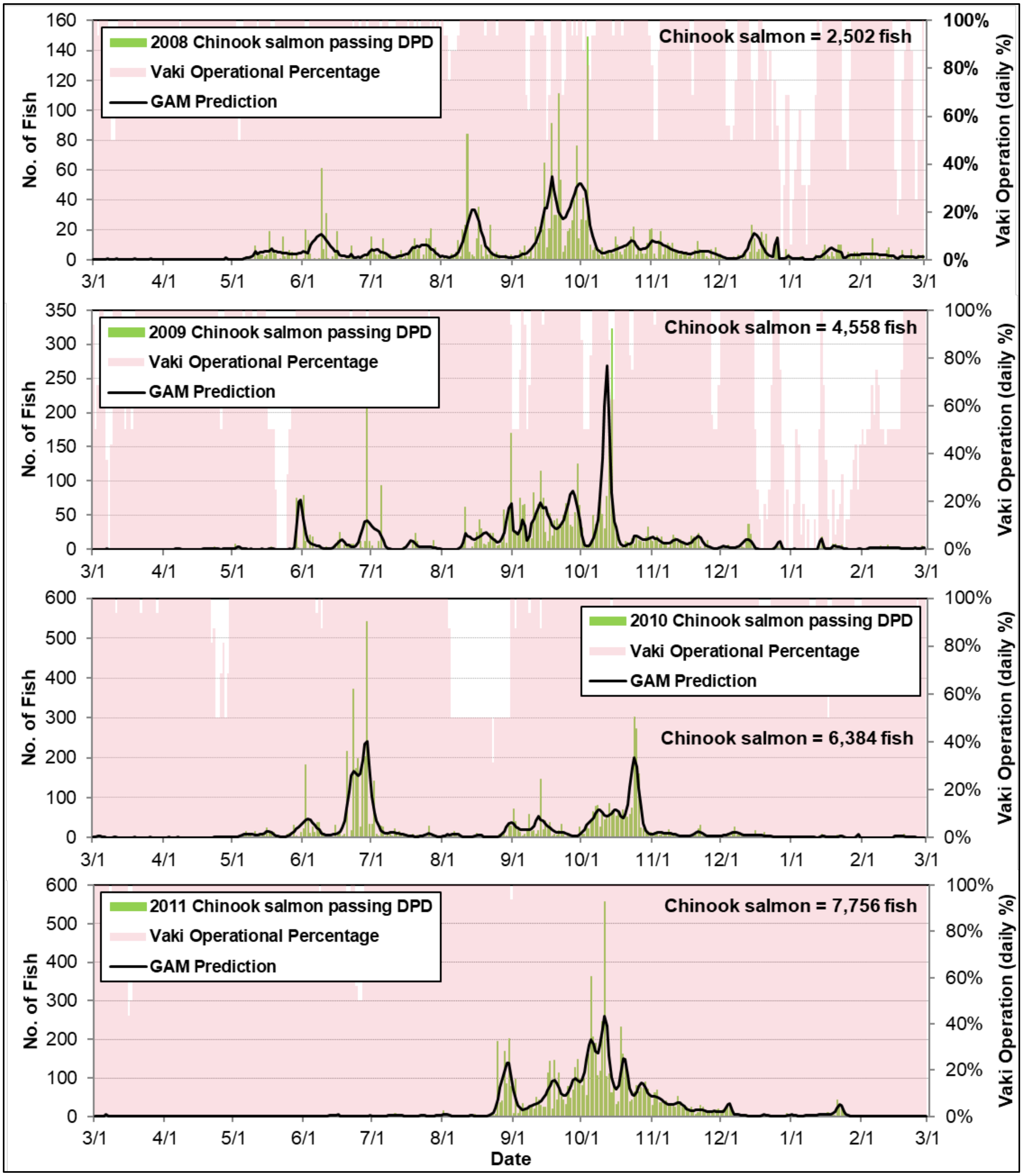



**Figure 6. Daily number of Chinook salmon passing upstream of Daguerre Point Dam during the 2008, 2009, 2010, and 2011 biological years (March 1 through February 28). Green bars indicate VAKI Riverwatcher™ daily counts, black lines indicate the predicted count resulting from the best fitting GAM, and pink shading indicates the daily percent of VAKI Riverwatcher™ system operation.**

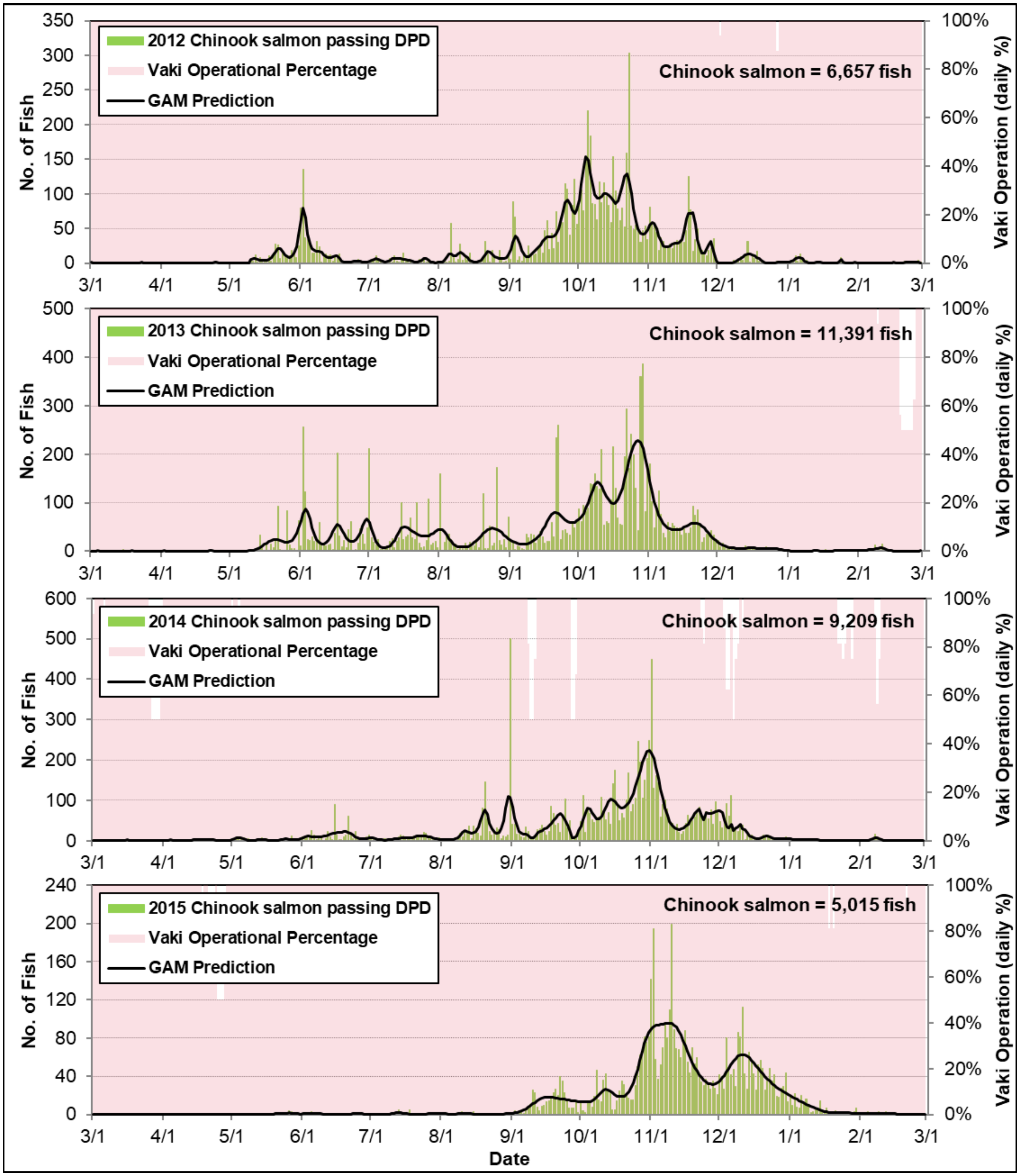



**Figure 7. Daily number of Chinook salmon passing upstream of Daguerre Point Dam during the 2012, 2013, 2014, and 2015 biological years (March 1 through February 28). Green bars indicate VAKI Riverwatcher™ daily counts, black lines indicate the predicted count resulting from the best fitting GAM, and pink shading indicates the daily percent of VAKI Riverwatcher™ system operation.**

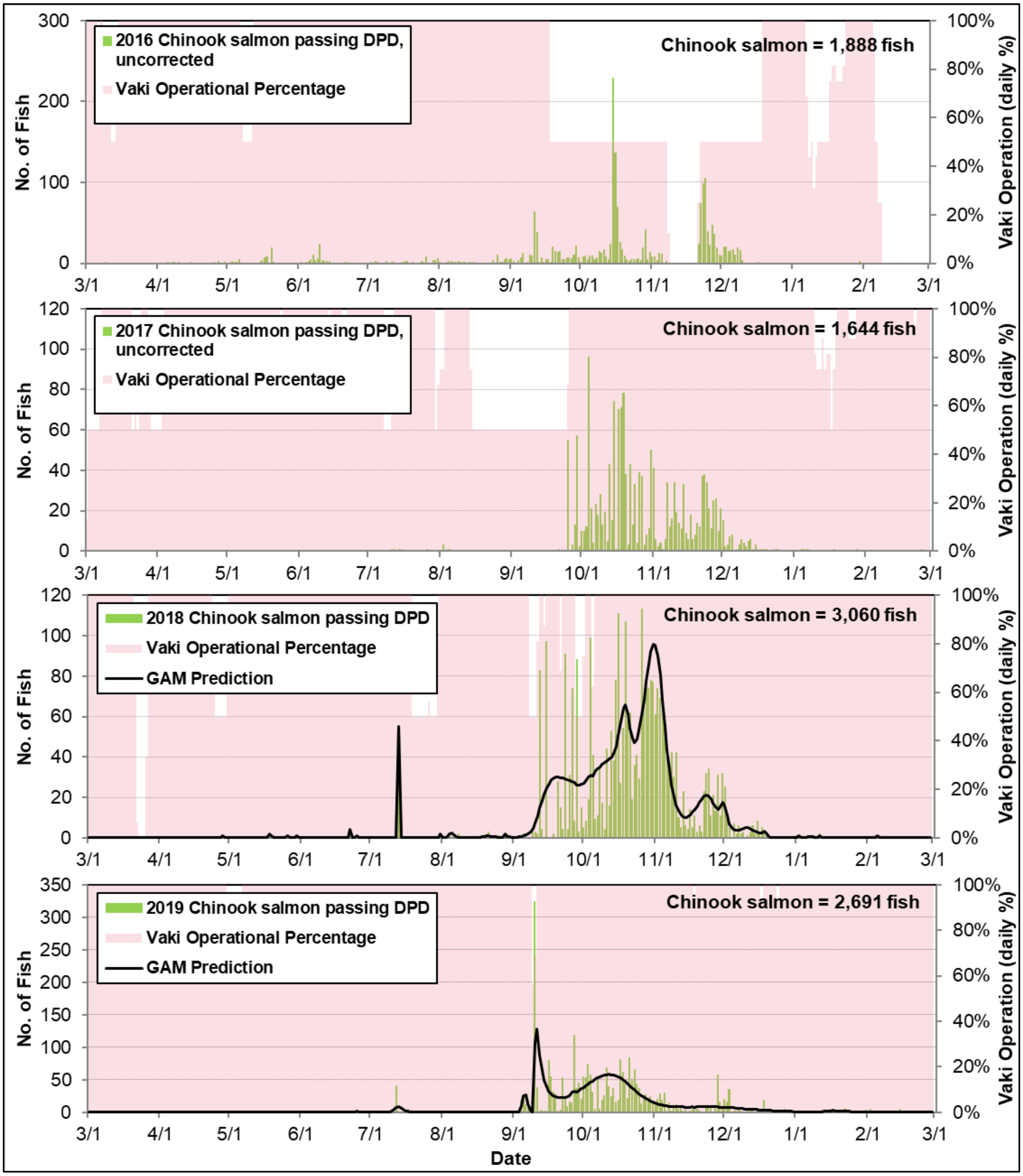



**Figure 8. Daily number of Chinook salmon passing upstream of Daguerre Point Dam during the 2016, 2017, 2018, and 2019 biological years (March 1 through February 28). Green bars indicate VAKI Riverwatcher™ daily counts, black lines indicate the predicted count resulting from the best fitting GAM, and pink shading indicates the daily percent of VAKI Riverwatcher™ system operation. Due to extended periods of VAKI Riverwatcher™ outages that occurred during the 2016 and 2017 biological years, counts for those years could not be corrected through GAM estimation.**

### 3.2 RUN DIFFERENTIATION

For each biological year, the daily proportions of late-migrating spring-run Chinook salmon and fall-run Chinook salmon predicted by the best fitting logistic functions were used to calculate the daily proportion of late-migrating spring-run Chinook salmon relative to total predicted Chinook salmon (see **Figure 9** for an example from the 2004 biological year). The daily proportion was then used to calculate the portion of the corrected daily passage occurring from July 15 through September 30 that is attributed to the late-migrating spring-run component. The daily proportion was also used to calculate the proportion of the observed daily counts of adipose-clipped Chinook salmon occurring from July 15 through September 30 that is attributed to the late-migrating spring-run component passing upstream of DPD within each biological year. The results of run differentiation analyses for the annual time series of daily counts of Chinook salmon passing upstream at DPD for each biological year (March through February, annually) from 2004 through 2019 are summarized in **Table 1** and are provided in **Figures 10 through 15**. Biological years 2016 and 2017 were not included in annual run differentiation due to the aforementioned long periods of VAKI Riverwatcher™ system outages during key migration periods that resulted in incomplete datasets for both years.

### 3.3 ANNUAL ABUNDANCE ESTIMATION AND HATCHERY CONTRIBUTION

For the period (2004-2019) during which VAKI Riverwatcher™ data are available, the annual number of spring-run Chinook salmon estimated to have passed upstream of DPD ranged from 235 in 2015 to 5,274 in 2005. The annual abundances of phenotypic spring-run Chinook salmon in the lower Yuba River appear to be strongly influenced by hatchery fish. The annual percentage of adipose fin-clipped spring-run Chinook salmon relative to the total number estimated to have passed upstream of DPD ranged from 3.5% in 2008 to 80.1% in 2018.

The annual percentage contribution of hatchery fish to the total number of Chinook salmon annually spawning upstream of DPD is reported by CDFW, based upon recovery of coded-wire tags during carcass surveys. In these reports, no distinction is made between spring-run and fall-run Chinook salmon populations in the lower Yuba River, and the percentage hatchery contribution is reported for "Chinook salmon". Reports for return years 2010–2015 (Kormos et al. 2012; Palmer-Zwahlen and Kormos 2013, 2015, 2018; Palmer-Zwahlen et al. 2019a; Palmer-Zwahlen et al. 2019b) were available at the time of preparation of this report and are summarized below.

- During 2010, adult Chinook salmon spawning in the Yuba River were 71% hatchery-origin (note: distinctions were not made between upstream and downstream of DPD).
- During 2011, adult Chinook salmon spawning in the Yuba River upstream of DPD were 65% hatchery-origin.
- During 2012, adult Chinook salmon spawning in the Yuba River upstream of DPD were 45% hatchery-origin.
- During 2013, adult Chinook salmon spawning in the Yuba River upstream of DPD were 34% hatchery-origin.

- During 2014, adult Chinook salmon spawning in the Yuba River upstream of DPD were 49% hatchery-origin.
- During 2015, adult Chinook salmon spawning in the Yuba River upstream of DPD were 60% hatchery-origin.

Comparison of the estimated percentage of all Chinook salmon exhibiting adipose fin clips in the VAKI Riverwatcher™ systems with the percentage of hatchery contribution to the annual spawning population upstream of DPD reported by CDFW indicates the CDFW estimates are 2-3 times higher. That is not unexpected because the direct observation of adipose fin clips in the VAKI Riverwatcher™ systems data do not include expansion based on constant fractional marking or mark-recapture ratios used in population estimation.

For the years (2004-2011) during which annual abundance estimates were available from the original methodology (as reported in RMT 2013) as well as the revised methodology presented in this report, the revised methodology resulted in somewhat higher estimates of annual spring-run Chinook salmon passing upstream of DPD (**Table 2**). The revised methodology of annual spring-run Chinook salmon abundance upstream of DPD yielded estimates ranging from 0.9 to 3.6 times those of the previous methodology, with an average over the eight overlapping years (2004-2011) of 1.9 times higher.

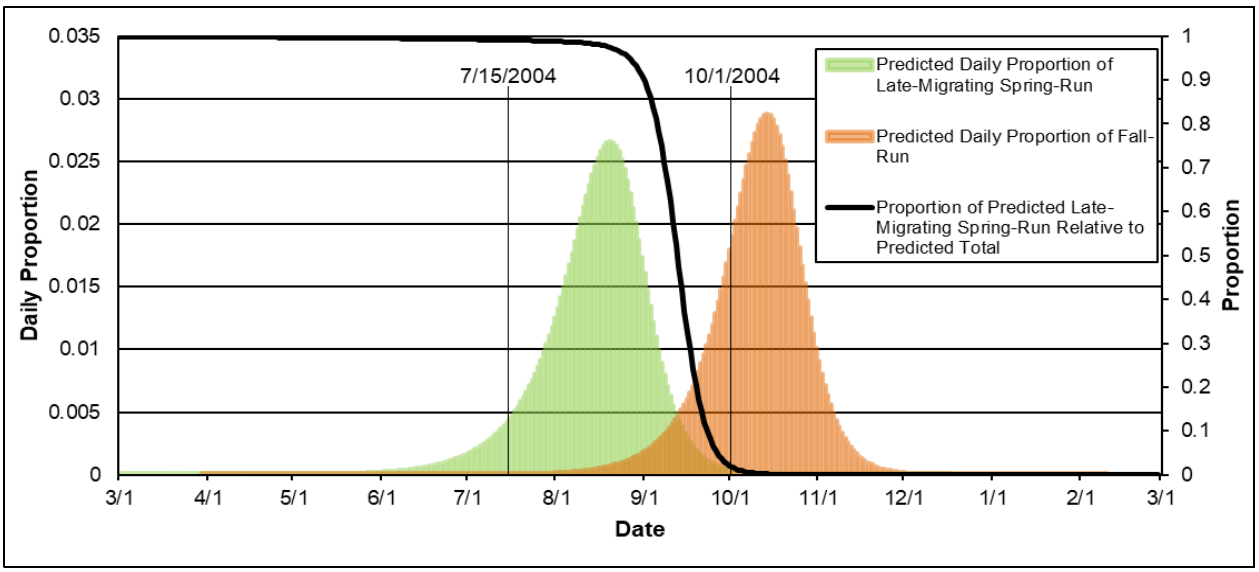



**Figure 9. Predicted daily proportion of late migrating spring-run Chinook salmon, predicted daily proportion of fall-run Chinook salmon, and daily ratio of predicted late migrating spring-run Chinook salmon to total predicted Chinook salmon (i.e., daily sum of predicted late migrating spring-run, as well as fall-run Chinook salmon). The daily ratio is used in the calculation of the portion of the corrected daily net upstream passage at DPD occurring from July 15 through September 30 that is attributed to the late-migrating spring-run Chinook salmon run component.**

