Chum Genetic and Environmental Management (ChumGEM) Model, Phase 1: Post-season run reconstruction

Final project report to the Pacific Salmon Commission, Chum Technical Committee

July 14, 2015

SC Fisheries, Ltd
4255 Cliffmont Road
North Vancouver, B.C., Canada, V7G 1J3

Contents

List of Tables

List of Figures

1 Introduction

Managing gauntlet fisheries for migrating Pacific salmon requires exposing an unknown abundance of co-migrating salmon stocks to a series of fisheries each potentially capable of imposing high mortality on very short time scales. To avoid mistakes that may cause serious conservation and economic harm, the salmon manager needs to understand the timing of stock arrivals to fishing areas, as well as the expected exploitation rates that stocks may experience as they migrate towards spawning grounds. This project developed a new database, run-reconstruction model, and visualization software (*ChumGEM*) to assist in post-season run reconstruction analysis for chum salmon fisheries. The run reconstruction analysis provides post-season estimates of abundance, arrival timing, exploitation rates, and escapement by day of season, management area, and genetic unit (GU) for co-migrating chum salmon stocks in southern British Columbia and Puget Sound Washington. The *ChumGEM* software package is orgranized into a set of 5 applications, or apps, that allow users to visualize the available monitoring and catch data for all chum salmon GUs for years in 2008-2013, configure and execute a Bayesian run-reconstruction model, and view results from alternative model configurations. The Bayesian approach to run reconstruction modeling allows the user integrate current knowledge about expected salmon abundance (i.e., forecast run size), migration timing and speed, and sampling variability with in-season fisheries catch and monitoring data to provide an objective post-season assessment of GU-specific run size, timing, and exploitation by management area. Although formal run reconstruction analyses were beyond the scope of this project, we also provide initial run reconstruction model analyses for each year to familiarize new users with the capabilities and typical output from *ChumGEM.*

This report provides the technical specifications of the Bayesian run reconstruction model, as well as a detailed step-by-step guide to completing one run reconstruction analysis for the 2012 fishing year. This model, which we refer to as CGEM, exists as an executable program developed using AD Model Builder software for non-linear optimisation. The larger *ChumGEM* software package forms a wrapper for managing the inputs and outputs from CGEM, as well as providing an environment and database for working with chum salmon data. A user manual for the *ChumGEM* software package is provided in a separate document. Financial reporting and the project summary are found in Appendices 1 and 2, respectively.

2 Model overview

The ChumGEM post-season run reconstruction model attempts to estimate stock-specific returning abundances, catch, and exploitation rates of chum salmon returning to spawning grounds in southern British Columbia, Canada and Puget Sound, Washington. Model spatial dimensions cover all ocean and terminal fisheries operating between Calvert Island (at the northern tip of Vancouver Island) and the southern end of Puget Sound, WA, including both sides of Vancouver Island. The model temporal dimension covers a 250-day fishing season (1 July to 7 March) on a daily time step. We define chum salmon stocks according to their Genetic Unit (GU) and general run timing group (i.e., Summer, Fall, or Winter), leading to notation such as PSS_Fall to denote chum salmon returning to the Puget Sound South GU in the Fall time period (Table 1). Within each putative stock, we further define migratory groups according to the number of possible migration paths between the ocean entry area (Calvert Island) and stock-specific spawning grounds. For instance, stocks within PSS_Fall can approach spawning grounds via an inside route through the Salish Sea to Puget Sound or via an outside route along the west coast of Vancouver Island and through Juan de Fuca Strait and Puget Sound. This scenario requires two migratory components PSS_Fall(1) and PSS_Fall(2), where a single PSS_Fall diversion rate parameter assigns the proportional abundances to each group. Figures 1 & 2 show all migratory paths for both Canadian and Puget Sound chum stocks included in the model.

Estimating the total return abundances and exploitation rates in each fishery management area (FMA) requires reconstructing the daily abundances of each stock within each FMA. Daily abundances are computed by the model based on (1) GU-specific total return abundance (we sometimes refer to this as run size) prior to all fisheries, (2) an arrival timing distribution to Calvert Island, (3) movement rates (km/day) along all migratory paths, and (4) stock-specific daily removals by fisheries within each FMA. All of these variables are unknown for most stocks and FMAs; therefore, CGEM must be either estimate them or take fixed input values based on other data. In its current form, the CGEM model uses movement rates (3) estimated from tagging, while estimating all other variables (1,2,4) by fitting the CGEM model predictions to empirical data from test fisheries, genetic stock identification (GSI) samples, and spawning ground escapement estimates. The model fitting procedure uses a complex non-linear parameter estimation scheme, which can be summarized as follows:

1. Propose initial guesses for (i) run sizes, (ii) run timing, and (iii) diversion rate parameters for each stock;
2. Apportion run sizes to stock-specific migratory groups based on diversion rates and predetermined migration paths;
3. For each time step, move migratory groups through FMAs, removing stock-specific catches in proportion to their among-stock relative abundances;
4. Generate predicted test fishery CPUE and spawning ground escapements for all FMAs and time steps in which they occur;
5. Generate predicted GSI stock proportions for all FMA and time steps where samples were taken;
6. Compute a goodness-of-fit statistic between model predictions and empirical data;
7. Revise parameter guesses in the direction of better goodness-of-fit;
8. Repeat 2-7 until parameter estimates no longer improve (i.e., estimates are converged).

The following sections provide detailed mathematical specifications for each of these steps.

3 Model notation

To differentiate Canadian and US management areas, Canadian area names are preceded by a "C", while US area names are preceeded by a "U" (Table 3). Fishery management areas within the model are indexed by a. For each area we specify l_a as the number of kilometers a fish must swim to traverse the area from one area boundary to another. There is some ambiguity to this definition because the "entrance" boundary and "exit" boundary of a given area may differ across stocks. Additionally, stocks may not traverse areas in a straight line. Area lengths were calculated using Google Earth version 7.1.2.2041.

Chum that have traversed all fishery management areas within their migration path subsequently enter their spawning grounds. While all base stocks have unique spawning areas, we use a common A to index all spawning ground areas. Migratory components within each base stock are indexed by x. Each migratory stock has a unique migration path $\boldsymbol{m}_{sx}$ expressed as a sequence of areas beginning in area C10 (Calvert Island) and ending in the generic spawning ground area A (Table 2).

4 Model Description

The run reconstruction model has five primary components: (1) *Parameters*, Θ, consisting of the stock-specific run sizes entering area C10 (R_s), mean dates of arrival to C10 (μ_s), variability around the mean dates of arrival to C10 (σ_s), and the proportions of stock abundance apportioned to each migratory component κ_{sx}; (2) *Data*, $\boldsymbol{Y}$, comprised of area-/time-specific test fishery catch-per-unit-effort (CPUE), GSI-based estimates of proportional stock composition, total catch, and escapement to spawning areas; (3) *Dynamic model* equations or rules for representing the changes in stock abundances over time and space. The dynamic model generates a set of predictions $\hat{\boldsymbol{Y}}$ that should agree reasonably well with the observed data $\boldsymbol{Y}$; (4) *Prior distributions* $\pi_i(\cdot)$ summarizing what is known about each of the parameters in Θ; and (5) a *Likelihood function*, $L(\Theta|\boldsymbol{Y}) = p(\boldsymbol{Y}|\Theta)$ describing the probability of observing the data given the parameters and dynamic model.

