## Project summary

(1) Organization title: University of Alaska Fairbanks (UAF), Department of Fisheries, College of Fisheries and Ocean Sciences
(2) Principal Investigator(s) (PI): Peter Westley, associate professor of fisheries (UAF); (co-PI) Curry Cunningham, assistant professor of fisheries (UAF); (project partners) Heather Finkle, research biologist, Division of Commercial Fisheries, Alaska Department of Fish \& Game and Chris Habicht, principal geneticist, gene conservation laboratory, Alaska Department of Fish \& Game
(3) Address, telephone number, and email address of Principal Investigator: 2150 Koyukuk Drive, Fairbanks, Alaska, 99775, 907-474-7458, pwestley@alaska.edu; of (co-PI): 17101 Point Lena Loop Road, Juneau, Alaska, 99801, 907-796-5457, cjcunningham@alaska.edu; of (partners): Alaska Dept. of Fish \& Game, Kodiak Island Limnology Lab, 351 Research Court,Kodiak, Alaska, 99615, 907-486-1848, heather.finkle@alaska.gov; 333 Raspberry Road, Anchorage, Alaska, 99518, chris.habicht@alaska.gov, 907-267-2247
(4) Project title. Revealing the drivers of Chignik sockeye salmon population dynamics with a stage-based life cycle model
(5) Project objectives for the project period: (1) Synthesize two decades of historical sockeye salmon smolt outmigration data (1994-2016) to understand the influence of biotic and abiotic factors on the timing and size of seaward migration; (2) Conduct genetic stock identification on archived tissue samples from 3000 individuals from 2013-2016 and expand existing database of already genotyped individuals that will yield a decade of stock-specific migration (2006-2016) that encompasses outmigration years that resulted in fishery collapse, and (3) parameterize an existing Bayesian stage-structured life history model for Black Lake and Chignik Lake stocks to quantify the direction and magnitude of hypothesized factors influencing the population dynamics of Chignik sockeye salmon.
(6) Summary of work to be performed within the project period: Here, we propose a series of complementary objectives (above) that aim to illuminate and quantify the factors associated with the 2018 Chignik sockeye salmon fishery disaster. We will synthesize and analyze archived, but never explored, datasets and samples. By doing so, we aim to not only understand what happened in 2018, but to provide information to help inform future management options in the light of on-going rapid climate change, natural habitat evolution in the watershed, and increasingly variable ocean conditions.
(7) Budget Information - Total funds requested from PSMFC: $\mathbf{\$ 5 8 9 , 1 8 8}$; Current and pending support (see attached); Cost sharing to be provided to this project: Although no cost matching is required or will be tracked, in-kind support from the Alaska Department of Fish \& Game will greatly leverage the project budget by providing lodging and boats for UAF personnel while in Chignik during the 3 project field seasons; - Total project cost: $\mathbf{\$ 5 8 9 , 1 8 8}$.

## Project description

Focal topic and approach: The 2018 collapse of the Chignik sockeye salmon run closed the consistently lucrative Area L fishery and severely limited access of community residents to subsistence-caught fish, which threatens food security and cultural well-being (Westley et al. 2021). Key data gaps exist as to the potential cause of the collapse and the extent to which freshwater or marine processes were most at play, limiting our understanding of what would cause both runs to fail and impeding the development of management options. Said another way, we do not understand the factors that shape the ups and downs of salmon production in the Chignik watershed. Here, we propose a research project that will synthesize decades of data on stock-specific size and timing of seaward migrating sockeye salmon smolts to help parameterize an existing stage-based life cycle model that is well-suited to assess the weight of evidence of abiotic and biotic drivers on salmon survival or population productivity. By doing so, we aim to better understand the leading factors associated with the 2018 disaster and potentially forecast the impacts of future climatic conditions, or other identified scenarios of system-change, upon Chignik sockeye salmon.