**Table 1. Results of phenotypic run differentiation analyses for the 2004 through 2019 biological years of lower Yuba River Chinook salmon based on VAKI Riverwatcher™ net upstream passage observations at DPD, corrected for periods of VAKI system outages. The 2016 and 2017 annual time series were deemed inappropriate for count correction and run separation analyses due to long periods of VAKI Riverwatcher™ system outages that resulted in incomplete datasets for both years.**

| Year | Early-Migrating Spring-Run Chinook Salmon | | Late-Migrating Spring-Run Chinook Salmon | | Total Spring-Run Chinook Salmon | | Fall-Run Chinook Salmon | | Total Chinook Salmon | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Estimate | % Ad-Clip | Estimate | % Ad-Clip | Estimate | % Ad-Clip | Estimate | % Ad-Clip | Estimate | % Ad-Clip |
| 2004 | 509 | 11.8% | 1,109 | 6.2% | 1,618 | 8.0% | 3,716 | 2.6% | 5,334 | 4.2% |
| 2005 | 1,978 | 23.1% | 3,296 | 10.1% | 5,274 | 15.0% | 5,816 | 2.1% | 11,090 | 8.2% |
| 2006 | 374 | 9.6% | 1,082 | 5.3% | 1,456 | 6.4% | 3,174 | 3.1% | 4,630 | 4.1% |
| 2007 | 285 | 13.0% | 68 | 0.0% | 353 | 10.5% | 1,007 | 2.7% | 1,360 | 4.7% |
| 2008 | 372 | 2.7% | 948 | 3.8% | 1,320 | 3.5% | 1,182 | 18.4% | 2,502 | 10.6% |
| 2009 | 693 | 30.2% | 1,923 | 16.4% | 2,616 | 20.1% | 1,942 | 11.9% | 4,558 | 16.6% |
| 2010 | 2,988 | 59.7% | 749 | 44.7% | 3,737 | 56.7% | 2,647 | 15.9% | 6,384 | 39.8% |
| 2011 | 59 | 47.5% | 2,301 | 24.8% | 2,360 | 25.4% | 5,396 | 21.0% | 7,756 | 22.4% |
| 2012 | 807 | 59.4% | 1,482 | 16.6% | 2,289 | 31.7% | 4,368 | 24.7% | 6,657 | 27.1% |
| 2013 | 2,053 | 4.7% | 1,993 | 7.3% | 4,046 | 6.0% | 7,345 | 13.5% | 11,391 | 10.9% |
| 2014 | 618 | 12.1% | 1,403 | 8.8% | 2,021 | 9.8% | 7,188 | 16.6% | 9,209 | 15.1% |
| 2015 | 50 | 8.0% | 185 | 15.1% | 235 | 13.6% | 4,780 | 20.3% | 5,015 | 20.0% |
| 2016 | NA | NA | NA | NA | NA | NA | NA | NA | 1,888 | 45.9% |
| 2017 | NA | NA | NA | NA | NA | NA | NA | NA | 1,644 | 55.4% |
| 2018 | 66 | 80.3% | 215 | 80.0% | 281 | 80.1% | 2,779 | 34.2% | 3,060 | 38.4% |
| 2019 | 48 | 18.8% | 753 | 19.0% | 801 | 19.0% | 1,890 | 18.7% | 2,691 | 18.8% |

# 4.0 DISCUSSION

## 4.1 DATA QUALITY AND CONTINUITY

The large data gaps in the datasets from the 2016 and 2017 biological years that render them inappropriate for these analyses highlight the importance of minimizing VAKI Riverwatcher™ system errors and outages to the greatest extent feasible. The RMT may want to consider system battery bank replacement on a more regular basis to ensure that power source failures are minimized. The RMT may also want to consider purchasing a backup VAKI Riverwatcher™ system so that equipment failures can be quickly rectified.

## 4.2 CORRECTING UPSTREAM PASSAGE COUNTS FOR VAKI RIVERWATCHER™ OUTAGES

Dispersion coefficients for initial and best-fitting models indicate that the negative binomial distribution is more appropriate for the count correction analysis of DPD passage data than the Poisson distribution utilized in the previous methodology. While this result was universal across each analyzed biological year time series, universally assuming the negative binomial distribution for the count correction analysis should not be adopted, and utilizing the assumption checking step presented in this methodology (see Section 2.2.3, Step 1) should be continued for future updates.

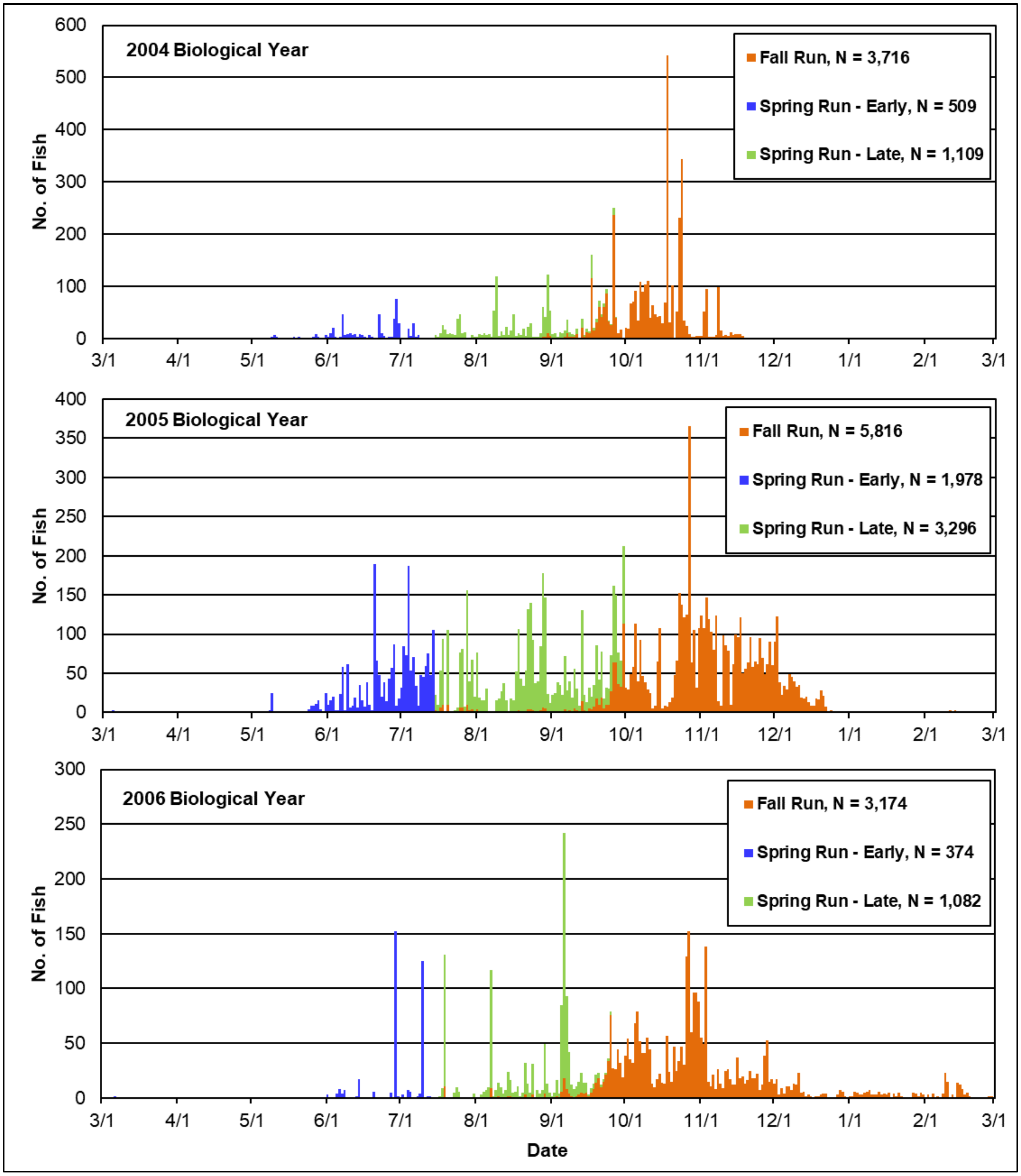



**Figure 10. Daily estimated counts of early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon passing Daguerre Point Dam during the 2004, 2005, and 2006 biological years (March 1 through February 28). Daily counts for each run component are presented as stacked bars, with annual run component escapement estimates provided in figure legends.**

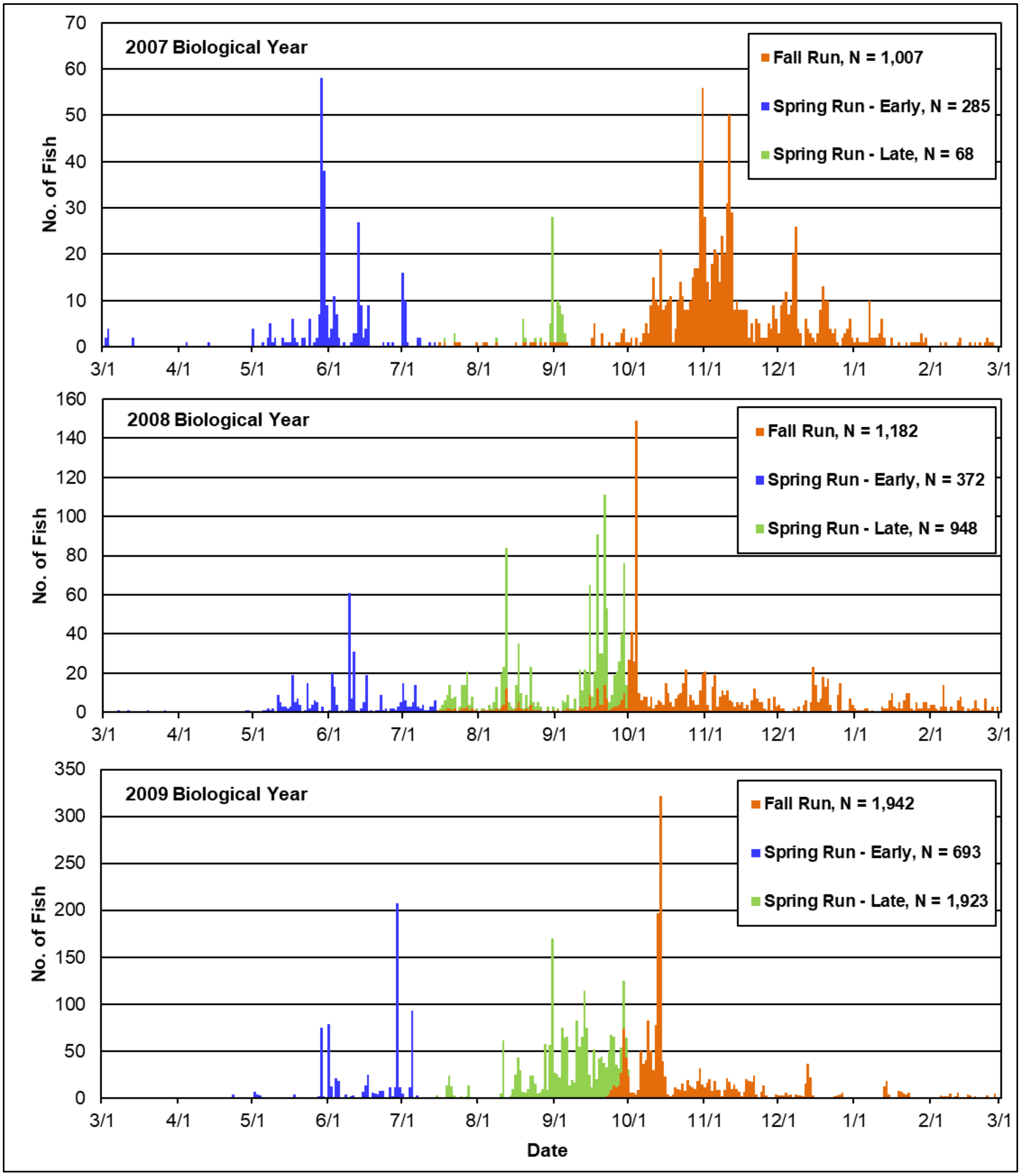



**Figure 11. Daily estimated counts of early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon passing Daguerre Point Dam during the 2007, 2008, and 2009 biological years (March 1 through February 28). Daily counts for each run component are presented as stacked bars, with annual run component escapement estimates provided in figure legends.**

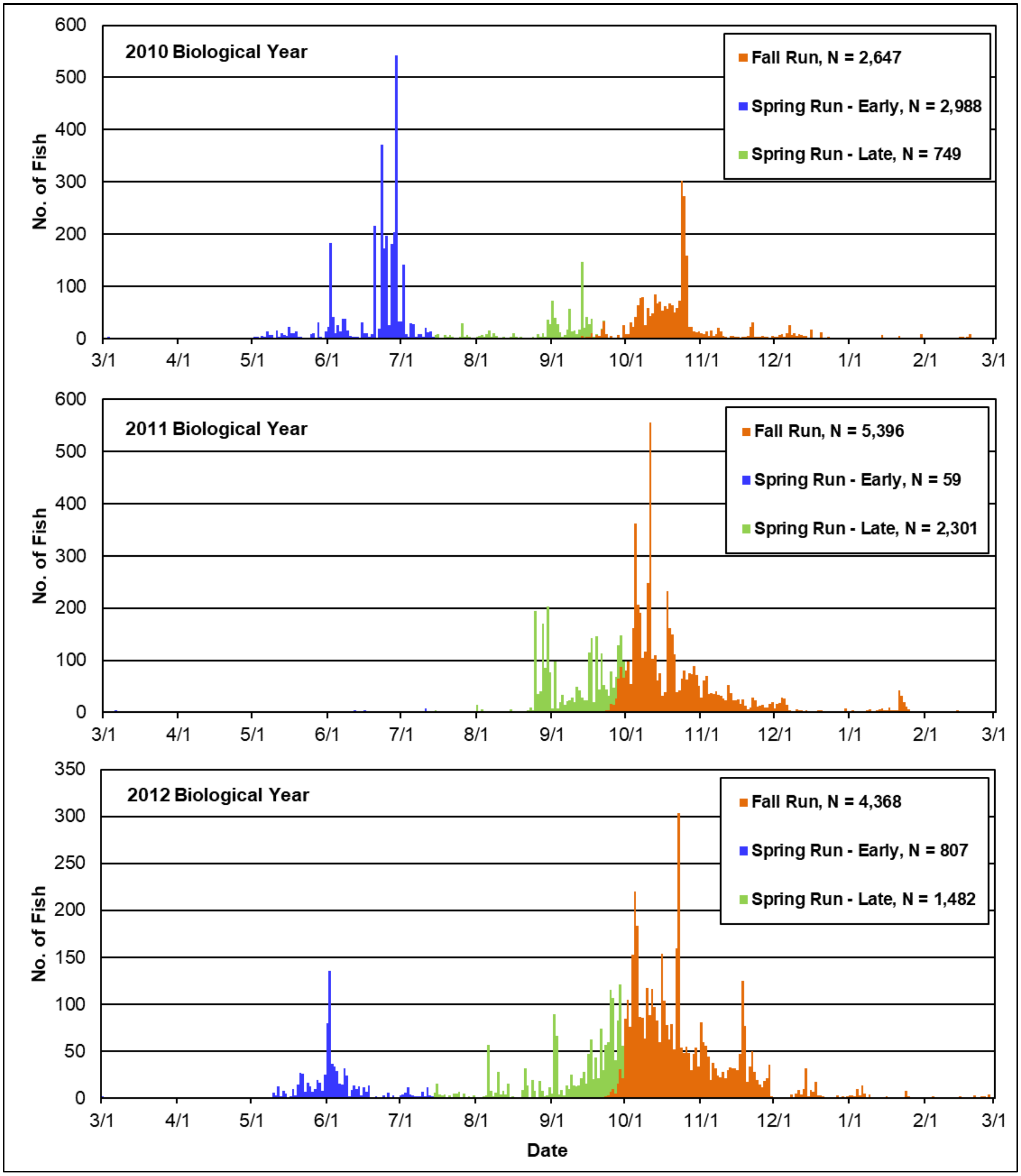



**Figure 12. Daily estimated counts of early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon passing Daguerre Point Dam during the 2010, 2011, and 2012 biological years (March 1 through February 28). Daily counts for each run component are presented as stacked bars, with annual run component escapement estimates provided in figure legends.**

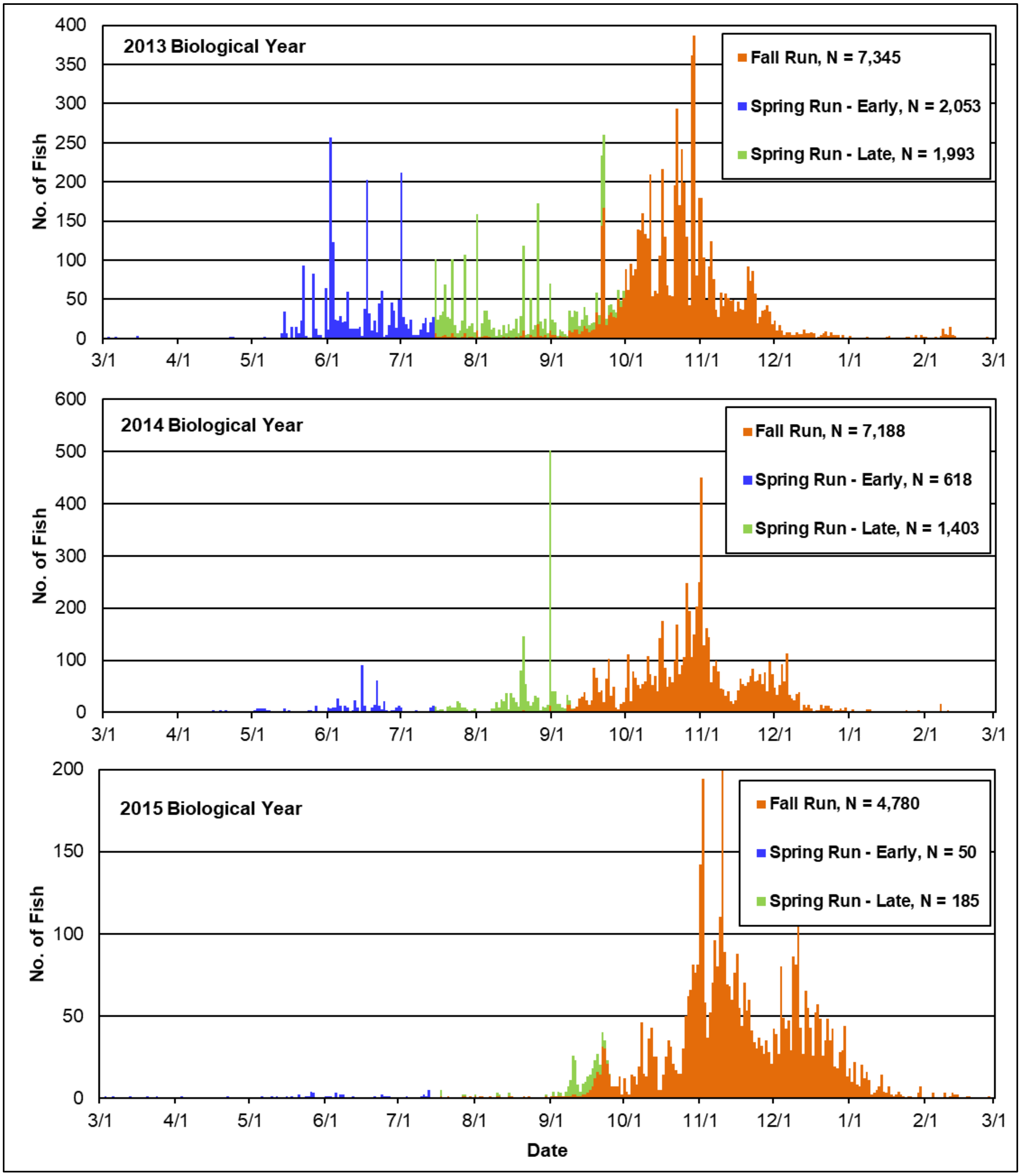



**Figure 13. Daily estimated counts of early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon passing Daguerre Point Dam during the 2013, 2014, and 2015 biological years (March 1 through February 28). Daily counts for each run component are presented as stacked bars, with annual run component escapement estimates provided in figure legends.**