A four dimensional abundance array $\boldsymbol{N}$ tracks the number of fish from component x of stock s in area a on day t. We initialize the model by calculating an arrival distribution at area C10 (N_{sx1t}) using the parameters Θ. We then "fill in" the rest of the abundance array using the dynamic model.

4.1 Arrival distribution

We assume that all components belonging to a particular stock have identical arrival timing to area C10. The size of each migratory component may differ via the diversion rate parameters γ_s, which we define as the proportion of the total run size of stock s that migrate around Vancouver Island via the northern diversion route (Johnstone Strait).

We define $X_{s,\text{north}}$ to be the set of components that migrate via the northern diversion route and we define $n_{s,\text{north}}$ to be the number of components in $X_{s,\text{north}}$. For each migratory component x in $X_{s,\text{north}}$, we calculate the proportion of the total run size of stock s attributable to component x (κ_{sx}) by dividing the northern diversion rate by the number of components taking the northern diversion route, i.e., $\kappa_{sx} = \gamma_s n_{s,\text{north}}^{-1}$. Similarly, for components taking the southern diversion route ($x \notin X_{s,\text{north}}$) we calculate the proportion of the total run size of stock s attributable to component x by dividing the southern diversion rate by the number of components taking the southern diversion route, i.e., $\kappa_{sx} = (1 - \gamma_s)(X_s - n_{s,\text{north}})^{-1}$

The diversion parameter γ_s will generally not be estimable from the current data available for chum salmon fisheries; therefore, it must be pre-specified. We calculated the movement rate ρ to be 37.7 km/day from tagging studies in areas C12 and C13 from 2000-2002.

The daily arrival proportions for stock s to C10 are assumed to arise from a normal distribution with mean μ_s and standard deviation σ_s, i.e.,

$$A_{st} = \frac{\exp\left(-\frac{(t-\mu_s)^2}{2\sigma_s^2}\right)}{\sum_i \exp\left(-\frac{(i-\mu_s)^2}{2\sigma_s^2}\right)} \tag{1}$$

Daily arrivals for component x of stock s are then generated by multiplying the daily arrival proportions by the stock-specific run size R_s and the proportion of that run size attributable to component x, i.e., $N_{sx1t} = \kappa_{sx} R_s A_{st}$.

4.2 Harvest

Total aggregate catches (i.e., input data) are apportioned to stocks in proportion to stock-specific abundances. We first estimate time- and area-specific harvest rates by dividing the observed catch by the predicted total number of fish available in the area, i.e.,

$$U'_{at} = \frac{C_{at}}{\sum_s \sum_x N_{sxat}} \tag{2}$$

Harvest rates are constrained to be less than 0.95 via

$$U_{at} = \begin{cases} U'_{at} & U'_{at} \leq 0.95 \\ 1 - \frac{1}{400(U'_{at} - 0.9)} & U'_{at} > 0.95 \end{cases} \tag{3}$$

After applying constraints, stock-specific catch is given by

$$C_{sat} = U_{at} \sum_x N_{sxat} \tag{4}$$

4.3 Movement

Movement of chum at each time step is controlled by an array $\boldsymbol{P}$ that defines the proportion of chum from component x of stock s that will migrate from area i to area a in one day. The number of chum moving into area a at each time step is computed from these proportions and the number of chum escaping fisheries on the previous day, i.e.,

$$N_{s,x,a,t+1} = \sum_i P_{sxia} N_{sxit}(1 - U_{it}) \tag{5}$$

Movement proportions are calculated from an ordered multinomial logit model (Gelman et al. 2007, pg 119). The basic idea of the movement model is that, given (i) a mean migration rate (ρ), (ii) a "departure" point ($\boldsymbol{D'}$) representing the source area, (iii) a sequence of "cutpoints" corresponding to the boundaries of fishing areas ($\boldsymbol{D}$), and (iv) a spread parameter (σ_P), we calculate the proportion of chum leaving the departure point that will land between each consecutive pair of cutpoints after a single day of migration.

The departure point and the cutpoints are defined based on their distances from area C10. We define Δ_{sxk} to be the distance between area C10 and the kth area in the migration path of component x, i.e.,

$$\Delta_{sxk} = \sum_{i=1}^{k} l_{m_{sxi}} \qquad 1 \leq k \leq \overline{m_{sx}} \tag{6}$$

where $\boldsymbol{l}$ is a vector of area lengths and $\overline{m_{sx}}$ is the total number of areas in the migration path of component x of stock s. The departure point for chum residing in the jth area of their migration path is defined as the distance between C10 and the jth area, i.e.,

$$D'_{sxj} = \Delta_{sxj} \tag{7}$$

while the cutpoints are defined as the boundaries of fishing areas along the migration path, i.e.,

$$D_{sxk} = \begin{cases} 0 & k = 1 \\ \Delta_{sxk} + 0.5 & 2 \leq k \leq \overline{m_{sx}} \end{cases} \tag{8}$$

The mean or expected landing area is the sum of the mean migration rate (km/day) and the departure point, i.e.,

$$\Psi_{sxj} = \rho + D'_{sxj} \tag{9}$$

Assuming $m_{sxj} = i$, the probability of a fish in area i moving at least as far as area k in a single time step is then given by:

$$Q_{sxik} = \frac{\exp\left(\frac{\Psi_{sxj} - D_{sxk}}{\sigma_P}\right)}{\exp\left(\frac{\Psi_{sxj} - D_{sxk}}{\sigma_P}\right) + 1} \tag{10}$$

Thus, the probability of moving to the kth area is given by

$$P'_{sxik} = \begin{cases} 1 - Q_{sxi1} & k = 1 \\ Q_{s,x,i,k-1} - Q_{sxik} & 2 \leq k \leq \overline{m_{sx}} \end{cases} \tag{11}$$

The final migration proportions are then be calculated by summing over all k which represent the same area, i.e.,

$$P_{sxia} = \sum_k P'_{sxik} \qquad \{\forall k | m_{sxk} = a\} \tag{12}$$

4.4 Escapement

Chum that reach the spawning grounds are considered “escaped” from fisheries and will no longer be present in the model at the next time step. Daily escapements are summed over migratory components within a stock according to

$$E_{st} = \sum_x N_{sxAt} \tag{13}$$

where A is the generic spawning ground area shared by all stocks. Any fish left in the model on day T are added to the escapement on day T:

$$E_{sT} = \sum_x \sum_a N_{sxat} \tag{14}$$

4.5 Data likelihoods

Parameters to be estimated for each stock include run size (R_s), the mean date of arrival to area C10 (μ_s) and the standard deviation around the mean date of arrival to area C10 (σ_s).