This proposal is partitioned by objective that will form the basis of a PhD graduate student dissertation at the University of Alaska Fairbanks in the Department of Fisheries, in partnership with ADF\&G and local tribal organizations ${ }^{1}$

Objective (1) Synthesize two decades of historical sockeye salmon smolt outmigration data (1994-2016) to understand the influence of biotic and abiotic factors on the timing and size of seaward migration.

Although it is increasingly clear that mortality at sea of large, maturing salmon has demographic consequences (Manishin et al.), the timing and size of seaward migrating smolts, which in turn impacts surivial, remains most associated with return strength of adult salmon in subsequent years. Despite its importance, it remains unknown what factors shape the body size and timing of migration by Chignik sockeye salmon smolts. Here we analyze a longterm dataset on smolt migration available from partners at ADFG and apply methodology recently used to understand the factors associated with body size and timing of migration of Kodiak sockeye salmon (Ulaski et al. 2022). In figure 1, we show the observed average size of age 1 and age 2 sockeye salmon migrating to sea from 1994-2016. These data, along with the average day of year (julian day) of migration will be described using dynamic linear models (DLM). DLMs have been effectively used to identify abiotic and biotic drivers of growth and survival for salmon while accounting for autocorrelation. A major advantage of a DLM approach is that the model structure allows regression coefficients of predictor variables to vary through time. In contrast, more typical GLM or GAM approaches assume stationary (i.e., temporally static) relationships with predictor variables. Accounting for non-stationarity of regression parameters is important, as the processes

[^0]linking climate and population parameters may not be static (Litzow et al., 2018). The DLM structure can be described by two equations. First, the observation equation can be written as,
$y_{t}=\boldsymbol{F}_{t}{ }^{\mathrm{T}} \boldsymbol{\theta}_{t}+v_{t} \quad v_{t} \sim N(0, r)$
where smolt length $y_{t}$ at time $t$ is modeled as a function of $\boldsymbol{F}_{t}$, a column vector of potential predictor variables, $\boldsymbol{\theta}_{t}$ is a column vector of fitted regression parameters, and $v_{t}$ is a vector of normally distributed observation errors with mean 0 and variance $r$. The second, the process equation, allows regression parameters to vary over time through an autoregressive process via a Kalman filter and smoother (Holmes et al. 2012),
$\boldsymbol{\theta}_{t}=\boldsymbol{\theta}_{t-1}+\boldsymbol{w}_{t} \quad \boldsymbol{w}_{t} \sim M V N\left(0, \boldsymbol{Q}_{\boldsymbol{t}}\right)$
where $\boldsymbol{w}_{t}$ is a vector of process errors and $\boldsymbol{Q}_{\boldsymbol{t}}$ is the system variance matrix. The variance of $\boldsymbol{w}_{t}$ is stored in the diagonal with zeroes elsewhere. We will use a random-walk model as a means to evaluate the relative performance of more complicated models containing abiotic and biotic predictor variables such as water temperature, large scale climate forcing, and proxies for intraspecific competition. Alternative models for each population will be compared using the Akaike Information Criterion, bias-corrected for small sample sizes (AICc). The relative effects of each predictor on smolt length will be estimated as the product of the regression coefficient and the value of the independent variable in each year of the time series (Ulaski et al. 2022). Effect sizes will be expressed in the units of the response variable in order to provide a simple interpretation of the effect of independent variables on smolt length of South Olga lakes sockeye salmon. All models were fitted using the MARSS package (Holmes et al., 2012) in the R programming language (www.r-project.org).


Figure 1. Average length (mm) of seaward migrating Age $1(1+$, blue) and Age 2 (2+, orange) sockeye salmon sampled in ADFG rotary screw trap during migration years 1994-2016. Migration years that included individuals that (based on ocean age) may have returned in 2018 (grey shading in years 2014-2016. It is interesting to note that the smallest recorded size of 1+ migrants lines up temporarily with the collapse in 2018.

## Objective (2) Conduct genetic stock identification on archived tissue samples from 3000 individuals from 2013-2016 to fill in a dataset of stock-specific smolt size and timing.