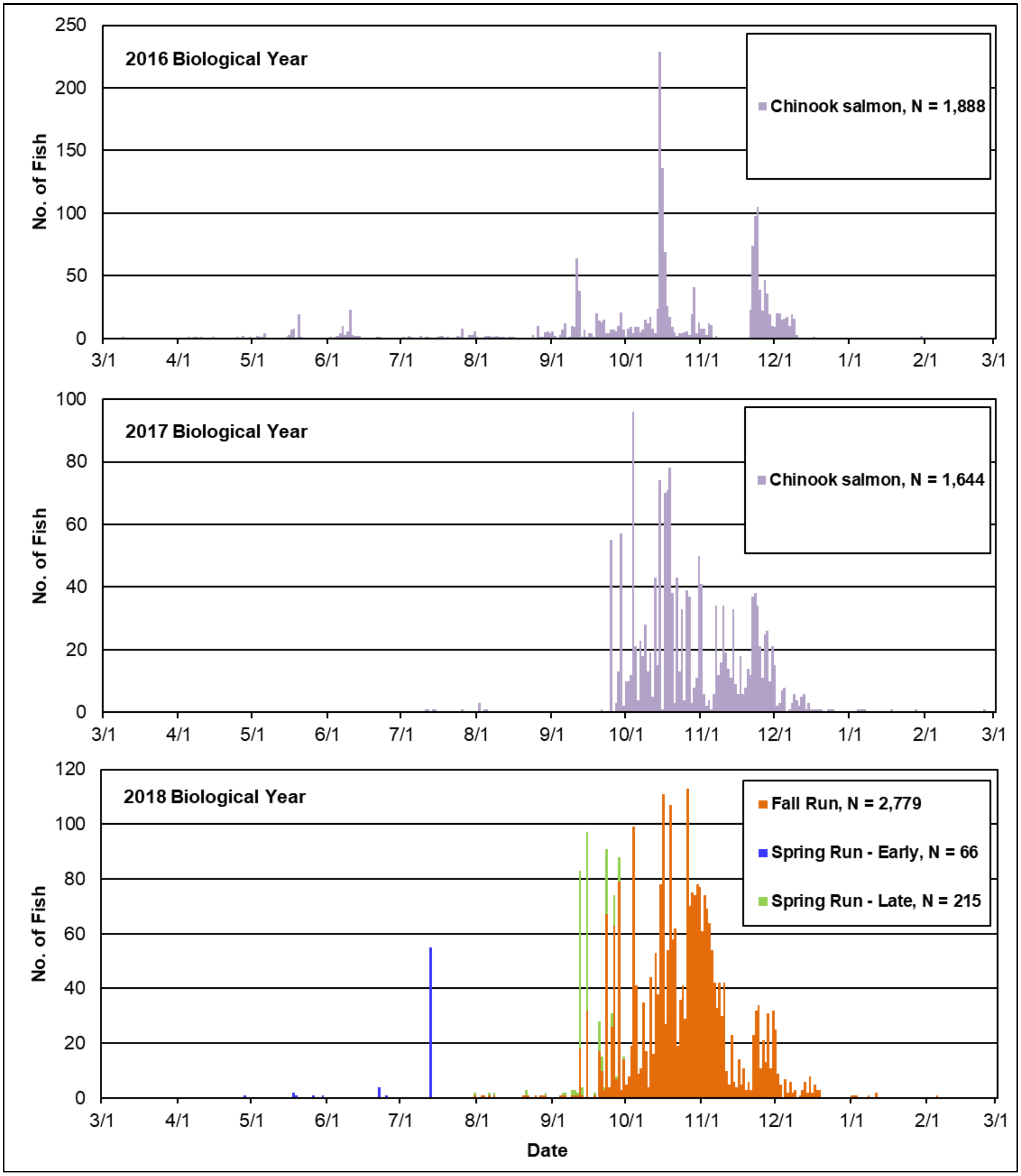



**Figure 14. Daily estimated counts of early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon passing Daguerre Point Dam during the 2016, 2017, and 2018 biological years (March 1 through February 28). Daily counts for each run component are presented as stacked bars, with annual run component escapement estimates provided in figure legends. Differentiation of run components was not conducted for the 2016 and 2017 biological years due to incomplete annual time series of daily counts in those years resulting from extended periods of VAKI Riverwatcher™ outages.**

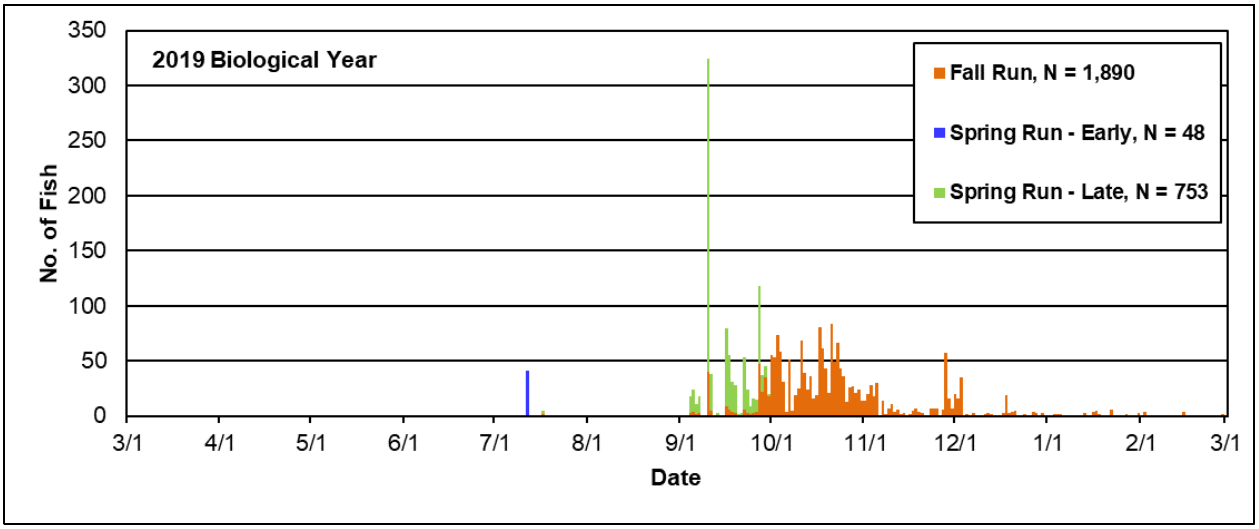



**Figure 15. Daily estimated counts of early-migrating spring-run Chinook salmon, late-migrating spring-run Chinook salmon, and fall-run Chinook salmon passing Daguerre Point Dam during the 2019 biological year (March 1, 2019 through February 29, 2020). Daily counts for each run component are presented as stacked bars, with annual run component escapement estimates provided in the figure legend.**

**Table 2. Results of phenotypic run differentiation analyses for the 2004 through 2019 biological years of spring-run Chinook salmon as presented in RMT (2013) and utilizing the revised methodology presented in this report. Estimates were only available for the years 2004 through 2011 in RMT (2013).**

| Year | Previous Method Spring-Run Chinook Salmon | | Revised Method Spring-Run Chinook Salmon | |
|---|---|---|---|---|
| | Escapement Estimate | % Ad-Clip | Escapement Estimate | % Ad-Clip |
| 2004 | 738 | 10% | 1,618 | 8.0% |
| 2005 | 3,592 | 19% | 5,274 | 15.0% |
| 2006 | 1,326 | 6% | 1,456 | 6.4% |
| 2007 | 372 | 10% | 353 | 10.5% |
| 2008 | 521 | 3% | 1,320 | 3.5% |
| 2009 | 723 | 29% | 2,616 | 20.1% |
| 2010 | 2,886 | 61% | 3,737 | 56.7% |
| 2011 | 1,159 | 28% | 2,360 | 25.4% |
| 2012 | NA | NA | 2,289 | 31.7% |
| 2013 | NA | NA | 4,046 | 6.0% |
| 2014 | NA | NA | 2,021 | 9.8% |
| 2015 | NA | NA | 235 | 13.6% |
| 2016 | NA | NA | NA | NA |
| 2017 | NA | NA | NA | NA |
| 2018 | NA | NA | 281 | 80.1% |
| 2019 | NA | NA | 801 | 19.0% |

Utilizing model AIC scores rather than model deviance for the model selection criterion results in a lower degree of count correction, resulting in more conservative annual passage estimates. In certain biological years (e.g., 2015, 2018, and 2019 biological years), this results in selection of smoother fitted functions that appear to follow the general trend in the data rather than closely approximating each datum. This result appears to be more common in years in which the annual time series exhibit a high degree of variation in daily counts. It should be noted that in most biological years, the AIC scores exhibit an undulating pattern across the range of model $k$ values, and that similarly low AIC scores sometimes existed at both lower and higher values of $k$. As such, and because $k$ is varied in this methodology by 5 value steps (i.e., from $k = 5$ to $k = 100$ by 5), it is possible that different top models associated with larger $k$ values would be selected if $k$ were varied by 1 value steps. While this adjustment to the methodology could result in increased precision in the GAM count correction model fitting process, the trade-off is that any potential improvement would come at a relatively high computational cost, as the GAM fitting process is computationally complex and time consuming.

### 4.3 Run Differentiation

The escapement estimates reported herein are point estimates of escapement without confidence intervals. While desirable, construction of confidence intervals around the individual run component escapement estimates is not feasible for the following reasons. First, construction of confidence intervals around non-parametric, non-linear models is difficult due to a lack of theoretical basis. Second, confidence intervals would need to incorporate and aggregate error from the count correction analysis in addition to the run differentiation analysis. Finally, the run differentiation methodology combines two individual non-linear models constructed on different datasets, which further complicates calculation of overall error.

Incorporating the passage timing information from the acoustic tagging study allowed for refinement of the run differentiation assumptions and methodology to more accurately reflect the available data. However, confronting the results of this revised approach with genetic information collected at DPD would allow further refinement and would likely result in the most accurate escapement estimation for each run component. As previously mentioned, recent advances in genetic analyses have shown improved success with differentiating spring-run and fall-run Chinook salmon on a genetic basis. Evaluation of potential methods for identifying the genetic makeup of adult Chinook salmon as they pass upstream through DPD fish ladders could prove to be beneficial to run differentiation.

## 5.0 Literature Cited

Afifi, A., V. A. Clark, and S. May. 2004. Computer-Aided Multivariate Analysis, Fourth Edition. Chapman and Hall/CRC, New York.

Allison, P. D. and R. P. Waterman. 2002. Fixed-Effects Negative Binomial Regression Models. Sociological Methodology 32(2002):247-265.

Burnham, K. P. and D. R. Anderson. 2002. Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach, Second Edition. Springer, New York.

Crawley, M. J. 2007. The R Book. John Wiley and Sons Ltd., England.

Grafen, A. and R. Hails. 2002. Modern Statistics for the Life Sciences. Oxford University Press, New York.

Grothendieck, G. 2013. nls2: Non-Linear Regression with Brute Force. R Package Version 0.2. URL https://CRAN.R-project.org/package=nls2.

Hastie, T. 2020. gam: Generalized Additive Models. R Package Version 1.20. URL https://CRAN.R-project.org/package=gam/.

Hess, J. E., J. S. Zendt, A. R. Matala, and S. R. Narum. 2016. Genetic Basis of Adult Migration Timing in Anadromous Steelhead Discovered Through Multivariate Association Testing. Proceedings of the Royal Society B 283:20153064.

Kock, I. J. and S. R. Narum. 2020. Validation and Association of Candidate Markers for Adult Migration Timing and Fitness in Chinook Salmon. Evolutionary Applications 13:2316-2332.

Kormos, B., M. Palmer-Zwahlen, and A. Low. 2012. Recovery of Coded-wire Tags from Chinook Salmon in California's Central Valley Escapement and Ocean Harvest in 2010. California Department of Fish and Wildlife, Fisheries Branch Administrative Report 2012-02.

Lambert, D. 1992. Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing. Technometrics 34(1):1-14.

Lower Yuba River Management Team (RMT). 2010. Lower Yuba River Accord Monitoring and Evaluation Program. Draft. June 28, 2010.

Lower Yuba River Management Team (RMT). 2013. Aquatic Resources of the Lower Yuba River: Past, Present, and Future. Draft Interim Report.

Meek, M. H., M. R. Stephens, A. Goodbla, B. May, and M. R. Baewald. 2020. Identifying Hidden Biocomplexity and Genetic Diversity in Chinook Salmon, An Imperiled Species with a History of Anthropogenic Influence. Canadian Journal of Fisheries and Aquatic Sciences 77:534-547.

Morandat, F., B. Hill, L. Osvald, J. Vitek. 2012. Evaluating the Design of the R Language: Objects and Functions for Data Analysis. ECOOP'12 Proceedings of the 26th European Conference on Object-Oriented Programming.

Narum, S. R., A. Di Genova, S. J. Micheletti, and A. Maass. 2018. Genomic Variation Underlying Complex Life-History Traits Revealed by Genome Sequencing in Chinook Salmon. Proceedings of the Royal Society B 285:20180935.

Palmer-Zwahlen, M. and B. Kormos. 2013. Recovery of Coded-wire Tags from Chinook Salmon in California's Central Valley Escapement and Ocean Harvest in 2011. California Department of Fish and Wildlife, Fisheries Branch Administrative Report 2013-02.

Palmer-Zwahlen, M. and B. Kormos. 2015. Recovery of Coded-wire Tags from Chinook Salmon in California's Central Valley Escapement, Inland Harvest, and Ocean Harvest in 2012. California Department of Fish and Wildlife, Fisheries Administrative Report 2015-04.

Palmer-Zwahlen, M. and B. Kormos. 2018. Recovery of Coded-wire Tags from Chinook Salmon in California's Central Valley Escapement, Inland Harvest, and Ocean Harvest in 2013. Prepared by Pacific Marine Fisheries Commission and California Department of Fish and Wildlife.

Palmer-Zwahlen, M., V. Gusman, and B. Kormos. 2019a. Recovery of Coded-wire Tags from Chinook Salmon in California's Central Valley Escapement, Inland Harvest, and Ocean Harvest in 2014. Prepared by Pacific Marine Fisheries Commission and California Department of Fish and Wildlife. March 2019.

Palmer-Zwahlen, M., V. Gusman, and B. Kormos. 2019b. Recovery of Coded-wire Tags from Chinook Salmon in California's Central Valley Escapement, Inland Harvest, and Ocean Harvest in 2015. Prepared by Pacific Marine Fisheries Commission and California Department of Fish and Wildlife. December 2019.

Prince, D. J., S. M. O'Rourke, T. Q. Thompson, O. A. Ali, H. S. Lyman, I. K. Saglam, T. J. Hotaling, A. P. Spidle, and M. R. Miller. 2017. The Evolutionary Basis of Premature Migration in Pacific Salmon Highlights the Utility of Genomics for Informing Conservation. Science Advances 2017(3):1-11.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richards, F. J. 1959. A Flexible Growth Function for Empirical Use. Journal of Experimental Botany 10:290-300.

Thompson, N. F., E. C. Anderson, A. J. Clemento, M. A. Campbell, D. E. Pearse, J. W. Hearsey, A. P. Kinziger, and J. C. Garza. 2020. A Complex Phenotype in Salmon Controlled by a Simple Change in Migratory Timing. Science 370:609-613.

Thompson, T. Q., M. R. Bellinger, S. M. O'Rourke, D. J. Prince, A. E. Stevenson, A. T. Rodrigues, M. R. Sloat, C. F. Speller, D. Y. Yang, V. L. Butler, M. A. Banks, and M. R. Miller. 2019. Anthropogenic Habitat Alteration Leads to Rapid Loss of Adaptive Variation and Restoration Potential in Wild Salmon Populations. Proceedings of the National Academy of Sciences of the United States of America 116(1):177-186.

U.S. Army Corps of Engineers (USACE). 2012. Biological Assessment for the U. S. Army Corps of Engineers Ongoing Operation and Maintenance of Englebright Dam and Reservoir, and Daguerre Point Dam on the Lower Yuba River. January 2012.

U.S. Bureau of Reclamation (Reclamation). 2008. Biological Assessment on the Continued Long-Term Operations of the Central Valley Project and the State Water Project. U.S. Department of the Interior, Bureau of Reclamation, Mid-Pacific Region. August 2008.

Ver Hoef, J. M. and P. L. Boveng. 2007. Quasi-Poisson vs. Negative Binomial Regression: How Should We Model Overdispersed Count Data? Ecology 88(11):2766-2772.

Vogel, D. A. and K. R. Marine. 1991. Guide to Upper Sacramento River Chinook Salmon Life History. U.S. Bureau of Reclamation Central Valley Project. July 1991.

Wood, S. N. 2020. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. URL https://CRAN.R-project.org/package=mgcv/.

Wood, S. N. 2017. Generalized Additive Models: An Introduction with R (Second Edition). Chapman and Hall/CRC.

Wood, S. N., N. Pya, and B. Safken. 2016. Smoothing Parameter and Model Selection for General Smooth Models. Journal of the American Statistical Association 111(516):1548-1575.

Yoshiyama, R. M., E. R. Gerstung, F. W. Fisher, and P. B. Moyle. 1996. Historical and Present Distribution of Chinook Salmon in the Central Valley Drainage of California. In: Sierra Nevada Ecosystem Project: Final Report to Congress, Volume III, Assessments, Commissioned Reports, and Background Information. University of California, Davis, Centers for Water and Wildland Resources.

Yuba County Water Agency, California Department of Water Resources, and U.S. Bureau of Reclamation. 2007. Draft Environmental Impact Report/Environmental Impact Statement for the Proposed Lower Yuba River Accord. State Clearinghouse (SCH) No: 2005062111. Prepared by HDR Engineering, Inc.|Surface Water Resources, Inc. June 2007.

Zar, J. H. 1999. Biostatistical Analysis, Fourth Edition. Prentice Hall, New Jersey.

# ATTACHMENT A

RMT (2013) Methodology to Correct Daily VAKI Riverwatcher™ Net Upstream Chinook Salmon Passage Counts for Periods of VAKI Riverwatcher™ System Non-Operation

# ATTACHMENT A

The following text is directly excerpted from RMT (2013, pp. 5-43 and 5-44) regarding the methodology previously used by the RMT to correct daily VAKI Riverwatcher™ net upstream Chinook salmon passage counts for periods of VAKI Riverwatcher™ system non-operation.

---

Prior to applying any analysis of temporal modalities to the 8 annual time series of Chinook salmon daily Vaki counts, the annual daily count series at each ladder were adjusted to account for days when the Vaki Riverwatcher systems were not fully operational. The procedure used to obtain complete annual daily count series of Chinook salmon migrating upstream of Daguerre Point Dam is summarized as follows.

- For each biological year (March 1 through February 28) and ladder, the series of observed daily counts of Chinook salmon passing upstream of Daguerre Point Dam was fitted to a smoothing spline, with days within the biological year as the only predictor (d), 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 were used as weights in the minimization process used to obtain the smoothing spline.
- Various smoothing splines were fitted to each of the 16 available time series (north and south ladder, respectively, for each of the 8 biological years) varying in the value of their smoothing parameter $\lambda$. The residual deviance for each fitted individual spline was plotted against its parameter $\lambda$ to detect the threshold at which the residual deviance was no longer minimized with increasing value of $\lambda$. The $\lambda$ value associated with that threshold was used to identify the best fitted spline for each time series.
- Once the smoothing spline that best fit the data was obtained, the values predicted by the spline on days of partial operation of the Vaki Riverwatcher systems were rounded to the nearest integer and compared with the corresponding observed count. If the prediction was higher than the observed count, the prediction was accepted as the corrected count. If the prediction was lower than the observed count, the prediction was discarded and the observed count was considered to be the corrected count.
- Once this process was finished for the series of daily counts of the north ladder and the south ladder, the resulting corrected series were combined into one series of corrected daily counts of Chinook salmon moving upstream of Daguerre Point Dam by summing the corresponding daily values for each ladder.