4.5.1 Total area-specific catch

Constraining harvest rates to less than 0.95 causes estimated area-specific catches to differ from observed catches. We attempt to minimize these occurrences by penalizing differences between the model and observed catch for each area and time step via

$$C_{at} = \sum_s C_{sat} \tag{15}$$

and

$$-\ln L_1 = \sum_a \sum_t \left(\ln \widehat{C_{at}} - \ln C_{at}\right)^2 \tag{16}$$

4.5.2 Contribution of observed escapement

Escapement to each spawning area is only available as annual totals. Additionally, the observed escapement represents only a fraction of the true escapement, since not all spawning areas are monitored. We define this fraction to be the escapement bias (E_s^*). Therefore, we compute the likelihood by summing model escapements over the whole season and multiplying by the escapement bias, i.e.,

$$E_s = E_s^* \sum_t E_{st} \tag{17}$$

and include this total in the likelihood

$$-\ln L_2 = \sum_s \left(\ln \widehat{E_s} - \ln E_s \right)^2 \tag{18}$$

4.5.3 Contribution of test fishery samples

Where they occur, test fishery CPUE provides critical information on the relative abundance and timing of chum salmon movement through each area. We assume that test fishery CPUE is directly proportional to the abundance of chum in an area, i.e.,

$$I_{at} = q_a \sum_s \sum_x N_{sxat} \tag{19}$$

We calculate daily log-catchability residuals by dividing CPUE by abundance and taking the natural logarithm:

$$\boldsymbol{z}_a = \ln \left(\frac{\widehat{I_{at}}}{\sum_s \sum_x N_{sxat}} \right) \tag{20}$$

Computing the mean of these residuals gives us the log-catchability by area

$$q'_a = \frac{\sum_{j=1}^{J_a} z_{aj}}{J_a} \tag{21}$$

where J_a is the number of CPUE observations in area a. Test fishery catchability by area is then given by

$$q_a = \exp(q'_a) \tag{22}$$

The likelihood contribution of test fishery samples is calculated by summing the squared deviations between log-catchability and the daily log-catchability residuals:

$$-\ln L_3 = \sum_a \sum \left(\boldsymbol{z}_a - q'_a \right)^2 \tag{23}$$

4.5.4 Contribution of GSI samples

To fit the model to GSI data, we used a method for comparing proportions presented by Fournier et al. (1998) The idea is to reduce the influence of (i) stocks that are not observed and (ii) large sample sizes. We calculate the estimated proportion of fish from stock s in sample g as

$$W_{gs} = \frac{\sum_x \sum_{t=t_{1g}}^{t_{2g}} N_{sxa_g t}}{\sum_s \sum_x \sum_{t=t_{1g}}^{t_{2g}} N_{sxa_g t}} \tag{24}$$

We set

$$\xi_{gs} = (1 - W_{gs})\, W_{gs} \tag{25}$$

and

$$\tau_g = \frac{1}{\min(\zeta_g, \zeta^*)} \tag{26}$$

where ζ^* is an "upper limit" sample size, above which samples are assumed not to be more accurate. We then assume the variance of $\widehat{W_{gs}}$ is

$$\upsilon_{gs} = \left(\xi_{gs} + \frac{0.1}{S}\right) \tau_g^2 \tag{27}$$

The log-likelihood function for the GSI data is then given by:

$$-\ln L_4 = -0.5 \sum_s \sum_g \ln\left(2\pi\left(\xi_{gs} + \frac{0.1}{S}\right)\right) + \sum_s \sum_g \ln\left(\exp\left(-\frac{1}{2\upsilon_{sg}}\left(\widehat{W_{gs}} - W_{gs}\right)^2\right) + 0.01\right) \tag{28}$$

4.5.5 Contribution of prior distributions

We assume that the prior distribution for the run size of stock s is normally distrbuted with mean R_s^* and standard deviation σ_{Rs}:

$$-\ln L_5 = \sum_s \frac{(R_s - R_s^*)^2}{2\sigma_{Rs}^2} \tag{29}$$

Similarly, we assume that the prior distribution for the mean date of arrival to C10 for stock s is normally distrbuted with mean μ_s^* and standard deviation μ_{Rs}:

$$-\ln L_6 = \sum_s \frac{(\mu_s - \mu_s^*)^2}{2\sigma_{\mu s}^2} \tag{30}$$

Standard deviations around the mean date of arrival to C10 are assumed to arise from the inverse-gamma distribution:

$$-\ln L_7 = (\alpha_1 + 1)\ln(\sigma_s) - \frac{\alpha_2}{\sigma_s} \tag{31}$$

4.5.6 Total likelihood

The total objective function is calculated by summing over all likelihood components:

$$-\ln L = \sum_{i=1}^{7} L_i \tag{32}$$

5 Example Analysis

This section provides a step-by-step guide to completing a typical post-season run reconstruction via the ChumGEM software package.

5.1 Step 1: Initialize ChumGem

Launch the R console and set the working directory to the ChumGem folder. Load the necessary files by entering the command `source(``chumGem.r'')` at the R command-line.

5.2 Step 2: Visualize the chum salmon migration and fishery conceptual model

Launch the *chumHotel* GUI by entering the command `chumHotel()` at the R command-line. The default behaviour of *chumHotel* is to plot migration paths (Figures 1 & 2). Examining the migration paths allows the user to identify areas in which intermingling may occur between stocks.

5.3 Step 3: Visualize catch and escapement data in chumData

Launch the *chumData* GUI by entering the command `chumData()` at the R command-line. To visualize barplots of the data, click on the "Bars" radio button. To view the escapement by stock for 2012, set the response variable to "Escapement" and set the independent variable to "Stock". Then select "Year" as the conditioning variable and choose 2012 from the year droplist. The resulting plot (Figure 3) shows that spawning ground escapement counts in 2012 are dominated by Fraser River Fall and Straight of Georgia Fall stocks.

To investigate the magnitude and location of catches in 2012, select "Catch" as the response variable and "Area" as the independent variable, leaving the conditioning variable and year unchanged. To partition the data by gear type instead, select "Area*Gear" as the independent variable. From the resulting plot (Figure 4), catches in 2012 exceeded 100,000 fish in eight management areas: C12, C13, C18, U10, U11, U12, U12C, and U12H. Referring to the migration path plots, we would expect fishing pressure to be high on Hood Canal chum, because they are the only stock migrating through U12 and its subareas. We would also expect fishing pressure to be high on Puget Sound Central and Puget Sound South chum because they are the only stocks migrating through U10 and U11. Inferences about Canadian stocks are more complicated because catches are concentrated in areas where many stocks co-migrate.

The different coloured bars in Figure 4 represent the different gear types in each fishery. The plot reveals the prevalence of gillnetting in Canadian fisheries, as well as significant seining in Johnstone Strait and Strait of Georgia. Catches in US waters do not have associated gear type information at this time because data were provided with multiple codes and no conversion table.

5.4 Step 4: Specify prior parameters and initial conditions

Clicking the <Est>button in the chumData GUI or entering chumEst() in the R command-line launches the chumEst GUI, where the user can create and execute a new run reconstruction model.

Three parameters estimated for each stock are: (1) run size, (2) mean date of arrival to Area C10, and (3) the standard deviation around the mean date of arrival to C10. Stock-specific prior distributions are specified for run size and mean data of arrival to C10, while a common inverse-gamma prior distribution on the standard deviation is shared across all stocks.