To be consistent with other years, we will genotype approximately 250 individuals during three temporal strata (early, middle, and late, for a total of ca. 750 individuals per year) of the seaward migration window. Briefly, we will follow standard DNA extraction and genotyping protocols used in the ADF\&G Gene Conservation Laboratory. Genomic DNA will be extracted from tissue samples using a NucleoSpin® 96 Tissue Kit by Macherey-Nagel (Düren, Germany). Samples will be genotyped using Fluidigm ${ }^{\circledR} 192.24$ Dynamic Array ${ }^{\mathrm{TM}}$ Integrated Fluidic Circuits (IFCs), which systematically combined 24 assays and 192 samples into 4,608 parallel reactions. The components will be pressurized into the IFC using the IFC Controller RX (Fluidigm) and conducted in a 9 nL volume chamber consisting of a mixture of 20X Fast GT Sample Loading Reagent (Fluidigm), 2X TaqMan ${ }^{\circledR}$ GTXpress ${ }^{\text {TM }}$ Master Mix (Applied Biosystems ${ }^{\text {TM }}$ ), Custom TaqMan® SNP Genotyping Assay (Applied Biosystems), 2X Assay Loading Reagent (Fluidigm), 50X ROX Reference Dye (Invitrogen ${ }^{\mathrm{TM}}$ ), and 60-400ng/ $\mu \mathrm{I}$ DNA. The Dynamic Array IFCs were read on a Biomark ${ }^{\mathrm{TM}}$ or EP1 $1^{\mathrm{TM}}$ System (Fluidigm) after amplification and scored using Fluidigm SNP Genotyping Analysis software.

Following rigid QA/QC procedures developed by the Gene Conservation Laboratory, we will use a Bayesian approach for genetic mixed-stock analysis (MSA), the Pella-Masuda Model as implemented in the program BAYES (Pella and Masuda 2001). The Bayesian method of MSA estimates the stock composition within each mixture strata using four pieces of information: 1) a baseline of allele frequencies for each population potentially contributing to the mixture, 2 ) the grouping of populations into the reporting groups desired for MSA, 3) prior information about the stock proportions of the mixture, and 4) the genotypes of fish sampled from the fishery. The baseline of allele frequencies for the 24 SNPs and 18 sockeye salmon populations and the groups into which the populations are combined are described in Dann et al. (2012).

Upon completion of genotyping and stock assignments we will compile a dataset describing stock-specific smolt size and migration timing for the years 2006-2016. We will again use DLM approaches to explain variation in stock-specific size and migration timing. Specifically we will assess the direction and magnitude of effects that biotic and abiotic factors have on stock-specific migration timing and size at outmigration. The relationships between smolt growth or timing and environmental factors are unknown in the Chignik system, and very rare in any salmonid system often due to a lack of data for this important life stage. We expect one PhD chapter and peerreviewed publication resulting from this objective.

## Objective (3) Parameterize an existing Bayesian stage-structured life history model for

 Black Lake and Chignik Lake stocks to quantify the direction and magnitude of factors hypothesized to influence the population dynamics of Chignik sockeye salmon.We will use an existing stage-based life cycle model recently applied to populations of Chinook salmon in the Yukon River watershed (Cunningham et al. 2018). This model is useful in its flexibility and ability to track cohorts originating from multiple populations (i.e., Black Lake and Chignik Lake) throughout their lifecycles and could allow those populations to interact. This is vital as we know that the Chignik watershed acts as a connected metacommunity
(Westley et al. 2008, 2010). Briefly, the population dynamics model will track cohorts of sockeye salmon from specific brood years forward in time across sequential freshwater and marine life stages. The number of salmon surviving to the end of a given stage ( $s$ ) depends upon the abundance of individuals from brood year $y$, surviving the prior stage of population $p$ ( $N_{y, s-1, p}$ ), and the year, stage, and population-specific survival rate $S R_{y, s, p}$ :

$$
\begin{equation*}
N_{y, s, p}=N_{y, s-1, p} * S R_{y, s, p} \tag{1}
\end{equation*}
$$

Survival rate though each life stage (Figure 1) is represented by a Beverton-Holt transition function (Moussalli \& Hilborn, 1986).