---

The text and materials provided below, which describe statistical analyses performed during development of the methodology used to correct daily VAKI Riverwatcher ™ net upstream passage counts for periods of VAKI Riverwatcher™ system non-operation, are excerpted from the Appendix (including two attachments, dated April 5, 2011) to the Lower Yuba River Accord Monitoring and Evaluation Program (RMT 2010).

# 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:

$$\sum_{i=1}^{n}(Y_i - \hat{\mu}(x_i))^2 + \lambda \int_{x1}^{x_n} \hat{\mu}''(x)^2 \, dx.$$

  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 $\lambda$. The parameter $\lambda$ 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 $\lambda$ 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 $\lambda$ to detect the threshold at which the residual deviance no longer minimized with increasing value of $\lambda$. The $\lambda$ 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).

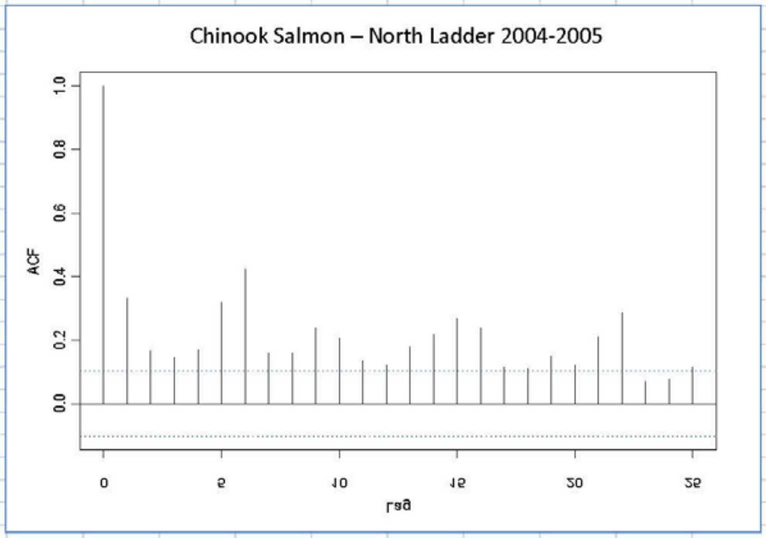
Results of the daily count estimation process incorporating all of the above-mentioned components and steps are presented in **Attachment 2**. On each individual page in Attachment 2, the top figure presents the fitted spline, and the bottom figure represents the resultant estimated daily count of Chinook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

## References

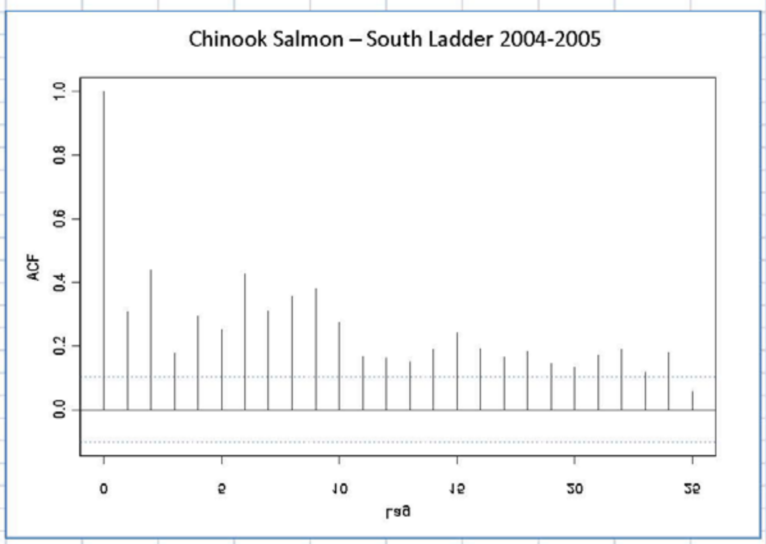
Freund, J. E. and G. A. Simon. 1991. Statistics: A first course, 5th ed. Prentice Hall, Englewood Cliffs, New York.

# Attachment 1

| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.332484 |
| 2 | 0.165417 |
| 3 | 0.143653 |
| 4 | 0.169026 |
| 5 | 0.319879 |
| 6 | 0.424415 |
| 7 | 0.15806 |
| 8 | 0.157998 |
| 9 | 0.23748 |
| 10 | 0.20496 |
| 11 | 0.13499 |
| 12 | 0.12107 |
| 13 | 0.178347 |
| 14 | 0.217564 |
| 15 | 0.267869 |
| 16 | 0.237551 |
| 17 | 0.114753 |
| 18 | 0.111175 |
| 19 | 0.149273 |
| 20 | 0.12123 |
| 21 | 0.210535 |
| 22 | 0.284462 |
| 23 | 0.07049 |
| 24 | 0.075578 |
| 25 | 0.114394 |

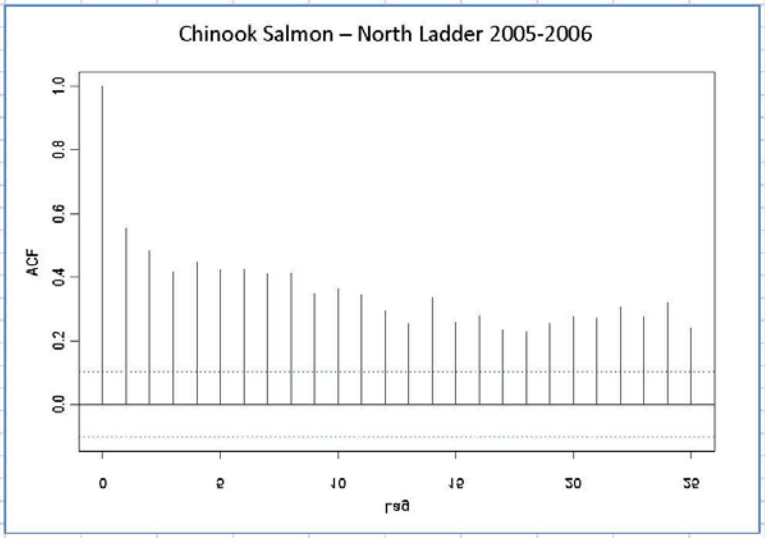



| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.307642 |
| 2 | 0.438355 |
| 3 | 0.17787 |
| 4 | 0.293988 |
| 5 | 0.252669 |
| 6 | 0.427168 |
| 7 | 0.309284 |
| 8 | 0.355975 |
| 9 | 0.381808 |
| 10 | 0.273496 |
| 11 | 0.16638 |
| 12 | 0.161429 |
| 13 | 0.150019 |
| 14 | 0.188376 |
| 15 | 0.242763 |
| 16 | 0.190439 |
| 17 | 0.165151 |
| 18 | 0.182692 |
| 19 | 0.144723 |
| 20 | 0.133552 |
| 21 | 0.171191 |
| 22 | 0.187775 |
| 23 | 0.117266 |
| 24 | 0.178831 |
| 25 | 0.058316 |

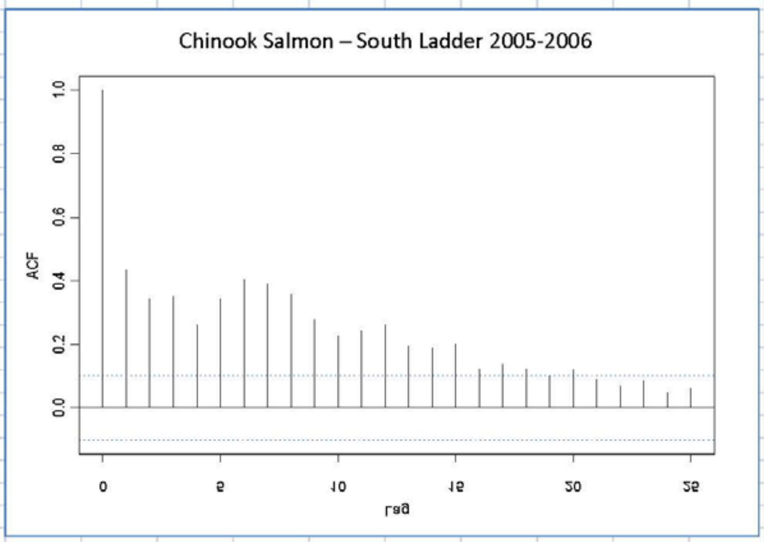


## Attachment 1

| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.553376 |
| 2 | 0.483803 |
| 3 | 0.417102 |
| 4 | 0.446273 |
| 5 | 0.421249 |
| 6 | 0.425389 |
| 7 | 0.40949 |
| 8 | 0.412346 |
| 9 | 0.349821 |
| 10 | 0.362973 |
| 11 | 0.345037 |
| 12 | 0.293446 |
| 13 | 0.254098 |
| 14 | 0.337314 |
| 15 | 0.256552 |
| 16 | 0.27989 |
| 17 | 0.231917 |
| 18 | 0.229148 |
| 19 | 0.25519 |
| 20 | 0.275675 |
| 21 | 0.272282 |
| 22 | 0.304775 |
| 23 | 0.275825 |
| 24 | 0.321216 |
| 25 | 0.239684 |

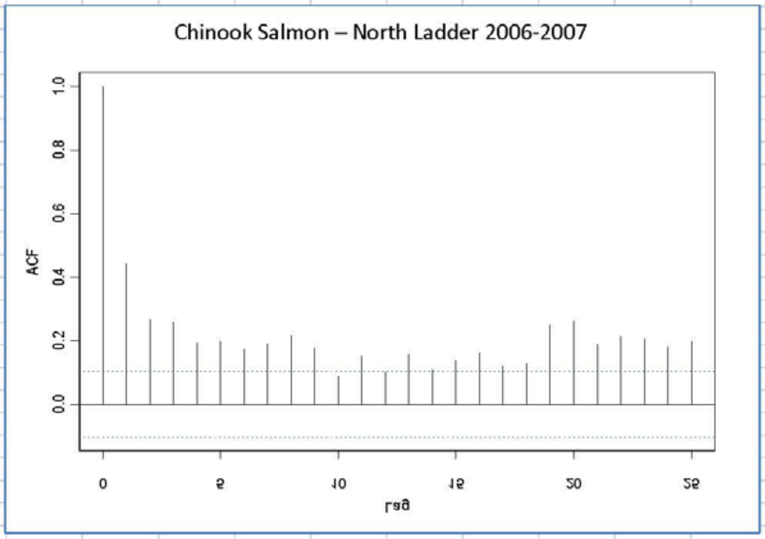



| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.432565 |
| 2 | 0.344246 |
| 3 | 0.350539 |
| 4 | 0.260078 |
| 5 | 0.344593 |
| 6 | 0.405363 |
| 7 | 0.389306 |
| 8 | 0.357419 |
| 9 | 0.279157 |
| 10 | 0.227528 |
| 11 | 0.243133 |
| 12 | 0.26209 |
| 13 | 0.1964 |
| 14 | 0.187886 |
| 15 | 0.200884 |
| 16 | 0.122342 |
| 17 | 0.135992 |
| 18 | 0.121425 |
| 19 | 0.103041 |
| 20 | 0.119391 |
| 21 | 0.090018 |
| 22 | 0.068281 |
| 23 | 0.086866 |
| 24 | 0.046505 |
| 25 | 0.06183 |

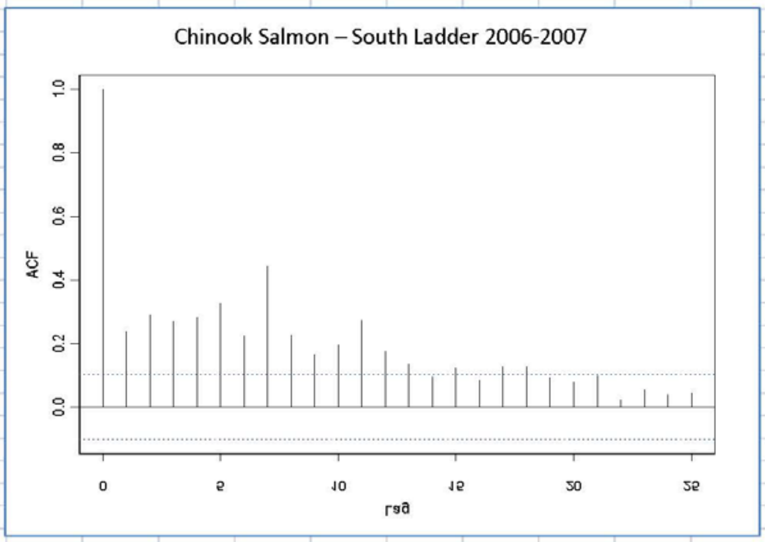


# Attachment 1

| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.441191 |
| 2 | 0.265152 |
| 3 | 0.258536 |
| 4 | 0.194257 |
| 5 | 0.199478 |
| 6 | 0.174619 |
| 7 | 0.19048 |
| 8 | 0.214758 |
| 9 | 0.176689 |
| 10 | 0.087137 |
| 11 | 0.151689 |
| 12 | 0.100053 |
| 13 | 0.156199 |
| 14 | 0.107656 |
| 15 | 0.137904 |
| 16 | 0.161752 |
| 17 | 0.119982 |
| 18 | 0.12828 |
| 19 | 0.250292 |
| 20 | 0.261741 |
| 21 | 0.189308 |
| 22 | 0.213384 |
| 23 | 0.205767 |
| 24 | 0.180438 |
| 25 | 0.197433 |

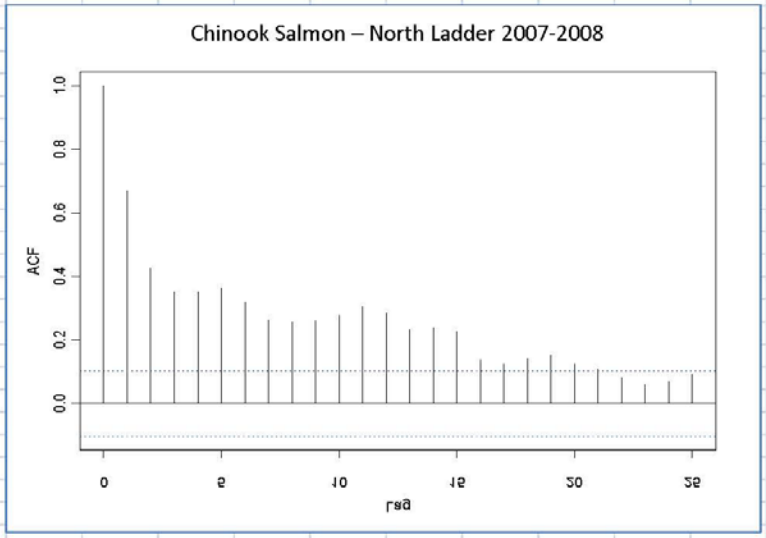



| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.238863 |
| 2 | 0.28969 |
| 3 | 0.271851 |
| 4 | 0.281755 |
| 5 | 0.326686 |
| 6 | 0.22527 |
| 7 | 0.443945 |
| 8 | 0.226582 |
| 9 | 0.1647 |
| 10 | 0.194372 |
| 11 | 0.272131 |
| 12 | 0.174876 |
| 13 | 0.133685 |
| 14 | 0.095035 |
| 15 | 0.123082 |
| 16 | 0.086169 |
| 17 | 0.127285 |
| 18 | 0.126093 |
| 19 | 0.092301 |
| 20 | 0.078624 |
| 21 | 0.098585 |
| 22 | 0.021444 |
| 23 | 0.053197 |
| 24 | 0.040213 |
| 25 | 0.044569 |

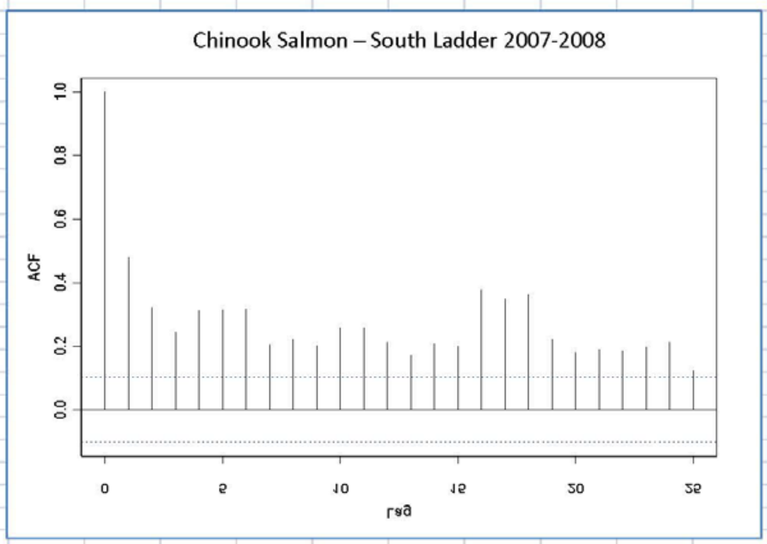


# Attachment 1

| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.670004 |
| 2 | 0.427063 |
| 3 | 0.350845 |
| 4 | 0.349958 |
| 5 | 0.362965 |
| 6 | 0.319424 |
| 7 | 0.264523 |
| 8 | 0.254928 |
| 9 | 0.260897 |
| 10 | 0.278569 |
| 11 | 0.306103 |
| 12 | 0.285958 |
| 13 | 0.232333 |
| 14 | 0.24032 |
| 15 | 0.226056 |
| 16 | 0.137542 |
| 17 | 0.12588 |
| 18 | 0.141077 |
| 19 | 0.152293 |
| 20 | 0.125519 |
| 21 | 0.107622 |
| 22 | 0.08056 |
| 23 | 0.05869 |
| 24 | 0.069272 |
| 25 | 0.090406 |

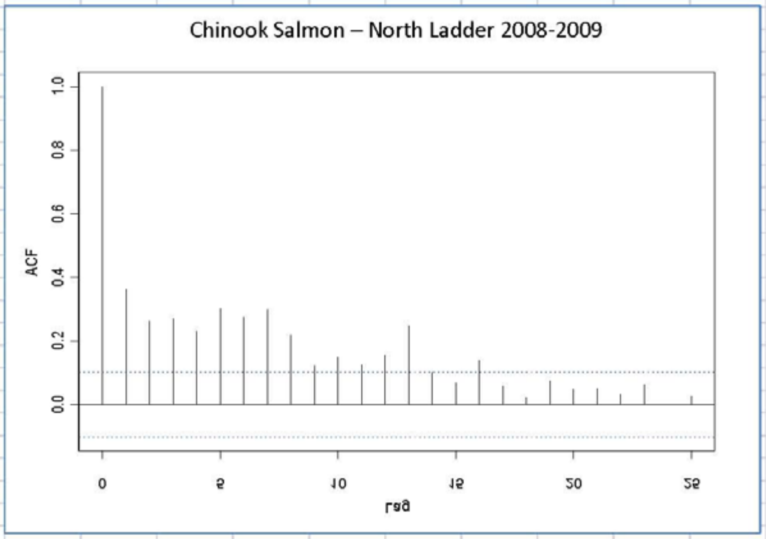



| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.480065 |
| 2 | 0.322231 |
| 3 | 0.243845 |
| 4 | 0.312133 |
| 5 | 0.314723 |
| 6 | 0.317313 |
| 7 | 0.205315 |
| 8 | 0.223184 |
| 9 | 0.202856 |
| 10 | 0.258921 |
| 11 | 0.258456 |
| 12 | 0.212155 |
| 13 | 0.171965 |
| 14 | 0.208169 |
| 15 | 0.200064 |
| 16 | 0.378357 |
| 17 | 0.348863 |
| 18 | 0.362148 |
| 19 | 0.22112 |
| 20 | 0.180931 |
| 21 | 0.189632 |
| 22 | 0.186111 |
| 23 | 0.197869 |
| 24 | 0.211154 |
| 25 | 0.122073 |