In the Stock Parameters section, set the initial conditions for the stock-specific priors. First select a stock from the droplist, then enter values for the prior parameters in the input boxes below. The stock-specific priors for run size and mean date of arrival to area C10 are assumed to be normally distributed. Specify the mean and standard deviation parameters in the boxes for each prior corresponding to the "Prior Mean" (somewhere between days 70 and 120) and

“Prior SD” (Within how many days do fish typically arrive around the Prior Mean?). The unit for arrival parameters is “model days”. Day 1 (July 1 by default) is specified by the Start Day and Start Month boxes in the Control Bounds sections.

We specify an inverse-gamma prior on the standard deviation around the mean date of arrival of each stock to C10. The shape parameter and scale parameter of this distribution are specified in the “Arr. Shape” box and “Arr. Scale” box in the Control Bounds section. To see the shape of the inverse-gamma prior distribution on the standard deviation around the mean date of arrival, make a barplot by typing the following command in the R console:

```
>barplot( dinvgamma( 1:20, shape=50, scale=100 ) )
```

Alter the shape and scale parameters until the distribution is something you like (e.g., has a peak corresponding to the most likely value value and spread reflecting the current uncertainty). Note that the shape parameter is currently bounded to be <50, so you might need to modify the scale as well.

Reducing prior distribution standard deviations usually leads to better/faster reconstruction model convergence since the user is imposing more information on the model in addition to the raw test fishery, catch, GSI, and escapement data. Lack of convergence most often occurs because the raw data simply are not informative enough to estimate parameters for some stocks. Therefore, some stocks may require very informative prior distributions until new information is generated via monitoring upgrades.

5.5 Step 5: Visualize parameters and data for a specific year in chumEst

Visualizations in the chumEst GUI are controlled by a row of radio buttons beneath the Stock Parameters section. Initial conditions and prior parameters can be visualized using the “Pars” and “Priors” radio buttons, respectively. It is helpful to check these using a plot to make sure all stock information has been updated prior to launching the CGEM estimator.

Clicking the Catch radio button produces a plot of daily catches by area. For example, Figure 5 shows that catches mainly occur in October and November, with Canadian catches more prominent in October and US catches more prominent in November.

Clicking the Test radio button produces a plot of daily test fishery CPUE by area (Figure 6). The CPUE time series for area C29F appears to be most informative, showing a symmetric pattern that peaks in mid-October. The next area with the most CPUE observations is area C12, but the pattern displayed by this time series is much more erratic. CPUE observations in the remaining areas are sparse.

Clicking the GSI radio button produces a plot of GSI proportions by sample (Figure 7). All samples are generally dominated by Fraser River and Straight of Georgia stocks, though some samples in area U7 and U7A show a stronger presence of Puget Sound stocks.

5.6 Step 6: Run estimation and visualize results

Once the run reconstruction model is setup, run it by clicking the green <Run>button on the chumEst GUI. Once the estimation has completed (∼5 minutes), the ChumView GUI will launch provided that the model actually converged. If the model failed to converge, outputs will still be saved in the chumProject folder, but they will not be viewable in chumView (mainly because non-converged models should be revised until they do converged prior to making inferences). When models fail to converge, chumView will not open automatically, so the user should return to chumEst via the `chumEst()` command.

For converged models, check the parameter estimates and standard errors in chumView, by clicking the Par or Std radio button. The former plots parameter estimates on either logarithmic (base e) or natural scales, while latter does the same except includes the parameter standard errors. Clicking the exp(Std) button plots parameter estimates and standard errors on their original scale (Figure 8).

Investigate the run reconstruction model fit to the observed CPUE, GSI, and Catch data. To compare the estimated catch by area with the observed catch, first click the Bars radio button, then select Obs_Catch as the response variable and Area as the independent variable. For this example, the model routinely underestimated catches in areas with large catches (Figure 9). This could be the result of setting prior run sizes or arrival timing that is inconsistent with observed catch, as well as not having data to suggest anything different.

To examine model fits to escapement data, select Escapement as the response variable and Stock as the independent variable. The reconstruction fits to the observed escapement data fairly well, though the escapement of some stocks (i.e. Fraser River) is overestimated (Figure 10)

Examine model fits to the test fishery data: select CPUE as the response variable, day as the independent variable and area as the conditioning variable. By default, the y-axis is identical on each panel. This plot is easier to interpret if we allow the y-axes to vary, since magnitudes of abundance vary greatly from one area to another. We can allow the y-axes to vary by clicking the Free XY checkbox. Note that the model does well to capture the overall pattern of the CPUE time series, though abundances are again frequently underestimated (Figure 11).

6 ChumGEM run reconstruction model limitations

Models are imperfect representations of reality and ChumGEM (Phase 1) is clearly no exception. There is always room to improve input data, model formulations, computer implementation, solution methods, visualizations, etc. A significant limitation for chum salmon run reconstruction is the need for a large and detailed run reconstruction model capable of tracking different GUs and their multiple migration paths through fisheries to spawning grounds. Other run reconstruction models for ocean fisheries, such those used for Fraser River sockeye, do not require this added complexity because there are only two possible migration paths between northern Vancouver Island and the Fraser River. For chum salmon (and perhaps other species such as Chinook and Coho), stocks exit ocean fisheries almost continuously from northern Vancouver Island to spawning grounds along inside and outside migration routes, as well as through Puget Sound.

Data limitation is probably the most important current gap for ChumGEM run reconstruction modeling. Despite the high degree of model complexity, our preliminary run reconstruction analyses suggest that the ChumGEM reconstruction model could do a reasonable job in estimating arrival timing, abundance, and escapement for stocks that are most prevalent in the test fishery CPUE and GSI data (i.e., Fraser River, Strait of Georgia West, and Puget Sound North). Although there are 7 test fishery locations, only two of these provide potentially useful CPUE indices of abundance and timing over the full course of the migration and fishing seasons. The remaining areas provide too few CPUE observations (e.g., as few as 1 per year) to be informative. The GSI data is critical to providing information about stock-specific movements through fishing areas; however, in this case, too few areas are monitored with Area C12 (Johnstone Strait) dominating the GSI sampling. Low prevalence of Puget Sound stocks in Area C12 suggests that either (a.) these stocks utilize the outside migration paths where test fishing and GSI sampling doesn't exist or (b.) that GSI sampling does not have sufficient sample size and/or precision to detect low abundance stocks. Therefore, run reconstruction model estimates will be highly dependent on prior information about arrival timing, migration rates, run size, and the scale of unmonitored escapements for most of these stocks.

7 Recommendations for future work

The ChumGEM (Phase 1) software package provides a single environment in which to integrate multiple sources of data and analyses of chum salmon fisheries. Within the ChumGEM environment, users can (i) store catch, escapement, test fishery, and genetic stock identification data; (ii) visualize all of this information in various combinations of genetic units, management units, and day-of-season; and (iii) apply a state-of-the-art run reconstruction model to estimate patterns of chum salmon arrival and abundance through management areas to spawning grounds. This continuous linkage — from raw data to run reconstruction modeling — represents substantial progress toward the first (Post-season Run Reconstruction) of three modeling outputs specified in the long-term ChumGEM vision and strategic plan.