The Beverton-Holt equation (Beverton \& Holt, 1957) provides a useful approximation for density-dependent survival of individuals and has been utilized in a wide range of life cycle models for salmon (Cunnigham et al. (2018)). The survival rate was calculated as:

$$
\begin{equation*}
S R_{y, s, p}=\frac{p_{y, s, p}}{1+\frac{p_{y, s, p} N_{y, s-1, p}}{k_{s, p}}} \tag{2}
\end{equation*}
$$

where the productivity parameter $\left(p_{y, s, p}\right)$ represents the maximum survival rate in the absence of density-dependent compensation, and the capacity parameter ( $k_{s, p}$ ) describes the maximum number of individuals that could potentially survive through the stage. Capacity could be assumed specific to a single Chignik stock or shared across populations.

Stage and population-specific productivity parameters ( $p_{y, s, p}$ ) are treated as time-varying functions of the environmental covariates being evaluated. Each productivity parameter is an inverse-logit function of a basal productivity $\beta_{s, p, 0}$, or the average survival rate at low density across years (in logit space), and the sum of covariate effects

$$
\begin{align*}
& p_{y, s, p}=\frac{1}{1+\exp \left(-\beta_{s, p, 0}-\sum_{c=1}^{N c_{s, p}} \beta_{p, c} X_{t, c}\right)}  \tag{3}\\
& t=y+\delta_{c}
\end{align*}
$$

where covariate $c$ values at time $t\left(X_{t, c}\right)$ were multiplied by coefficients $\left(\beta_{p, c}\right)$ describing their respective influence on stage and population-specific survival rates. The parameter $\delta_{c}$ defines the difference between the brood year $y$ and the calendar year $t$ in which the cohort interacts with each covariate, and $N c_{s, p}$ is the number of productivity covariates linked to each in each stage. Following Cunningham et al. (2018) we will formulate a set of biologically-based hypotheses reflecting various biotic and abiotic factors that may influence population dynamics. All covariates will be z-standardized to aid interpretation of effect sizes and explored for correlation prior to analyses (covariates with correlations beyond 0.6 will be removed).

To select the most parsimonious set of explanatory covariates and generate predictions that are robust to model structural uncertainty, we will use Bayesian model selection methods.
Specifically, we will employ random effects version of the indicator variable method described
by Kuo and Mallick (1998), where the covariate influences in equation $3\left(\beta_{p, c}\right)$ are the product of the estimated covariate effect $\left(\varepsilon_{p, c}\right)$ and an indicator variable $\left(\tau_{p, c}\right)$ :

$$
\begin{align*}
\beta_{p, c} & =\tau_{p, c} \varepsilon_{p, c} \\
\tau_{p, c} & \sim \operatorname{Bernouli}\left(v_{p}\right) \tag{15}
\end{align*}
$$

The indicator variable is Bernoulli distributed with a population-specific prior inclusion probability of $v_{p}$ that is estimated as a hyperparameter. For each sample from the posterior, $\tau_{p, c}$ takes a value of 0 or 1 , functionally including or excluding that covariate in each step. As such, the average value of $\tau_{p, c}$ across posterior samples represents the inclusion probability for each covariate for each population, or the weight of evidence from the data for its inclusion. For this analysis we will implement a random effects version of the Kuo and Mallick (1998) method by placing a hyperprior on $\gamma_{p}$, the standard deviation of the normal prior on covariate effects $\varepsilon_{p, c}$.

Given the flexibility of the model structure and its capacity to explore future scenarios of change, we will work with project partners to identify future outcomes that are of most interest. This may include incorporation of intercept harvest while at sea, or increased warming, or changes to spawner densities that may reflect different management strategies.

## Project communication, outreach, and performance measures

We will organize and host town hall style meetings to share project updates and results with Chignik residents and stakeholders during annual visits to the watershed by project personnel. In addition we will present findings at regional (e.g., AK Chapter of American Fisheries Society) and national meetings (e.g., Ecological Society of America) leveraging travel funds from other sources available to the PI. We will measure performance of this work primarily through the submission and acceptance of results in peer-reviewed scientific journals. To increase accessibility of the results, we have budgeted for Open Access costs in year 3 of the project.