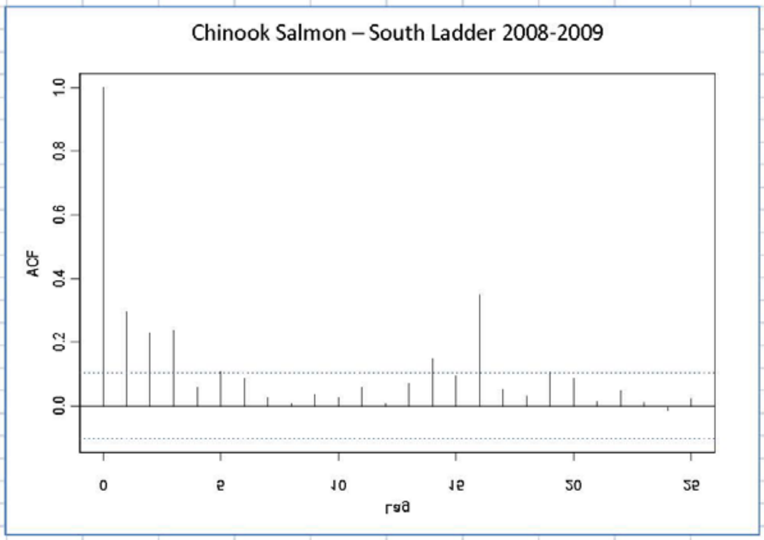


# Attachment 1

| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.362083 |
| 2 | 0.260895 |
| 3 | 0.269651 |
| 4 | 0.231396 |
| 5 | 0.301314 |
| 6 | 0.273984 |
| 7 | 0.298659 |
| 8 | 0.218334 |
| 9 | 0.120172 |
| 10 | 0.14852 |
| 11 | 0.123675 |
| 12 | 0.154663 |
| 13 | 0.246943 |
| 14 | 0.102397 |
| 15 | 0.068596 |
| 16 | 0.138702 |
| 17 | 0.056955 |
| 18 | 0.021132 |
| 19 | 0.074142 |
| 20 | 0.047498 |
| 21 | 0.050664 |
| 22 | 0.030331 |
| 23 | 0.063366 |
| 24 | -0.00066 |
| 25 | 0.026859 |

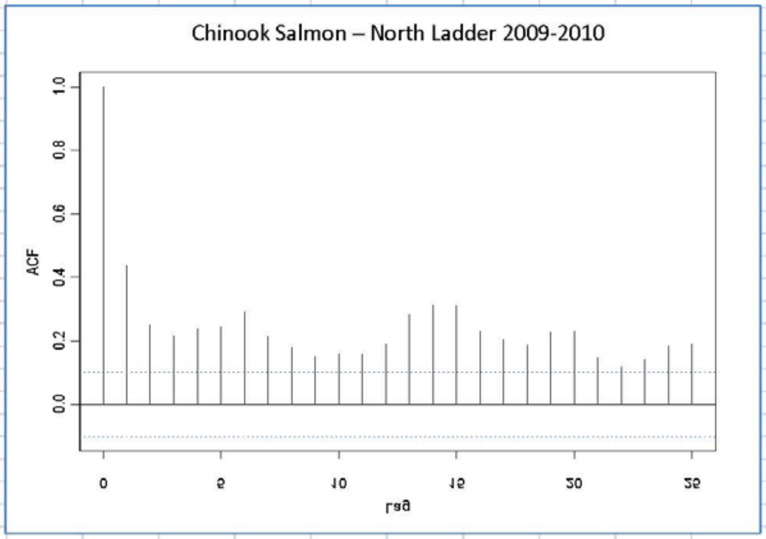



| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.296342 |
| 2 | 0.229869 |
| 3 | 0.236689 |
| 4 | 0.057241 |
| 5 | 0.107466 |
| 6 | 0.087441 |
| 7 | 0.025183 |
| 8 | 0.007519 |
| 9 | 0.036176 |
| 10 | 0.024724 |
| 11 | 0.057819 |
| 12 | 0.008872 |
| 13 | 0.069081 |
| 14 | 0.149479 |
| 15 | 0.092768 |
| 16 | 0.349323 |
| 17 | 0.051449 |
| 18 | 0.031122 |
| 19 | 0.105752 |
| 20 | 0.086979 |
| 21 | 0.014071 |
| 22 | 0.048323 |
| 23 | 0.012022 |
| 24 | -0.01274 |
| 25 | 0.022397 |

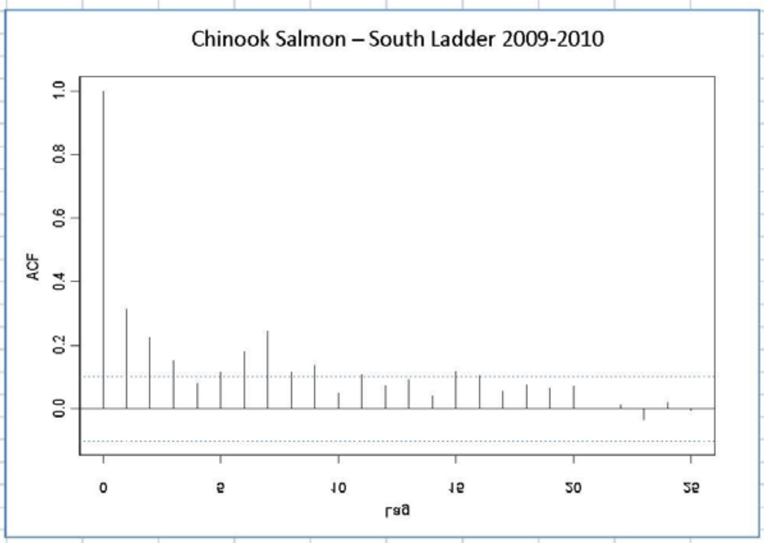


# Attachment 1

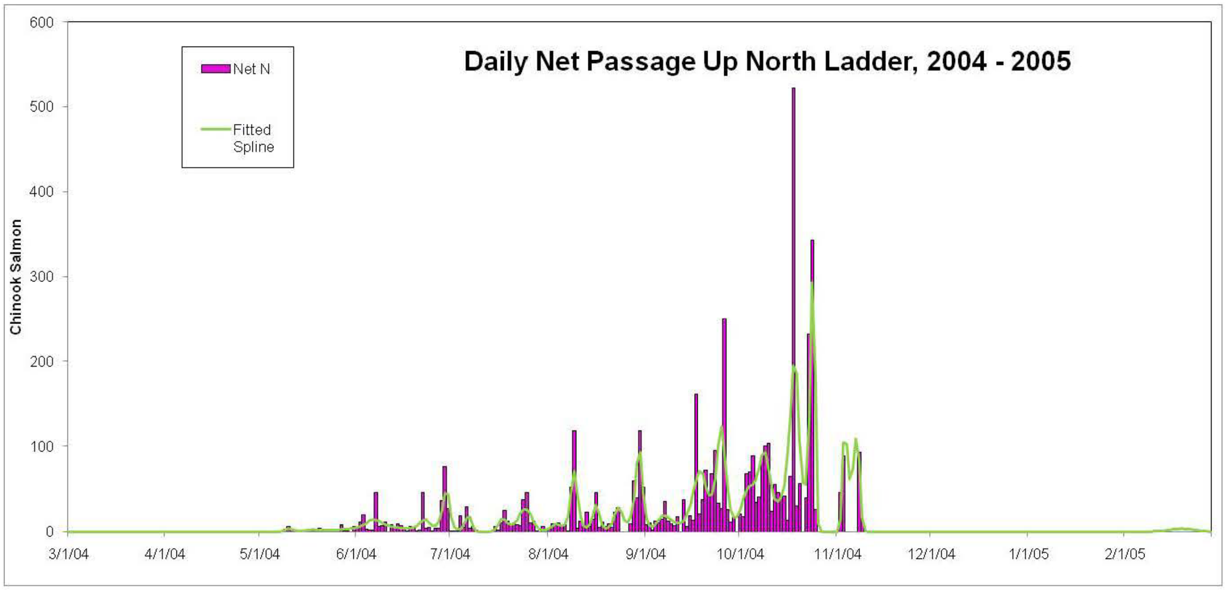
| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.437242 |
| 2 | 0.249558 |
| 3 | 0.217107 |
| 4 | 0.236967 |
| 5 | 0.24443 |
| 6 | 0.291428 |
| 7 | 0.212802 |
| 8 | 0.179927 |
| 9 | 0.150371 |
| 10 | 0.160072 |
| 11 | 0.157404 |
| 12 | 0.19021 |
| 13 | 0.281277 |
| 14 | 0.312418 |
| 15 | 0.311458 |
| 16 | 0.230073 |
| 17 | 0.203133 |
| 18 | 0.187144 |
| 19 | 0.229549 |
| 20 | 0.229864 |
| 21 | 0.146808 |
| 22 | 0.11913 |
| 23 | 0.140469 |
| 24 | 0.184012 |
| 25 | 0.188514 |

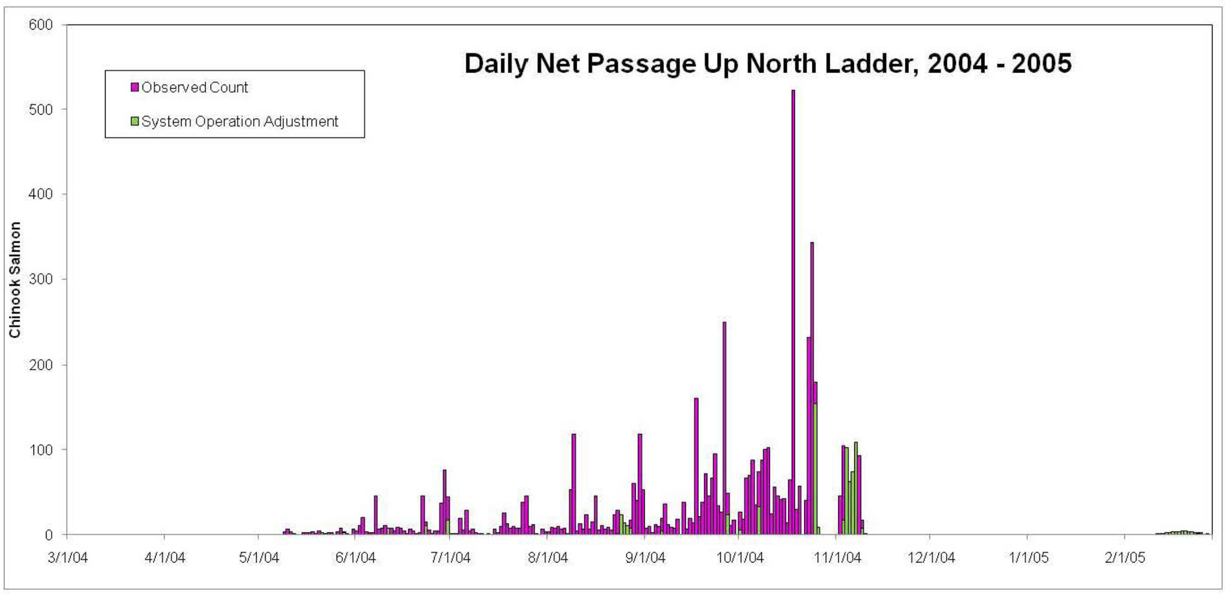



| Lag | Auto Correlation |
|---|---|
| 0 | 1 |
| 1 | 0.313693 |
| 2 | 0.222393 |
| 3 | 0.151083 |
| 4 | 0.079773 |
| 5 | 0.112845 |
| 6 | 0.180291 |
| 7 | 0.242446 |
| 8 | 0.114693 |
| 9 | 0.13628 |
| 10 | 0.048507 |
| 11 | 0.107722 |
| 12 | 0.072865 |
| 13 | 0.0921 |
| 14 | 0.039605 |
| 15 | 0.115282 |
| 16 | 0.103944 |
| 17 | 0.055564 |
| 18 | 0.074798 |
| 19 | 0.064047 |
| 20 | 0.070347 |
| 21 | 0.000213 |
| 22 | 0.012392 |
| 23 | -0.03481 |
| 24 | 0.018524 |
| 25 | -0.00516 |