As noted above, data limitation represents the main obstacle to run reconstruction analyses for chum salmon fisheries. Therefore, we recommend that future research focus on identifying the most influential data improvements possible given current resources available for chum fishery and escapement monitoring. Improving the data is most important at this stage because run reconstruction models form the foundation for future annual and long-term fishery planning.

The ChumGEM run reconstruction model could be used within a simulation framework to identify the most valuable improvements to data collection for chum salmon fisheries and stocks (Figure 12). The basic approach would be to first complete formal run reconstruction analyses for all years that have sufficient data in the ChumGEM database. Run reconstruction outputs including stock-specific abundances, arrival timing, and fishery exploitation rates could then be used to parameterize a simulation model of in-season abundance and migration of chum salmon stocks, fishery exploitation rates by management area, and data generation via test fisheries, GSI, and escapement monitoring. The data generation steps could be simulated using alternative sampling designs ranging from, say, status quo to a full dataset in which every management area generates catch, GSI, and test fishery indices of abundance, and every spawning ground escapement is fully enumerated. A feasible sampling design should exist somewhere along this gradient from data-limited to data-rich: the purpose of the simulation exercise is to reveal the trade-off between investment in data collection and run reconstruction assessment performance.

A comprehensive run reconstruction assessment for all years in 2008-2013 would be the most logical next step toward the simulation framework described in Figure 12 and in further developing the ChumGEM long-term vision. Such an exercise would clearly identify what existing data can tell us about GU-specific exploitation rates in chum salmon fisheries. The current ChumGEM package is well-suited to such analyses, so the timeframe for completing the exercise should be on the order of a few to several months depending on analyst experience.

Constructing a simulation framework for evaluating alternative monitoring designs would be a significant undertaking, probably requiring 1-1.5 years to complete. Setting up and running simulation analyses requires infrastructure (i.e., computer code) to efficiently (i) configure monitoring choices, (ii) model in-season abundance and fishery dynamics so that simulated data are as realistic as possible, (iii) store large volumes of simulation results (e.g., 500 replications of chum salmon fishery data on multiple stocks could generate hundreds of Mb of output), and (iv) compare performance across alternative designs (again, challenging because each design has large output of information). A properly designed simulation framework could be flexible enough to add other topics, such as tagging programs, in-season management, and environmental drivers of chum salmon dynamics, to the suite of future monitoring and research options.

8 Tables

Table 1: Base stocks represented in the ChumGEM run reconstruction model.

Index (s)	Stock	Genetic Unit	Run timing
1	$\mathrm{FR_F}$	Fraser River	Fall
2	$\mathrm{JS_F}$	Johnstone Strait	Fall
3	$\mathrm{SOGE_F}$	Strait of Georgia East	Fall
4	$\mathrm{SOGW_F}$	Strait of Georgia West	Fall
5	$\mathrm{WCVI_F}$	West Coast Vancouver Island	Fall
6	$\mathrm{HC_S}$	Hood Canal	Summer
7	$\mathrm{HC_F}$	Hood Canal	Fall
8	$\mathrm{JDF_S}$	Juan de Fuca	Summer
9	$\mathrm{JDF_F}$	Juan de Fuca	Fall
10	$\mathrm{PSC_F}$	Puget Sound Central	Fall
11	$\mathrm{PSC_S}$	Puget Sound Central	Summer
12	$\mathrm{PSC_W}$	Puget Sound Central	Winter
13	$\mathrm{PSN_F}$	Puget Sound North	Fall
14	$\mathrm{PSS_S}$	Puget Sound South	Summer
15	$\mathrm{PSS_F}$	Puget Sound South	Fall
16	$\mathrm{PSS_W}$	Puget Sound South	Winter

Table 2. Stock-specific migration paths through fishery management areas. Each row represents one migratory component stock from the base stock indicated in Column 1. The SG symbol is a generic indicator for the spawning grounds, which is where escapements are enumerated.

Stock	x	Migration path
FR_F	1	C10,C11,C12,C13,C15,C16,C17,C29M,C29F,SG
FR_F	2	C10,C11,C12,C13,C14,C17,C29M,C29F,SG
FR_F	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,C19,U7,C18,C29M,C29F,SG
FR_F	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,C19,U7,U7A,C29M,C29F,SG
JS_F	1	C10,C11,SG
JS_F	2	C10,C11,C12,SG
JS_F	3	C10,C11,C12,C13,SG
SOGE_F	1	C10,C11,C12,C13,C15,SG
SOGE_F	2	C10,C11,C12,C13,C15,C16,SG
SOGE_F	3	C10,C11,C12,C13,C15,C16,C17,C29M,C28,SG
SOGE_F	4	C10,C11,C12,C13,C14,C17,C29M,C28,SG
SOGW_F	1	C10,C11,C12,C13,SG
SOGW_F	2	C10,C11,C12,C13,C14,SG
SOGW_F	3	C10,C11,C12,C13,C14,C17,SG
SOGW_F	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,C19,U7,C19,SG
WCVI_F	1	C10,C111,C127,C27,SG
WCVI_F	2	C10,C111,C127,C126,C26,SG
WCVI_F	3	C10,C111,C127,C126,C125,C25,SG
WCVI_F	4	C10,C111,C127,C126,C125,C124,C24,SG
WCVI_F	5	C10,C111,C127,C126,C125,C124,C123,C23,SG
WCVI_F	6	C10,C111,C127,C126,C125,C124,C123,C121,C21,C22,SG
HC_S	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U12,U12B,U12A,SG
HC_S	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_S	3	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_S	4	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_S	5	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12A,SG
HC_S	6	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12A,SG
HC_S	7	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_S	8	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_S	9	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12A,SG
HC_S	10	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12A,SG
HC_F	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U12,U12B,U12A,SG
HC_F	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_F	3	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_F	4	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_F	5	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12A,SG

HC_F	6	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6B,U9,U12,U12B,U12A,SG
HC_F	7	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_F	8	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12C,U12D,SG
HC_F	9	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12A,SG
HC_F	10	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U12,U12B,U12A,SG
JDF_S	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,SG
JDF_S	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,SG
JDF_S	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,SG
JDF_S	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6D,SG
JDF_S	5	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,SG
JDF_S	6	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6B,SG
JDF_S	7	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6B,SG
JDF_S	8	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6D,SG
JDF_S	9	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6D,SG
JDF_S	10	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6C,SG
JDF_S	11	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6C,SG
JDF_S	12	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6C,U5,SG
JDF_S	13	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6C,U5,SG
JDF_S	14	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6C,U5,U4B,SG
JDF_S	15	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6C,U5,U4B,SG
JDF_S	16	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,SG
JDF_S	17	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,SG
JDF_S	18	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6D,SG
JDF_S	19	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6D,SG
JDF_S	20	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6C,SG
JDF_S	21	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6C,SG
JDF_S	22	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6C,U5,SG
JDF_S	23	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6C,U5,SG
JDF_S	24	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6C,U5,U4B,SG
JDF_S	25	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6C,U5,U4B,SG
JDF_F	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,SG
JDF_F	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,SG
JDF_F	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,SG
JDF_F	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6D,SG
JDF_F	5	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,SG
JDF_F	6	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6B,SG
JDF_F	7	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6B,SG
JDF_F	8	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6D,SG
JDF_F	9	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6D,SG
JDF_F	10	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6C,SG
JDF_F	11	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6C,SG
JDF_F	12	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6C,U5,SG