## Literature cited

Beverton, Raymond JH, and Sidney J. Holt. On the dynamics of exploited fish populations. Vol. 11. Springer Science \& Business Media, 1957.

Creelman, Elisabeth K., Lorenz Hauser, Ryan K. Simmons, William D. Templin, and Lisa W. Seeb. "Temporal and geographic genetic divergence: characterizing sockeye salmon populations in the Chignik Watershed, Alaska, using single-nucleotide polymorphisms." Transactions of the American Fisheries Society 140, no. 3 (2011): 749762.

Cunningham, Curry J., Peter AH Westley, and Milo D. Adkison. "Signals of large scale climate drivers, hatchery enhancement, and marine factors in Yukon River Chinook salmon survival revealed with a Bayesian life history model." Global change biology 24, no. 9 (2018): 4399-4416.

Dann, Tyler H., James R. Jasper, Heather A. Hoyt, Heather Hildebrand, and Chris Habicht. "Western Alaska Salmon Stock Identification Program Technical Document 6: Selection of the 96 SNP Marker Set for Sockeye Salmon." Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J12-11, Anchorage (2012).

Kuo, Lynn, and Bani Mallick. "Variable selection for regression models." Sankhyā: The Indian Journal of Statistics, Series B (1998): 65-81.

Moussalli, Elie, and Ray Hilborn. "Optimal stock size and harvest rate in multistage life history models." Canadian Journal of Fisheries and Aquatic Sciences 43, no. 1 (1986): 135-141.

Holmes, Elizabeth E., Eric J. Ward, and Wills Kellie. "MARSS: multivariate autoregressive state-space models for analyzing time-series data." $R$ J. 4, no. 1 (2012): 11.

Litzow, Michael A., Lorenzo Ciannelli, Curry J. Cunningham, Bethany Johnson, and Patricia Puerta. "Nonstationary effects of ocean temperature on Pacific salmon productivity." Canadian Journal of Fisheries and Aquatic Sciences 76, no. 11 (2019): 1923-1928.

Pella, Jerome, and Michele Masuda. "Bayesian methods for analysis of stock mixtures from genetic characters." Fishery Bulletin 99, no. 1 (2001): 151-151.

Rich, Harry B., Thomas P. Quinn, Mark D. Scheuerell, and Daniel E. Schindler. "Climate and intraspecific competition control the growth and life history of juvenile sockeye salmon (Oncorhynchus nerka) in Iliamna Lake, Alaska." Canadian Journal of Fisheries and Aquatic Sciences 66, no. 2 (2009): 238-246.

Ulaski, Marta E., Heather Finkle, Anne H. Beaudreau, and Peter AH Westley. "Climate and conspecific density inform phenotypic forecasting of juvenile Pacific salmon body size." Freshwater Biology 67, no. 2 (2022): 404-415.

Westley, P. A. H., R. Hilborn, T. P. Quinn, G. T. Ruggerone, and D. E. Schindler. "Long-term changes in rearing habitat and downstream movement by juvenile sockeye salmon (Oncorhynchus nerka) in an interconnected Alaska lake system." Ecology of Freshwater Fish 17, no. 3 (2008): 443-454.

Westley, Peter AH, Daniel E. Schindler, Thomas P. Quinn, Gregory T. Ruggerone, and Ray Hilborn. "Natural habitat change, commercial fishing, climate, and dispersal interact to restructure an Alaskan fish metacommunity." Oecologia 163, no. 2 (2010): 471-484.

RC140



[^0]:    ${ }^{1}$ The PI spoke at length with Mr. George Anderson, president of Chignik Intertribal Coalition (CIC) to assess the most urgent needs of the tribes and the research reflected in this proposal reflects the outcome of that conversation. If awarded we would invite tribal participation on the student committee recognizing the need for Indigenous as well as western science to best understand the Chignik sockeye salmon collapse.