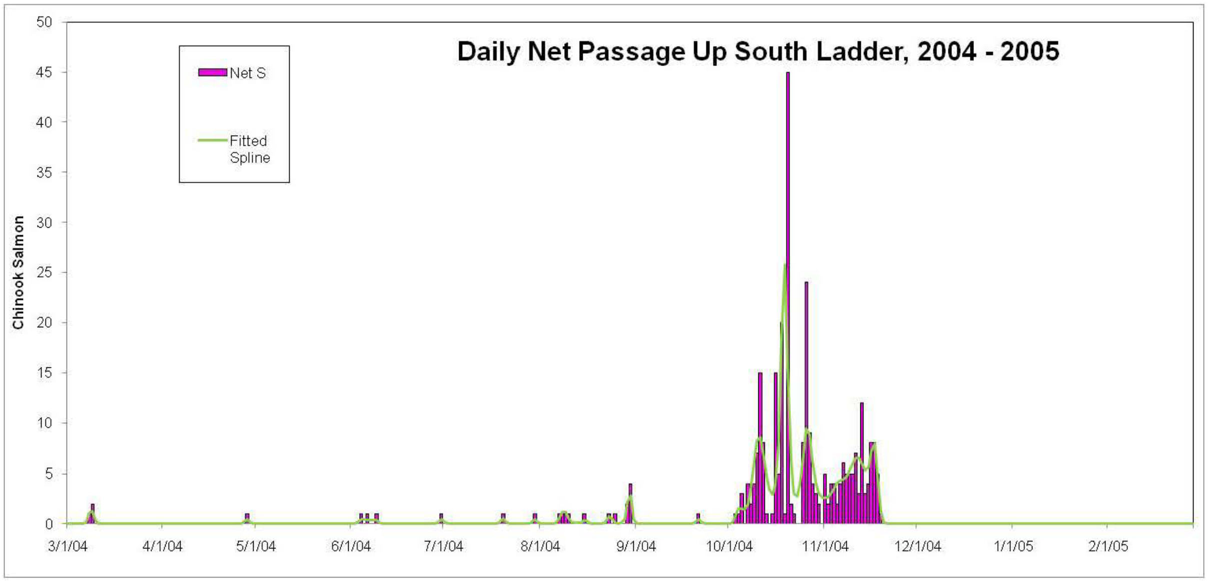


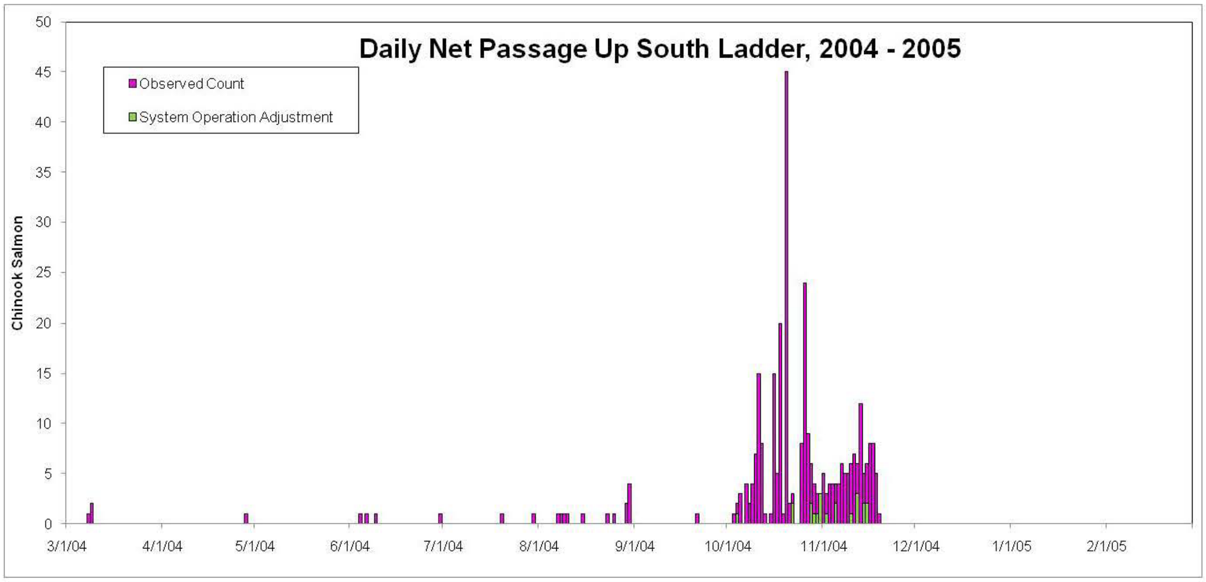
# Attachment 2



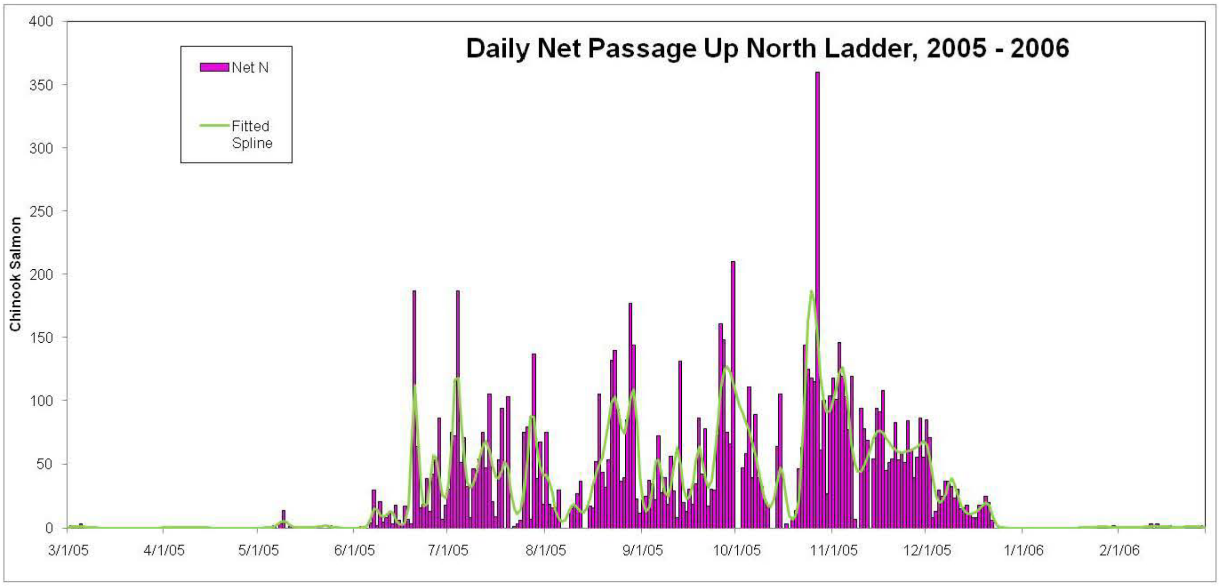



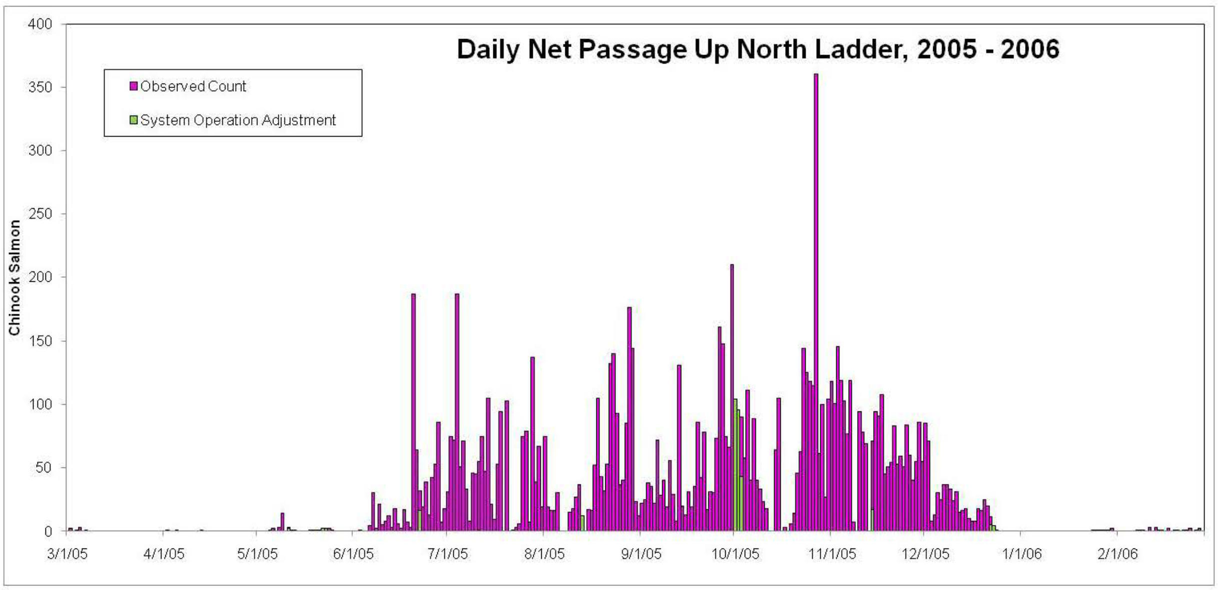
# Attachment 2



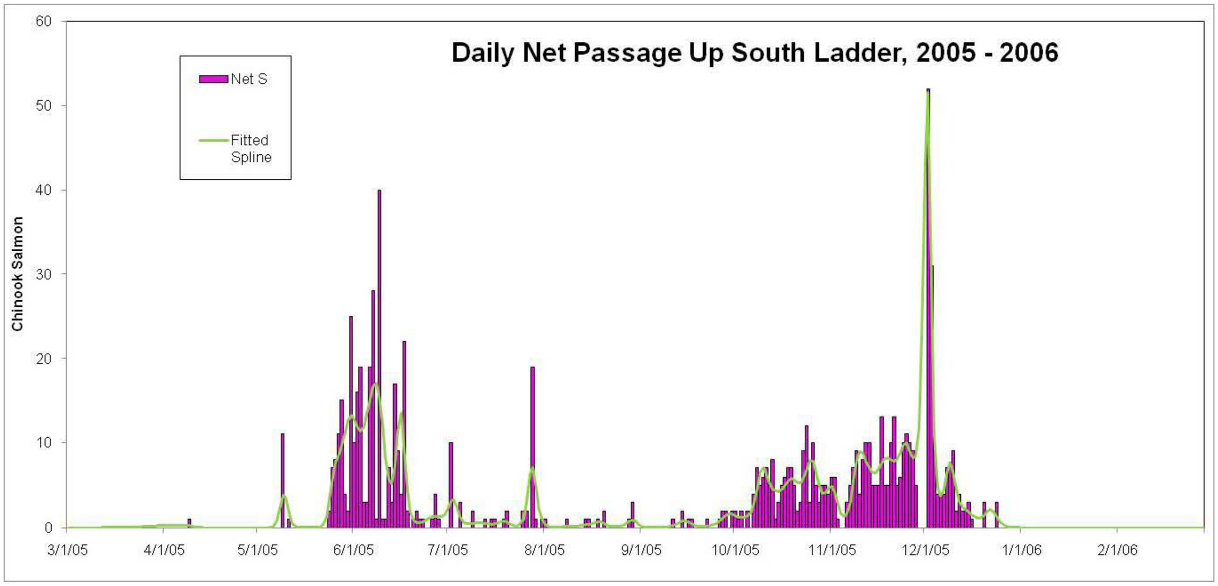



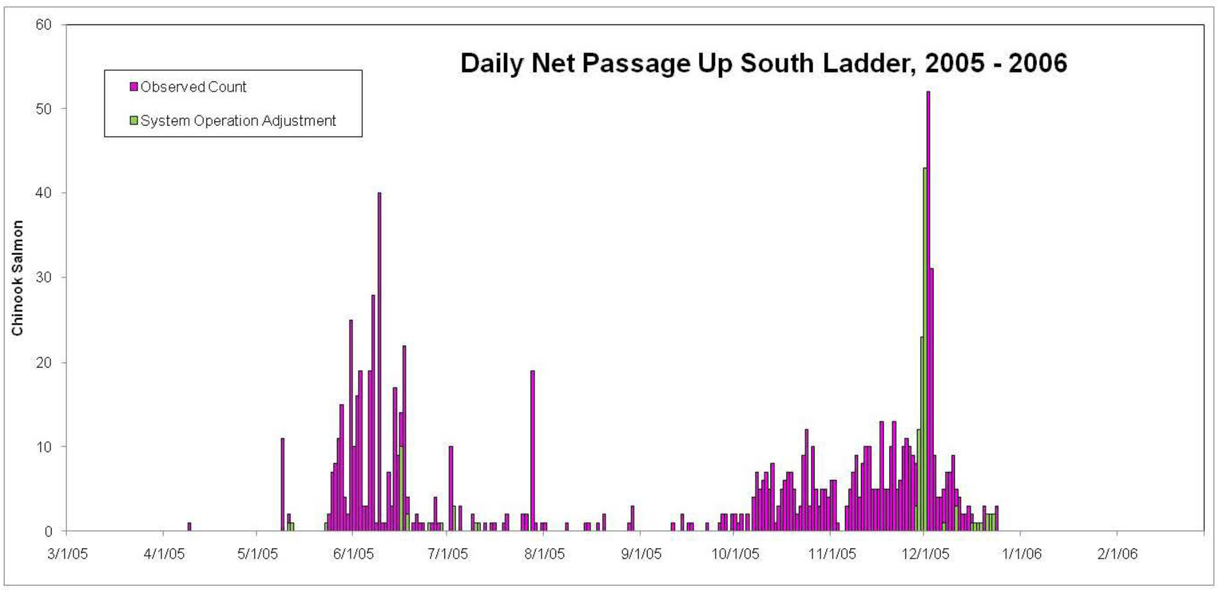
# Attachment 2



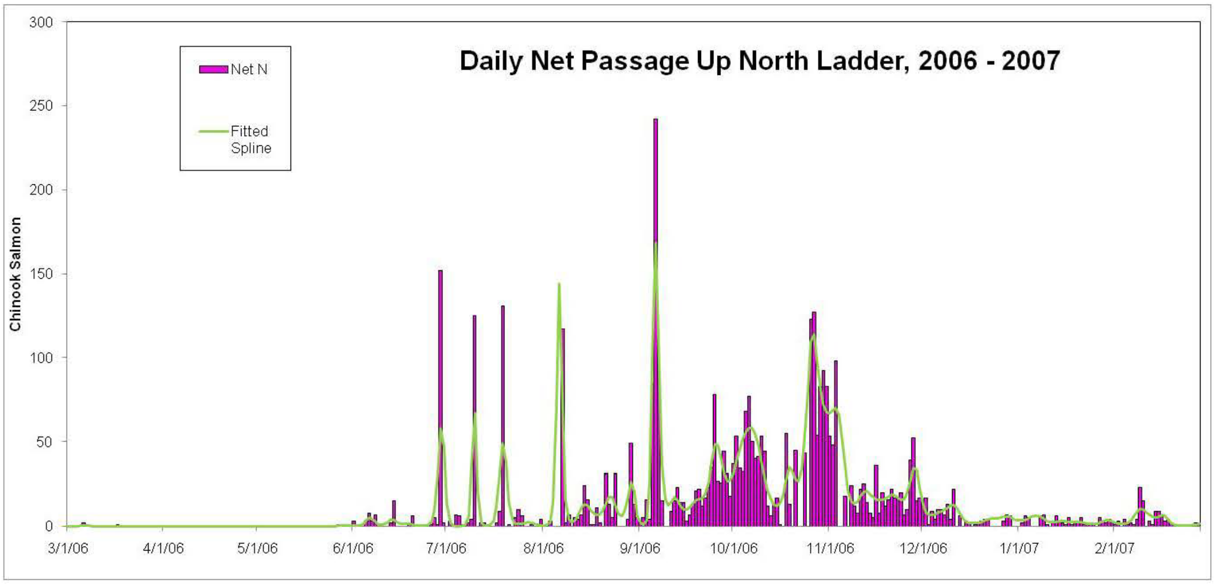



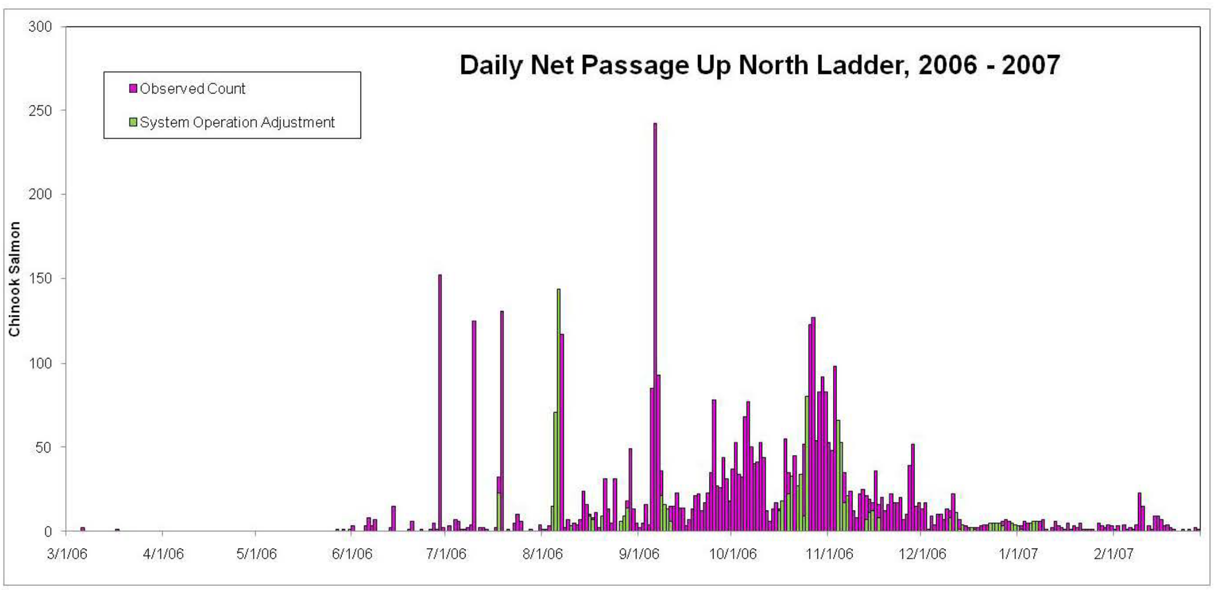
# Attachment 2



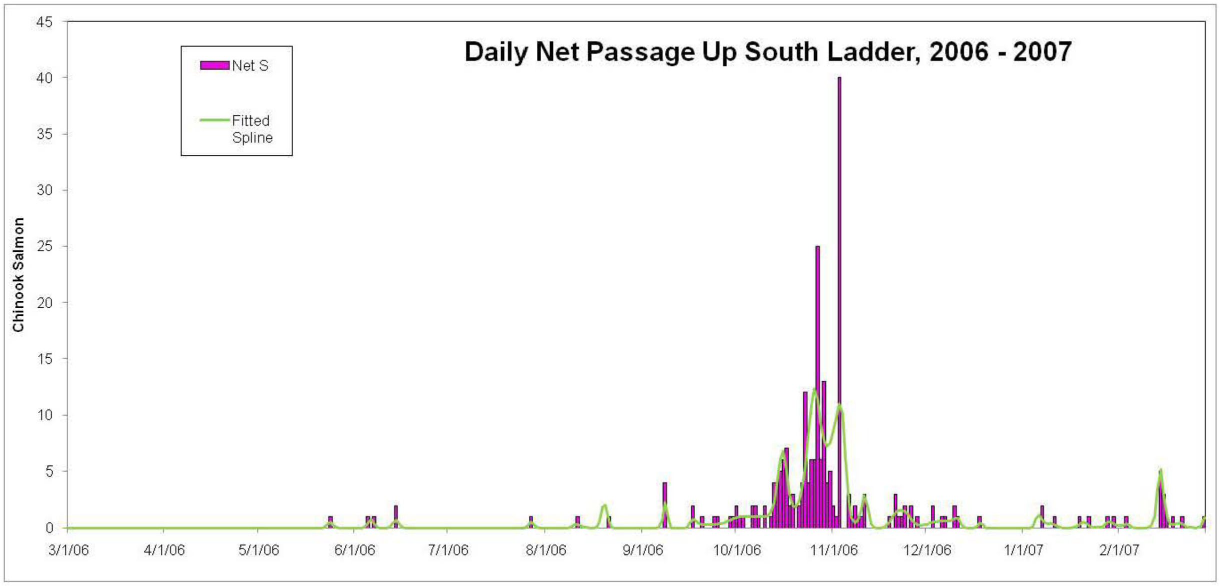



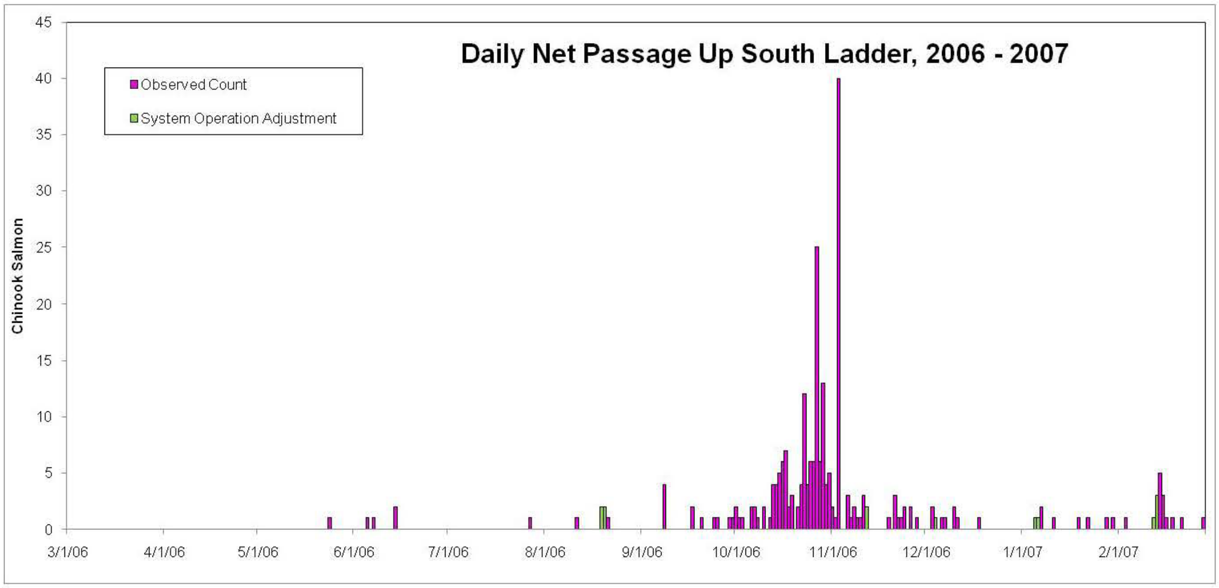
# Attachment 2



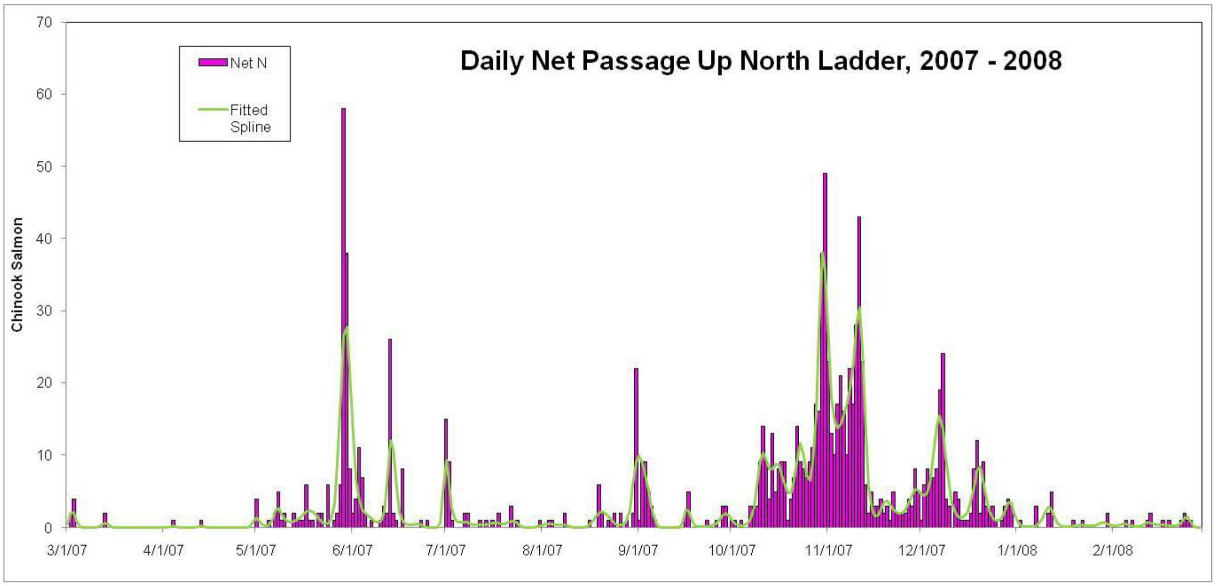



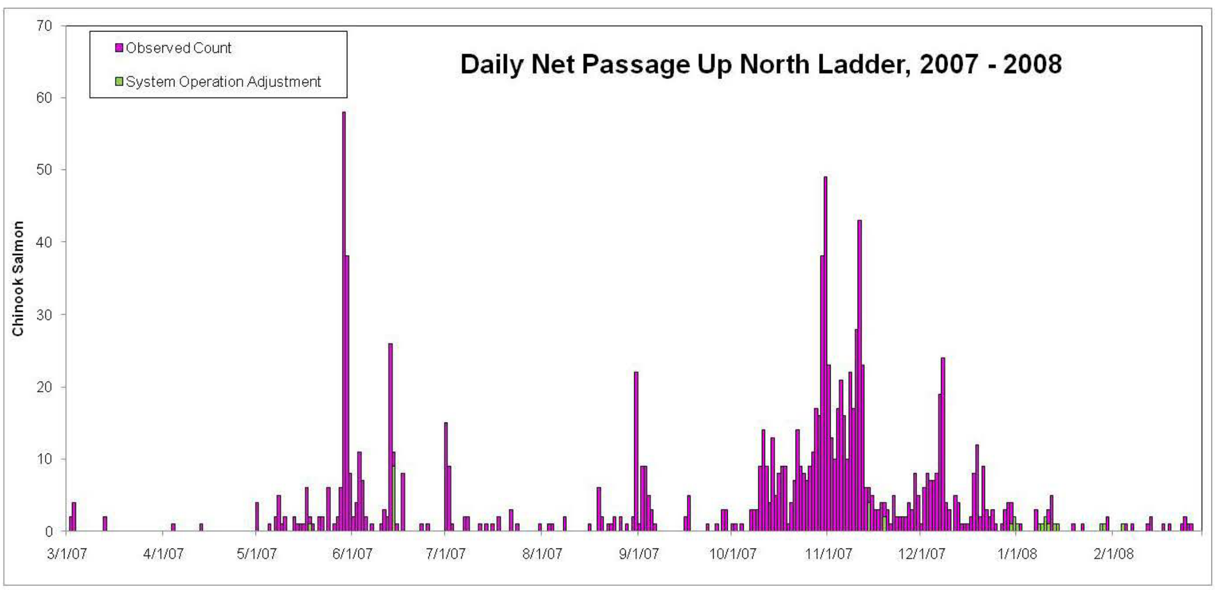
# Attachment 2



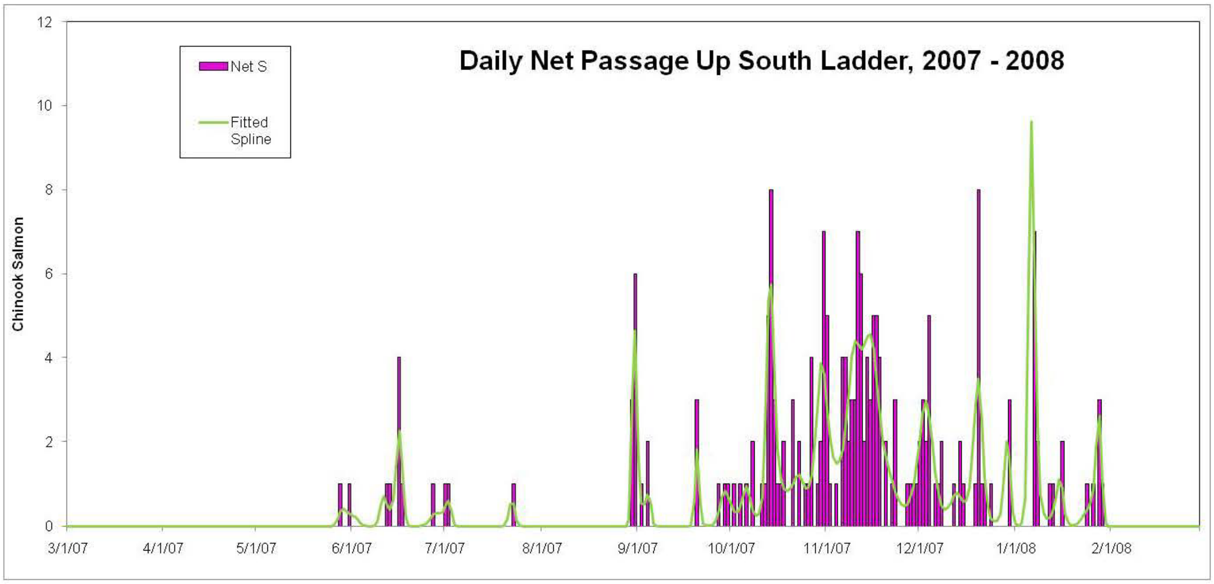



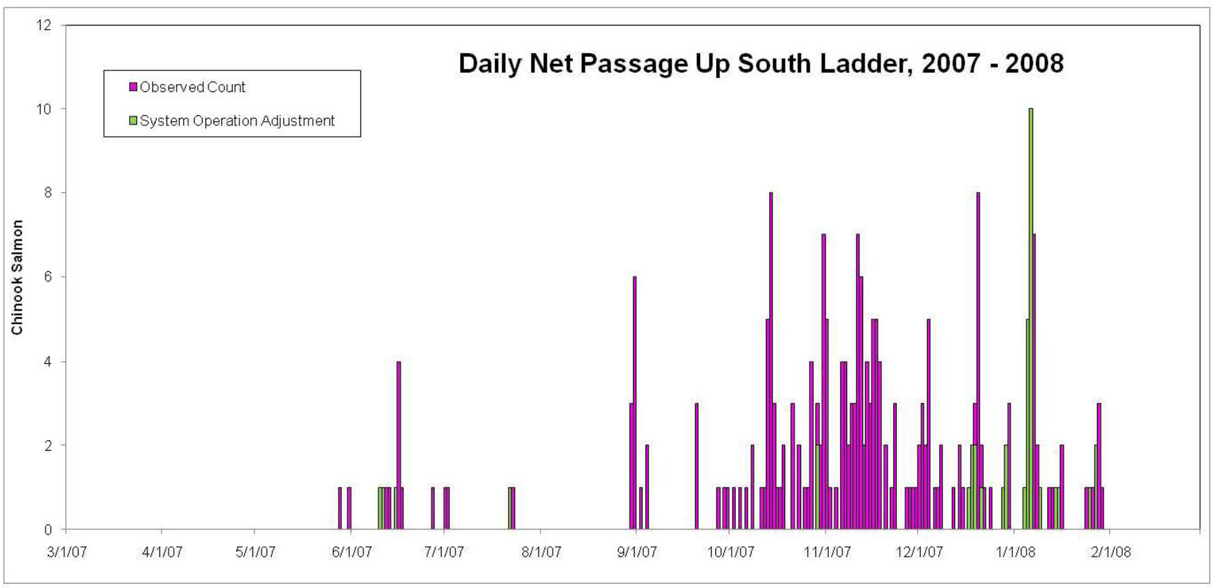
# Attachment 2



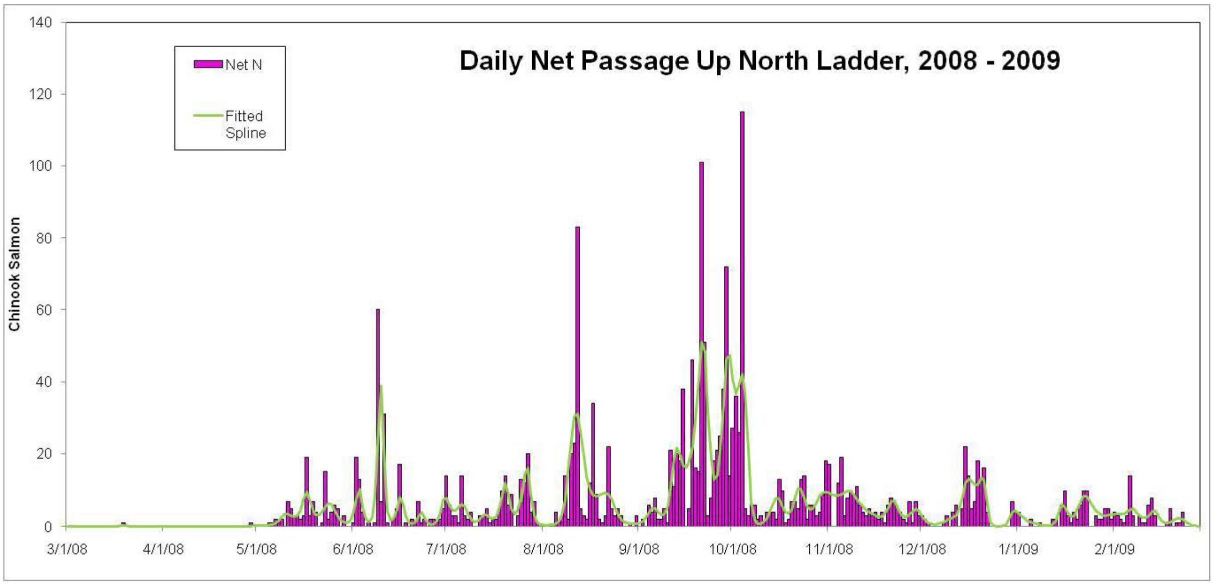



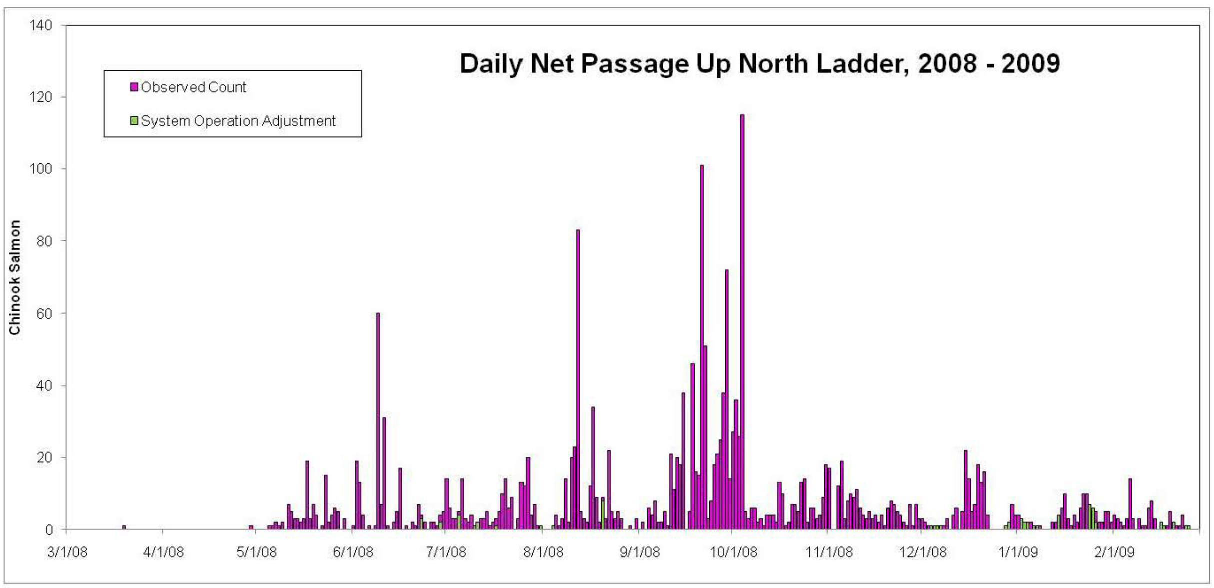
# Attachment 2



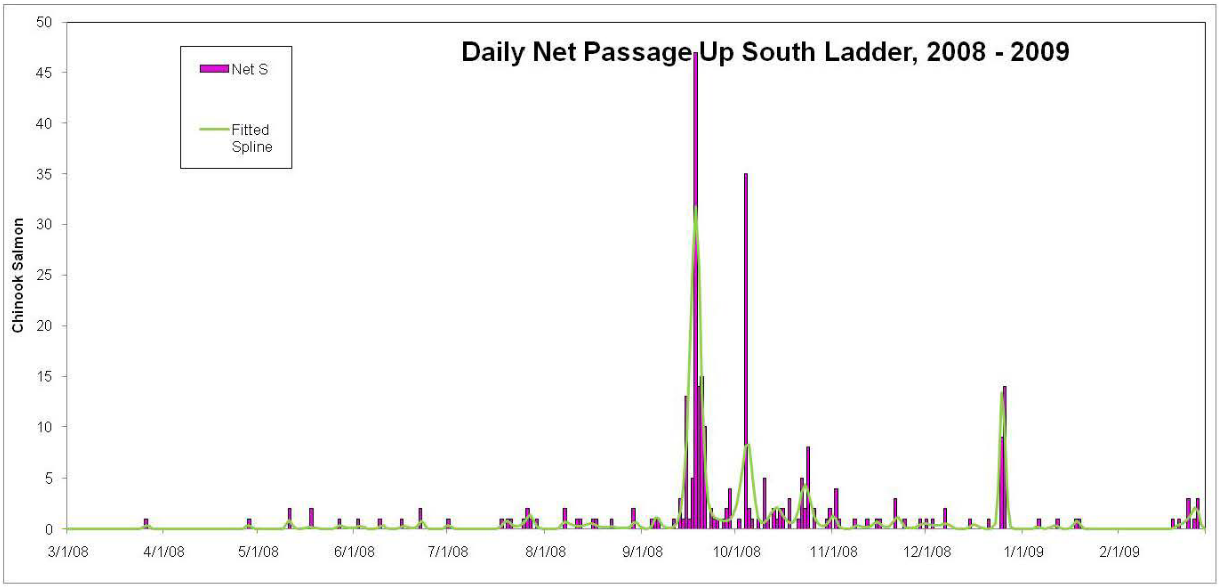



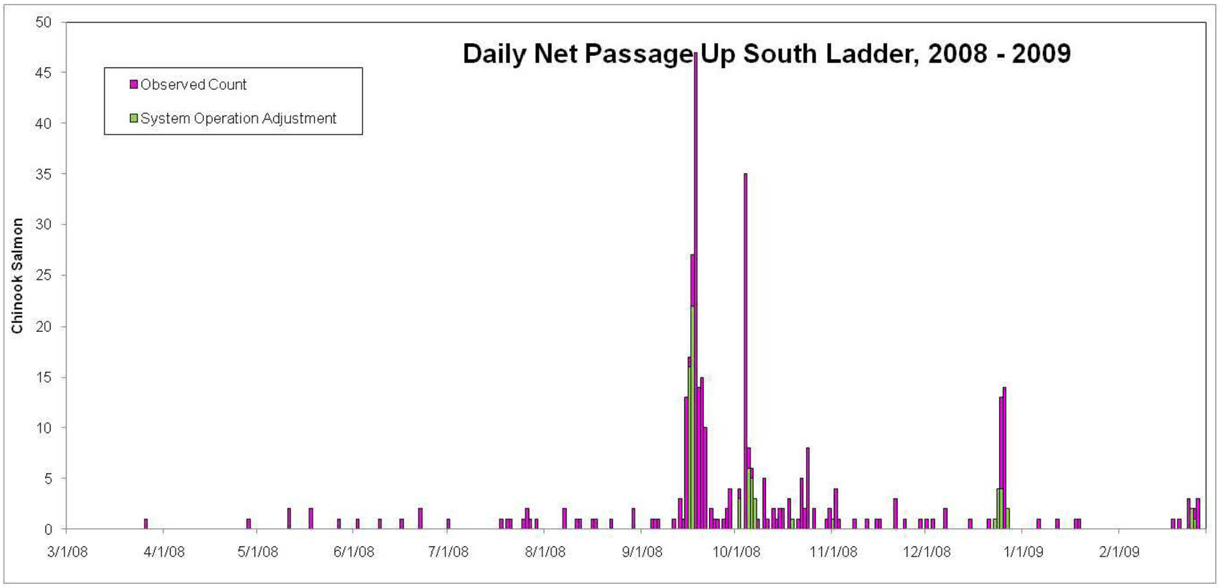
# Attachment 2



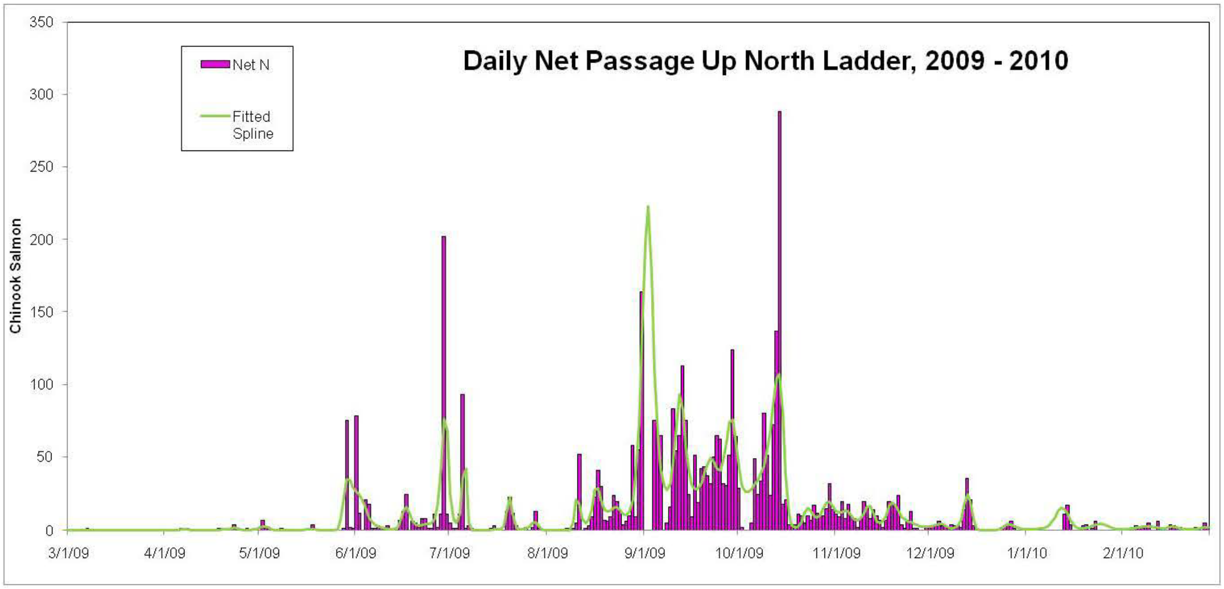



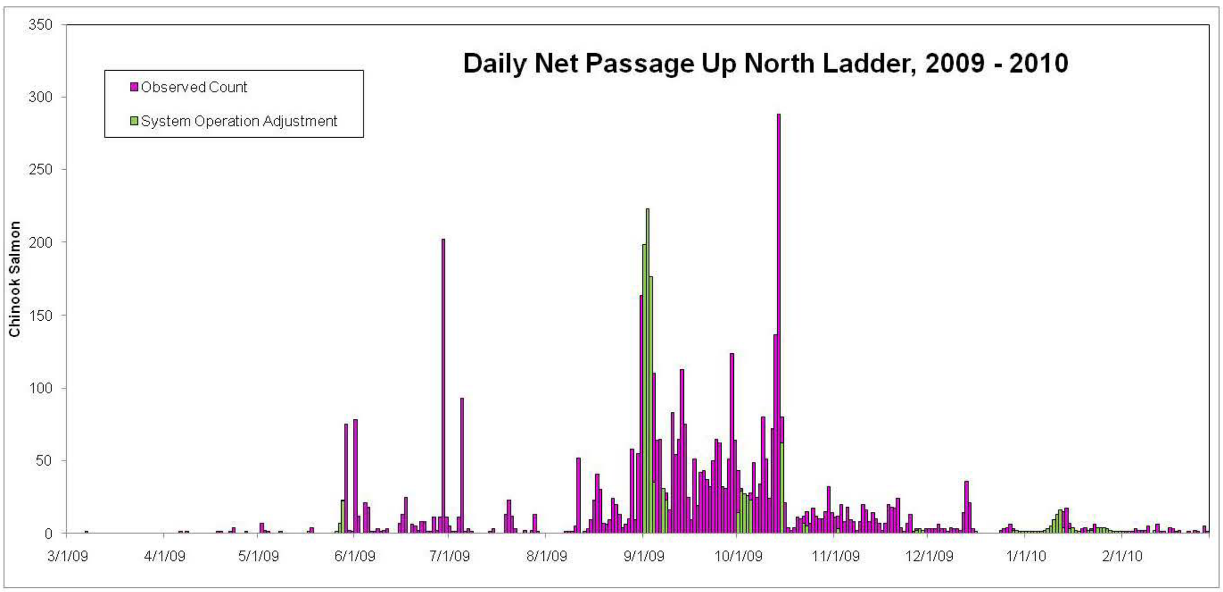
# Attachment 2



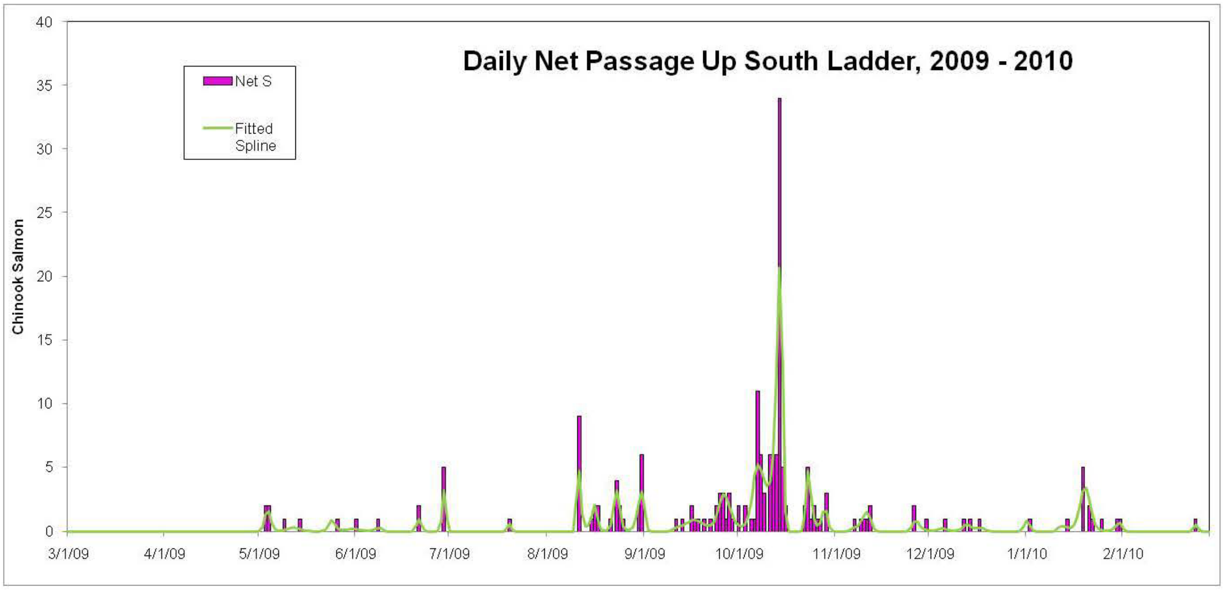



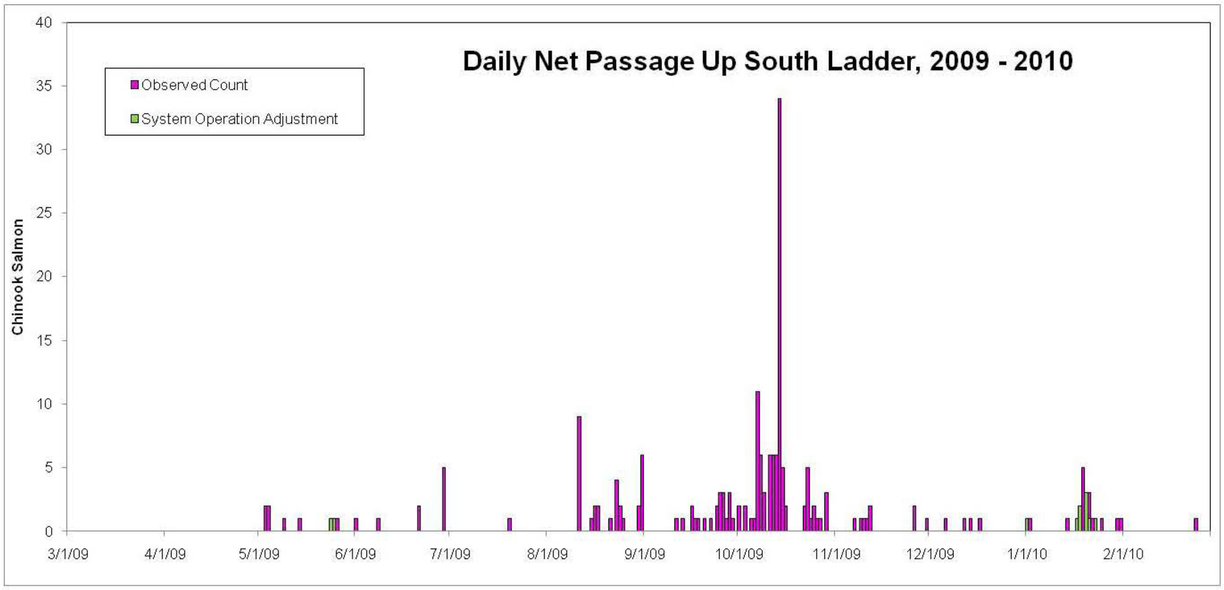
# Attachment 2



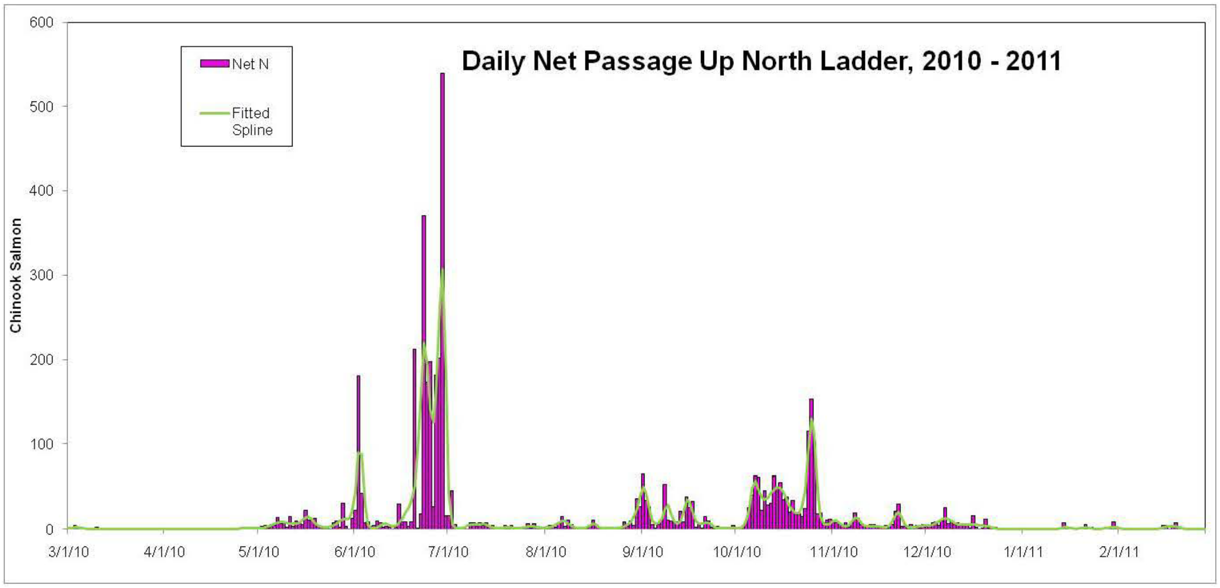



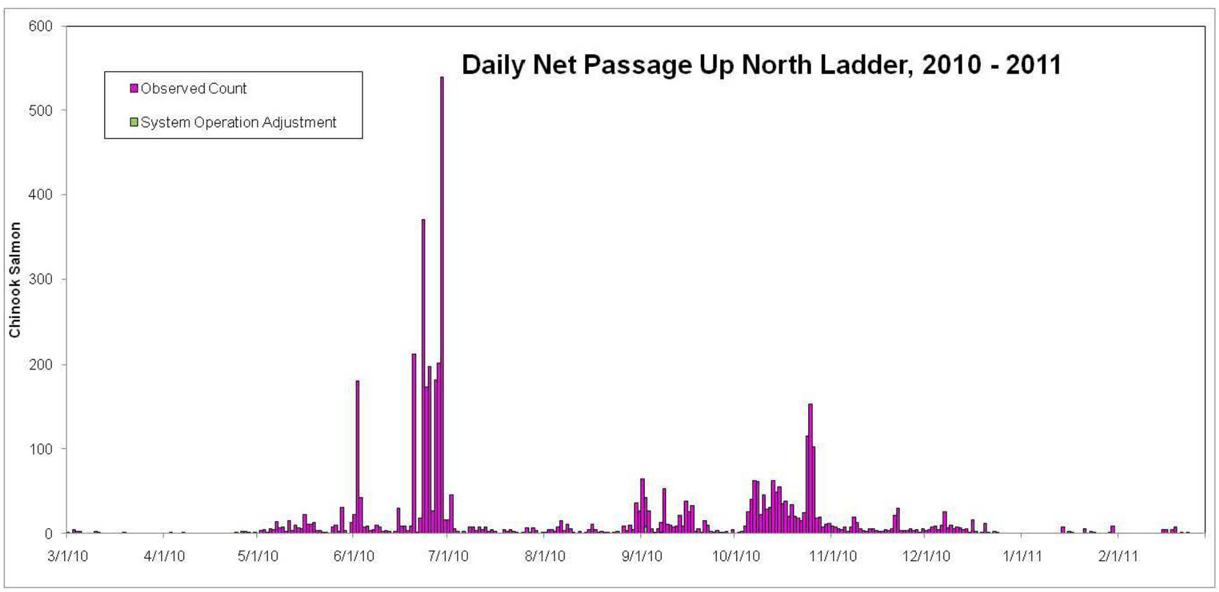
# Attachment 2



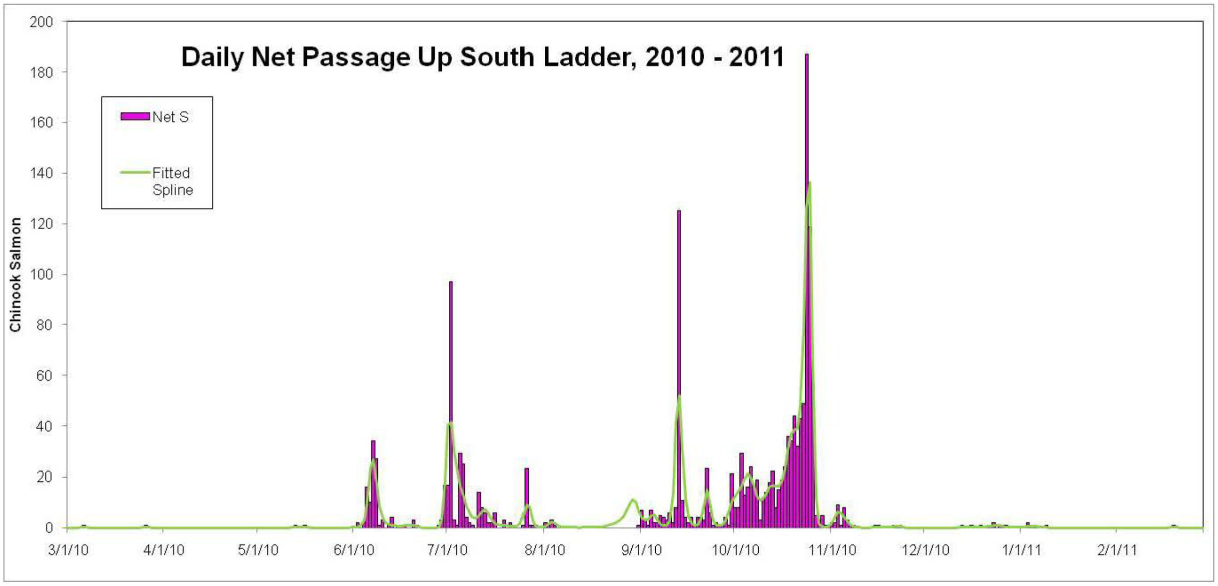



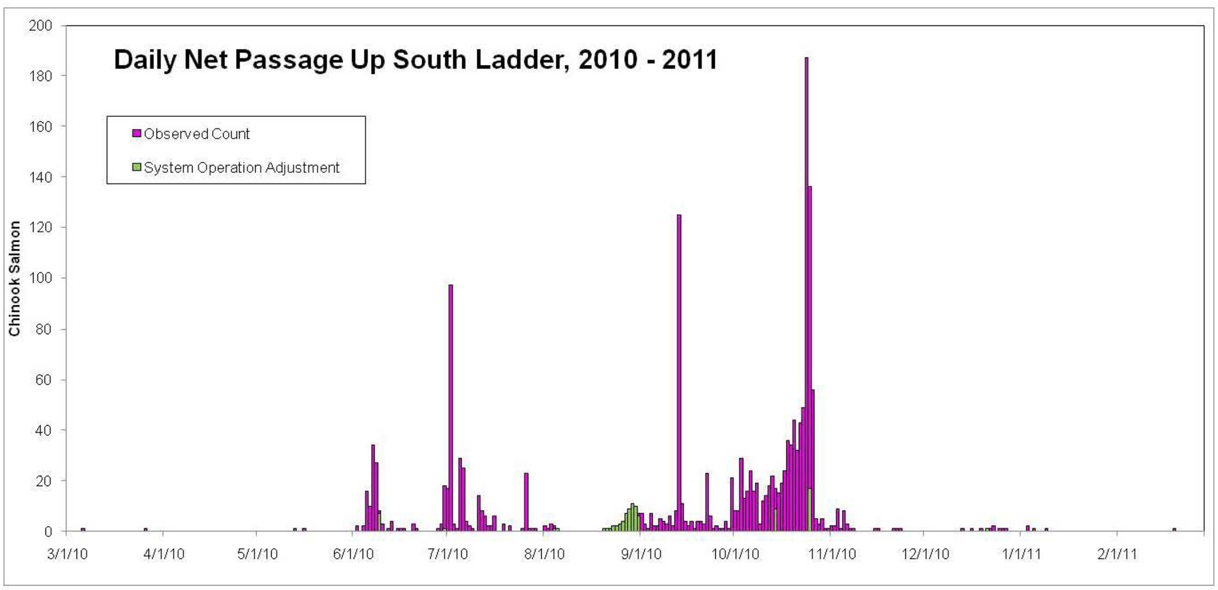
# Attachment 2



# Attachment 2



# ATTACHMENT B

RMT (2013) Methodology to Differentiate Spring-run and Fall-run Chinook Salmon Based Upon Corrected VAKI Riverwatcher$^{TM}$ Net Upstream Counts

# ATTACHMENT B

The following text is directly excerpted from RMT (2013, pp. 5-53 through 5-55) regarding the methodology previously used by the RMT to differentiate spring-run and fall-run Chinook salmon based on corrected VAKI Riverwatcher™ net upstream counts of Chinook salmon at DPD on the lower Yuba River.

---

The approach described in the following paragraphs represents the most objective way currently identified by the RMT to select an annually variable temporal demarcation date for the separation of spring-run and fall-run Chinook salmon passing upstream of Daguerre Point Dam for the eight years of available Vaki data.

The following conventions are used throughout the following paragraphs:

- The dates to which a particular daily count was assigned are referred as $d$, with $d$ ranging from 1 (March 1 of a given calendar year) to 365 or 366 (February 28 or February 29 of the next calendar year).