JDF_F	13	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6C,U5,SG
JDF_F	14	C10,C11,C12,C13,C15,C16,C17,C29M,C18,U7,U6,U6C,U5,U4B,SG
JDF_F	15	C10,C11,C12,C13,C14,C17,C29M,C18,U7,U6,U6C,U5,U4B,SG
JDF_F	16	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,SG
JDF_F	17	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,SG
JDF_F	18	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6D,SG
JDF_F	19	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6D,SG
JDF_F	20	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6C,SG
JDF_F	21	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6C,SG
JDF_F	22	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6C,U5,SG
JDF_F	23	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6C,U5,SG
JDF_F	24	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6C,U5,U4B,SG
JDF_F	25	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6C,U5,U4B,SG
PSC_F	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U10E,SG
PSC_F	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U10A,SG
PSC_F	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U11A,SG
PSC_F	4	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10A,SG
PSC_F	5	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10A,SG
PSC_F	6	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10E,SG
PSC_F	7	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10E,SG
PSC_F	8	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U11A,SG
PSC_F	9	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U11A,SG
PSC_S	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U10E,SG
PSC_S	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U10A,SG
PSC_S	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U11A,SG
PSC_S	4	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10A,SG
PSC_S	5	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10A,SG
PSC_S	6	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10E,SG
PSC_S	7	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10E,SG
PSC_S	8	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U11A,SG
PSC_S	9	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U11A,SG
PSC_W	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U10E,SG
PSC_W	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U10A,SG
PSC_W	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U11A,SG
PSC_W	4	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10A,SG
PSC_W	5	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10A,SG
PSC_W	6	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10E,SG
PSC_W	7	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U10E,SG
PSC_W	8	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U11A,SG
PSC_W	9	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U11A,SG

PSN_F	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U7,U7B,U7C,SG
PSN_F	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U7,U7B,U7D,SG
PSN_F	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6A,U8,U8A,SG
PSN_F	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U8A,SG
PSN_F	5	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U7B,U7D,SG
PSN_F	6	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U7B,U7D,SG
PSN_F	7	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U7B,U7C,SG
PSN_F	8	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U7B,U7C,SG
PSN_F	9	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6A,U8,U8A,SG
PSN_F	10	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6A,U8,U8A,SG
PSN_F	11	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U8A,SG
PSN_F	12	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U8A,SG
PSS_S	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_S	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_S	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_S	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_S	5	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_S	6	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_S	7	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_S	8	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_S	9	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_S	10	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_S	11	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_S	12	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_S	13	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_S	14	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_S	15	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_S	16	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_S	17	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_S	18	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_S	19	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_S	20	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_S	21	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_S	22	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_S	23	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_S	24	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_F	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_F	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_F	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_F	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_F	5	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG

PSS_F	6	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_F	7	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_F	8	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_F	9	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_F	10	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_F	11	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_F	12	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_F	13	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_F	14	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_F	15	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_F	16	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_F	17	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_F	18	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_F	19	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_F	20	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_F	21	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_F	22	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_F	23	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_F	24	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_W	1	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_W	2	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_W	3	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_W	4	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_W	5	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_W	6	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_W	7	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_W	8	C10,C111,C127,C126,C125,C124,C123,C121,C20,U4B,U5,U6C,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_W	9	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_W	10	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13A,SG
PSS_W	11	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_W	12	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13C,SG
PSS_W	13	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_W	14	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13E,SG
PSS_W	15	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_W	16	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13F,SG
PSS_W	17	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_W	18	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13G,SG
PSS_W	19	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_W	20	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13K,SG
PSS_W	21	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_W	22	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13J,SG
PSS_W	23	C10,C11,C12,C13,C15,C16,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG
PSS_W	24	C10,C11,C12,C13,C14,C17,C29M,U7A,U7,U6,U6B,U9,U10,U11,U13,U13D,U13H,U13I,SG

Table 3: Fishery managemet areas and corresponding lengths. The M and F symbols for Area C29 indicate marine and freshwater areas, respectively.

Index (a)	Area	Length (km)
1	C10	33.3
2	C11	25.7
3	C12	133
4	C13	99.4
5	C14	95.8
6	C15	59
7	C16	33.8
8	C17	30.9
9	C18	60
10	C19	59
11	C20	57
12	C21	4.5
13	C22	47.8
14	C23	30
15	C24	27
16	C25	30
17	C26	31
18	C27	8.8
19	C29M	52.4
20	C29F	100
21	C28	25.1
22	C111	28
23	C121	36.7
24	C123	45.3
25	C124	77.5
26	C126	61.2
27	C125	74.2
28	C127	84.6
29	CR2	999
30	U4B	29.1
31	U5	34
32	U6	11.6
33	U6A	4.7
34	U6B	7.8
35	U6C	30.6
36	U6D	10
37	U7	57.9
38	U7A	38.7
39	U7B	31.5
40	U7C	10
41	U7D	10
42	U8	46
43	U8A	37.8
44	U8D	5

45	U9	45
46	U9A	37.8
47	U10	33.8
48	U10A	3.8
49	U10D	10
50	U10E	22.5
51	U10F	11.7
52	U10G	14.4
53	U11	31.7
54	U11A	5.1
55	U12	28.9
56	U12A	11.3
57	U12B	24.2
58	U12C	19.9
59	U12D	19.2
60	U13	24.5
61	U13A	13.5
62	U13C	1.3
63	U13D	17.8
64	U13E	5.3
65	U13F	11.5
66	U13G	13
67	U13H	16
68	U13I	6.1
69	U13J	17.4
70	U13K	9.2
71	SG	999

Table 4: Notation and symbols for run reconstruction model.

