Hello Mary and Zac

I wanted to give you an update regarding progress on the Sitka Salmon modeling project. I’ve spent the past year focused on how best to approach the issue of modeling and inferences related to estimating pink salmon in the Indian River and the role of hatchery strays. To that end, I’ve had a number of meetings with ADF&G seeking additional data and trying out different models and approaches that best suit these objectives. I’ve also had to focus on writing my dissertation/research proposal based on that trial-and-error approach. Last month, I presented that proposal formally to my committee and will be finalizing the draft, based on feedback, later this month.

To summarize, the research seeks to assess whether the abundances of pink salmon observed in recent decades at Indian River exist within a natural range of variation relative to pink salmon streams elsewhere in the region. The additional and related objective is to quantify the effect SJH operations has had on abundances of pink salmon at Indian River over time. My focus has been based on the assumption that harm to ecosystems can be driven by the hyperabundance of a native species due either to shifting regional trends in habitat suitability, to direct (and often anthropogenic) intervention, or to a combination of these factors.

As you know numbers of pink salmon have increased rapidly in the last 30 years, coincident with operations of the SJH, and there have been example when high fish densities can impact the Indian River’s ability to sustain other resident fish species by crashing in-stream dissolved oxygen concentrations.

As of now I seek to focus on 2 fundamental questions/objectives and will write at least 2 stand-alone manuscripts) that I hope will be of strong value to you. These are X and Y. My approach involve integrating data from Indian River and the wider region to make inferences about trends in pink salmon populations, and to consider the impact of SJH releases of pink salmon on Indian River populations using data from operations reports filed by the hatchery with ADFG. To these ends, I’m utilizing ADFG surveys of 700 pink salmon index streams in SEAK, including the 35 streams grouped near Sitka and the Indian River. I’m also utilizing data collected intermittently by ADFG of pink salmon escapement at Indian River. For estimating abundance and trends, I’ve decided on an approach (MARSS) that’s well suited for parsing time series data. Of course I don’t want to mislead by inferring that these results are the final results but preliminary estimates demonstrate IR pink salmon are higher than average but not outlying after controlling for stream length and observer. Model results demonstrate that SITK salmon follow the region-wide trend of increases in pink salmon, meaning that the Indian River would expect to see increases even if SJH operations didn’t exist. However, it also demonstrates that abundance is higher than most suggesting……ETC ETC.

I’ll be working to refine these model and build on these inferences….

I’m also using difference-in-difference analysis in an effort to isolate the causal impact of hatchery operations on Indian River pink salmon populations. This is a quais-experimental approach that examines treatment group and control group in pre- and post-treatment periods Indian River (treatment) and nearby index streams (control)

Will consider two treatments:

Hatchery operations begin in 1975

Operations expand in 2010

Preliminary results indicate a strong effect of time period and insignificant treatment effect

To understand impact of hatchery releases on Indian River pink salmon abundance I’ve decided on using MARSS models evaluate the impact of Indian River abundance and hatchery releases in year *t-1* on Indian River abundance in year *t (a brief description can be found below).*

Separate models will consider impact of hatchery releases (known) and impact of estimated hatchery returning adults

Most direct quantification of the impact of hatchery releases on Indian River abundances

The results are intended to provide you/the park management with a comprehensive description of both the impacts of Sheldon Jackson Hatchery operations and of the population dynamics of pink salmon throughout the broader region.

**Figure 1**

A graph showing the number of states

AI-generated content may be incorrect.*Peak pink salmon escapement observed annually from 1960 to 2023 in 35 index streams across ADFG’s designated Northern Southeast (Outside) region. Spawner counts are scaled against stream length (km), then log transformed and standardized at the level of each stream.*

A graph showing the growth of salmon

AI-generated content may be incorrect.**Figure 2**

*MARSS model results estimating the underlying states of even-year pink salmon abundance at Indian River (highlighted in pink) and throughout north southeast (outside) region (grey). Uncertainty is higher at Indian River due to fewer observations.*

**Timeline**

**August 2025: Finish and finalize dissertation proposal**

**March 2026: Finish draft manuscript focused on XXXX.**

**August 2026: Finish draft manuscript focused on YYYY**

**September 2026: Meeting at SITK to present preliminary results and seek feedback on how to enhance its utility for you. Invite managers at the Science Center and hatchery managers.**

Thank you again for the support for my project. I’m really happy I’ve now got a clear path for analyses……etc. Also, thanks for your feedback on the summary article I’ll be submitting to the next issue of the Alaska Park Science….I got really good feedback from the hatchery folks and from Zac (??)……..

**Models**

MARSS models are composed of a process (or state) model and an observation model, generically formulated as follows:

Observation equation: **yt** = **Zxt** + **a** + **Ddt**, + **vt**, where **vt**∼MVN(0,**R**)

Process/state equation: **xt** = **Bxt−1** + **u** + **Cct** + **wt**, where **wt**∼MVN(0,**Q**)

In the above, observed data in a given time period (**yt**) are used to estimate the underlying state of a system of interest (**xt**). In the observation equation, **Z** determines which time series observations in the observation equation will inform which underlying states in the process model. In the case of Indian River and the 35 index streams in its vicinity, specifying **Z** as a 36x36 identity matrix would allow each time series to be representative of a separate (in this case stream-level) state, while specifying **Z** as a 36x1 vector of 1s would force all time series in the data to represent samples from one comprehensive underlying state. **a** allows for trends in the observation process to be asserted or captured (depending on whether a is specified or estimated), and any correlation in observation error between time series is defined by **R**, the variance-covariance matrix. In the process equation, **B** defines the interaction between state estimates through time (i.e., **xt** and **xt−1**). If **B** is specified as an identity matrix, **u** then represents trends in the underlying state(s), otherwise **B** and **u** together determine the underlying mean and how quickly the time series would return to the mean following some perturbation. Finally, the correlation of process error is defined by **Q**. Finally, **ct** and **dt** are covariate data related to the state or sampling procedure respectively, with **C** and **D** capturing the effect of said covariates.

Difference-in-difference models compare a treatment group to a control group across two time periods, pre- and post- intervention. The general structure of a difference-in-difference model is as follows:

yi,t = β0 + β1Dtreatment + β2Dpost  + β3Dpost\*Dtreatment + ei,t

In the above, i indexes whether an observation (yi,t) relates to the treatment or control group, while t indexes whether that same observation comes from before or after the treatment is applied. Dtreatment is equal to 1 when yi,t is part of the treatment group and equal to 0 otherwise. Similarly Dpost is equal to 1 when yi,t is observed after treatment is applied, and equal to 0 when yi,t is observed before. These indicator variables allow for the estimation of a net effect of treatment on the treatment group (β3) by first controlling for differing characteristics in the treatment and control group (β1) and differences in the pre- and post-treatment time periods (β2). Additional control covariates may be included in order to further isolate the effects of interest.