- The temporal demarcations (i.e., cutting dates) are referred to as $D$. Examination of Figures 5-34 and 5-35 indicates that the cutting dates should lie between June 15 and September 15. Further justification for allowing the range of potential cutting dates to extend to September 15 includes: (1) the previously reported temporal periodicity of spring-run Chinook salmon migration extending through September (Yoshiyama et al. 1998); (2) the observed phenotypic extended duration of holding prior to spawning; (3) NMFS (2011) suggestion that Central Valley spring-run Chinook salmon adult holding primarily occurs through mid-September; and (4) results from the acoustic tagging study, demonstrating that phenotypic spring-run Chinook salmon passed upstream of Daguerre Point Dam by the end of September, with the majority passing upstream of Daguerre Point Dam by approximately mid-September.
- Thus, $D$ ranges from 107 (June 15 of a given calendar year) to 199 (September 15 of the same calendar year).
- The daily Vaki observations are referred as $n_d$, with $N$ reserved for the summation of the daily observations over particular periods of time.
- The cumulative distributions of daily Vaki observations relative to the total number of observations over particular periods are referred as $Y$. For example, the cumulative distributions of daily lagged observations through day $d = X$ relative to the total number of Vaki observations over the year is expressed as $Y_X = \sum_{d=1}^{d \le X} n_d \Big/ N$.

The procedure to select annually variable temporal demarcations to separate spring-run and fall-run Chinook salmon in the eight annual time series of Chinook salmon daily passage upstream of Daguerre Point Dam consists of the following steps:

- Step 1. Select the first year y from the 8 available years of Chinook salmon Vaki data (*i.e.*, from 2004 through 2011.
- Step 2. Select the first temporal demarcation (i.e., a cutting day) $D$ from the range $D = 107$ through $D = 199$ (June 15 through September 15).
- Step 3. Separate the set of daily observations $n_d$ for year y into two subsets. One subset that includes all observations that occurred prior to day $D$ will describe potential spring-run Chinook salmon counts upstream of Daguerre Point Dam, and will sum to a total $N_S$. The other subset includes all observations that occurred from day $D$ through day 365 or 366 (leap year) will describe potential fall-run Chinook salmon counts upstream of Daguerre Point Dam, and will sum to a total $N_F$.
- Step 4. Calculate the daily relative cumulative distributions for the two data subsets of the previous step. The cumulative distribution of potential spring-run Chinook salmon daily observations to day $d = X$ will be calculated as $YS_X = \sum_{d=1}^{X<D} n_d \Big/ N_S$, while the cumulative distribution of potential fall-run Chinook salmon daily observations through day $d = X$ will be calculated as $YF_X = \sum_{d \geq D}^{X \leq 365} n_d \Big/ N_F$.