Symbol	Value	Description
Indexes		
s	$\{1, 2, ..., S\}$	Base stock
x	$\{1, 2, ..., X_s\}$	Component of stock s
a	$\{1, 2, ..., A\}$	Fishery management area
t	$\{1, 2, ..., T\}$	Day
g	$\{1, 2, ..., G\}$	GSI sample
Parameters		
R_s	Estimated	Run size of stock s
μ_s	Estimated	Mean date of arrival to area C10 for stock s
σ_s	Estimated	Standard deviation around μ_s for stock s
κ_{sx}		Proportion of base stock s run size attributable to component x
E^*_s		Escapement bias for stock s
X_s		Number of components for stock s
$X_{s,north}$		Number of components from stock s taking the northern diversion route
$\boldsymbol{m}_{sx}$		Migration path of component x for stock s
$\overline{m_{sx}}$		Number of areas in migration path of component x for stock s
l_a		Length of area a
ρ	37.7	Migration rate (km/day)
σ_P	6	Spread of migration proportions
R^*_s		Mean of prior distribution for run size of stock s
σ_{Rs}		Standard deviation around R^*_s
μ^*_s		Mean of prior distribution for the mean arrival of stock s to C10
$\sigma_{\mu s}$		Standard deviation around μ^*_s
α_1		Shape parameter for inverse gamma prior on σ_s
α_2		Scale parameter for inverse gamma prior on σ_s
a_g		Area in which GSI sample g was taken
t_{1g}		Start date of GSI sample g
t_{2g}		End date of GSI sample g
Unobserved states		
A_{st}		Proportion of total run for stock s arriving on day t
N_{sxat}		Number of fish from component x of stock s in area a on day t
R_s		Run size of stock s
U'_{at}		Raw harvest rate in area a on day t
U_{at}		Adjusted harvest rate in area a on day t
C_{sat}		Estimated catch from stock s in area a on day t
C_{at}		Estimated catch in area a on day t
E_{st}		Estimated daily spawning ground escapement for stock s
I_{at}		Estimated catch per unit effort in area a on day t
Δ_{sxk}		Distance between area C10 and the kth area in the migration path of component x of stock s
$\boldsymbol{D'}$		Starting points of movement step calculation
$\boldsymbol{D}$		Cutpoints of movement step calculation
$\boldsymbol{\Psi}$		Mean landing point in movement step calculations
Q_{sxak}		Proportion of chum from component x of stock s in area a moving at least as far as the kth area in its migration path at each time step
P'_{sxak}		Proportion of chum from component x of stock s in area a moving to the kth area in its migration path at each time step
P_{sxaj}		Proportion of chum from component x of stock s in area a that will move to area j at each time step
$\boldsymbol{z}_a$		Daily log-catchability residuals

q'_a	Log-catchability for the area a test fishery
q_a	Catchability for the area a test fishery
t_{1g}	First day of sampling for GSI sample g
t_{2g}	Final day of sampling for GSI sample g
J_a	Number of CPUE observations in area a
W_{gs}	Proportion of fish from stock s in area a_g between t_{1g} and t_{2g}
ξ_{gs}	
τ_g	
ζ_g	Size of GSI sample g
ζ^*	Upper limit sample size, above which GSI samples are assumed not to be more accurate
υ_{gs}	Variance the proportion of fish in sample g attributed to stock s
Observed states	
$\widehat{C_{at}}$	Number of fish caught in area a on day t
$\widehat{E_s}$	Total spawning ground escapement for stock s
$\widehat{I_{at}}$	Catch per unit effort (catch/number of sets) in area a on day t
$\widehat{W_{gs}}$	Proportion of fish in sample g attributed to stock s

9 Figures

Figure 1: Conceptual model of chum salmon migration routes through Canadian waters.

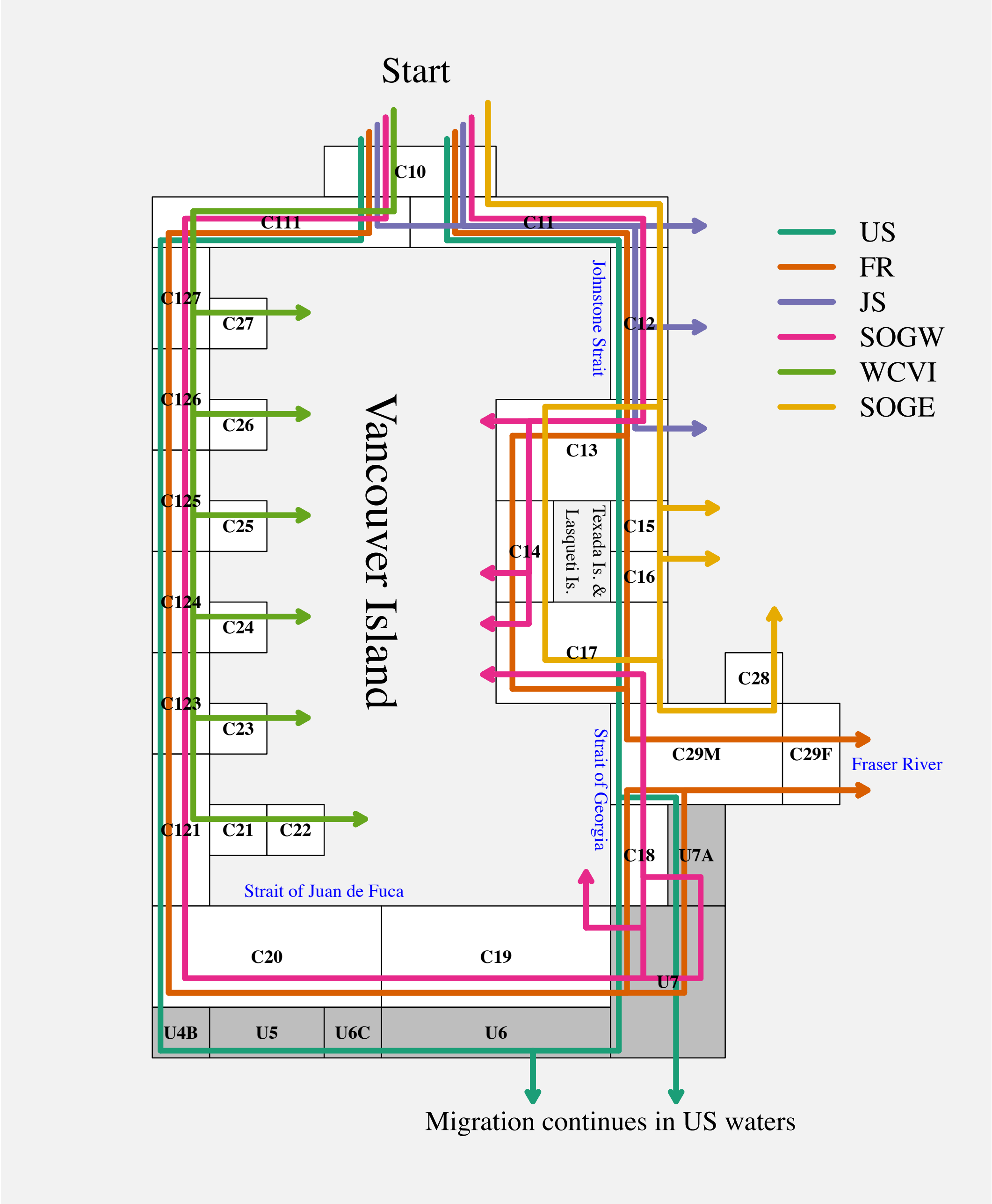


Figure 2: Conceptual model of chum salmon migration routes through U.S.A. waters.