- Step 5. Fit a generalized logistic function (Richards 1959) to each of the two sets of daily relative cumulative distributions calculated in the previous step using a nonlinear minimum least squares procedure. The generalized logistic distribution was selected because it was desired that the smoothed function representing each group should be continuous, unimodal and plastic enough to allow for asymmetry.
- The expected cumulative distribution of potential spring-run Chinook salmon daily passage upstream of Daguerre Point Dam to day $d = X$ as represented by the corresponding fitted generalized logistic function has the formula: $\hat{Y}S_X = 1 \Big/ \left[1 + \exp\left(\alpha_S - \beta_S \times X\right)\right]^{(1/\delta_S)}$, and the expected cumulative distribution of potential fall-run Chinook salmon daily passage upstream of Daguerre Point Dam through day $d = X$ as represented by the corresponding fitted generalized logistic function has the formula: $\hat{Y}F_X = 1 \Big/ \left[1 + \exp\left(\alpha_F - \beta_F \times X\right)\right]^{(1/\delta_F)}$. The $\alpha_S$, $\alpha_F$, $\beta_S$, $\beta_F$, $\delta_S$ and $\delta_F$ are the fitted parameter values that describe the shapes of the resulting distribution functions. In particular, the parameter $\delta$, whose value was constrained to be greater than or equal to 0.1 and less than or equal to 10, determines the asymmetry of the resulting functions.
- Step 6. Using the fitted generalized logistic functions from the previous step, calculate the expected daily observations $\hat{n}_d$ with the following formula:
- $\hat{n}_d = N_S \times \left(\hat{Y}S_{d+1} - \hat{Y}S_d\right) + N_F \times \left(\hat{Y}F_{d+1} - \hat{Y}F_d\right)$.

- Step 7. Calculate and record the proportion of the annual daily variability explained by the expected daily observations $\hat{n}_d$ using the following formula:

$$\varphi_D^2 = 1 - \frac{\sum_{d=1}^{365}\left(n_d - \hat{n}_d\right)^2}{\sum_{d=1}^{365}\left(n_d - \bar{n}\right)^2}$$

- where $\bar{n}$ is the annual average of daily Chinook salmon observations.
- Step 8. Select the next temporal demarcation (i.e., a cutting date) D from the range *D* = 107 through *D* = 199, and repeat Steps 3 through 7.
- Step 9. Repeat Step 8 with each of the remaining cutting dates *D*.
- Step 10. Once an $\varphi_D^2$ has been obtained for each of the *D* from the range 107 through 199 corresponding to the selected year y, select the maximum $\varphi_D^2$ in the set. The selected temporal demarcation $D_{max}$ for year y is the cutting date *D* associated with the maximum $\varphi_D^2$ in the set.
- Step 11. Repeat Steps 2 through 10 with each of the remaining years.

------------------------------------------------------------------------------------------------------------