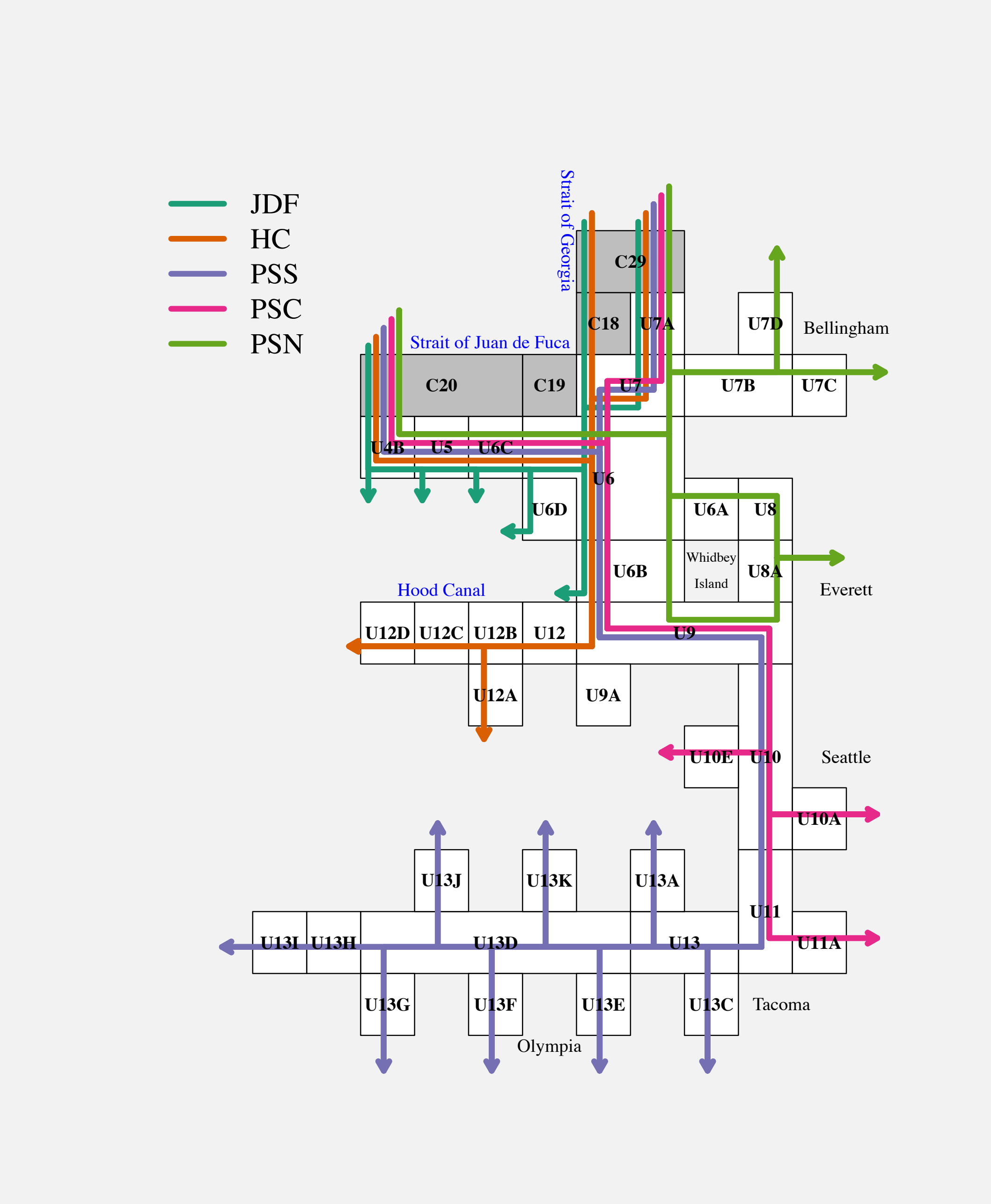


Figure 3: Observed escapement by stock for the year 2012.

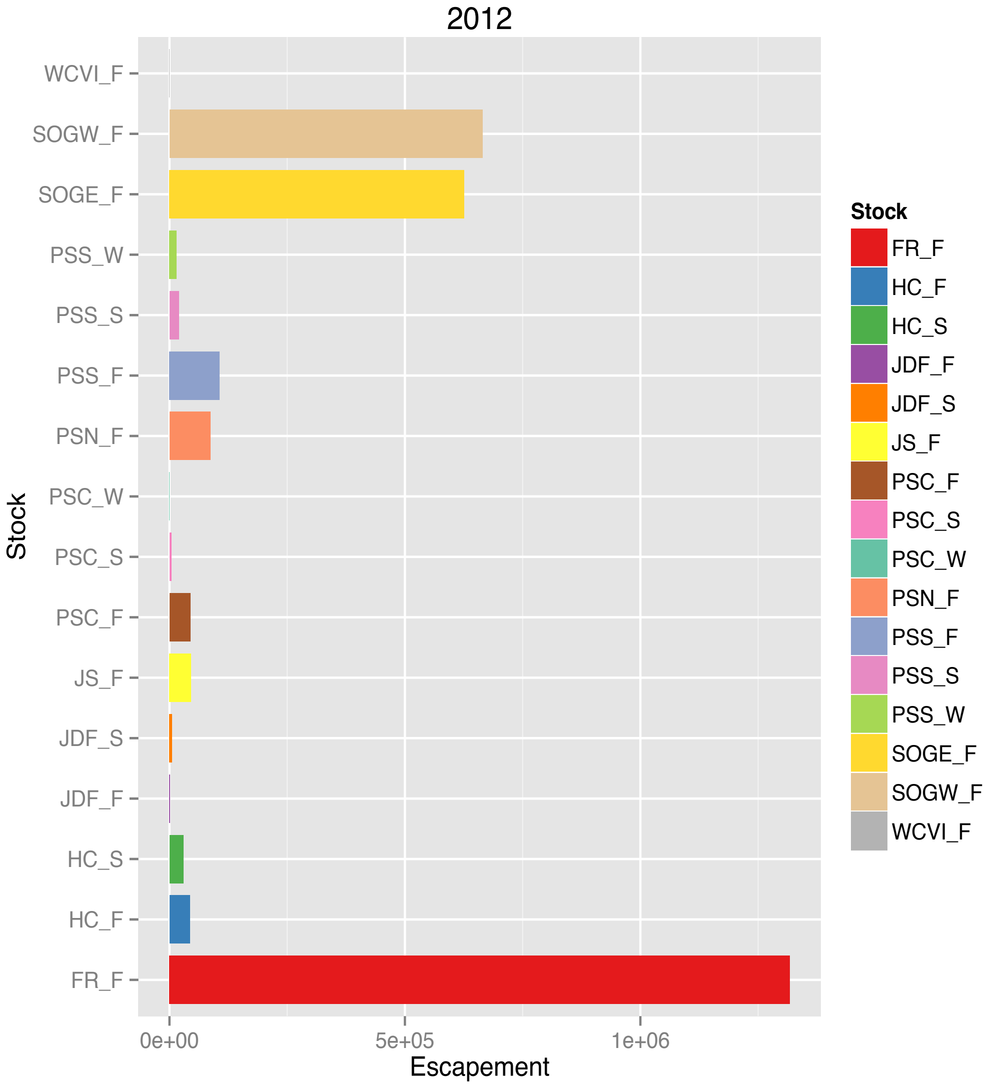


Figure 4: Observed catches by area and gear type for year 2012.

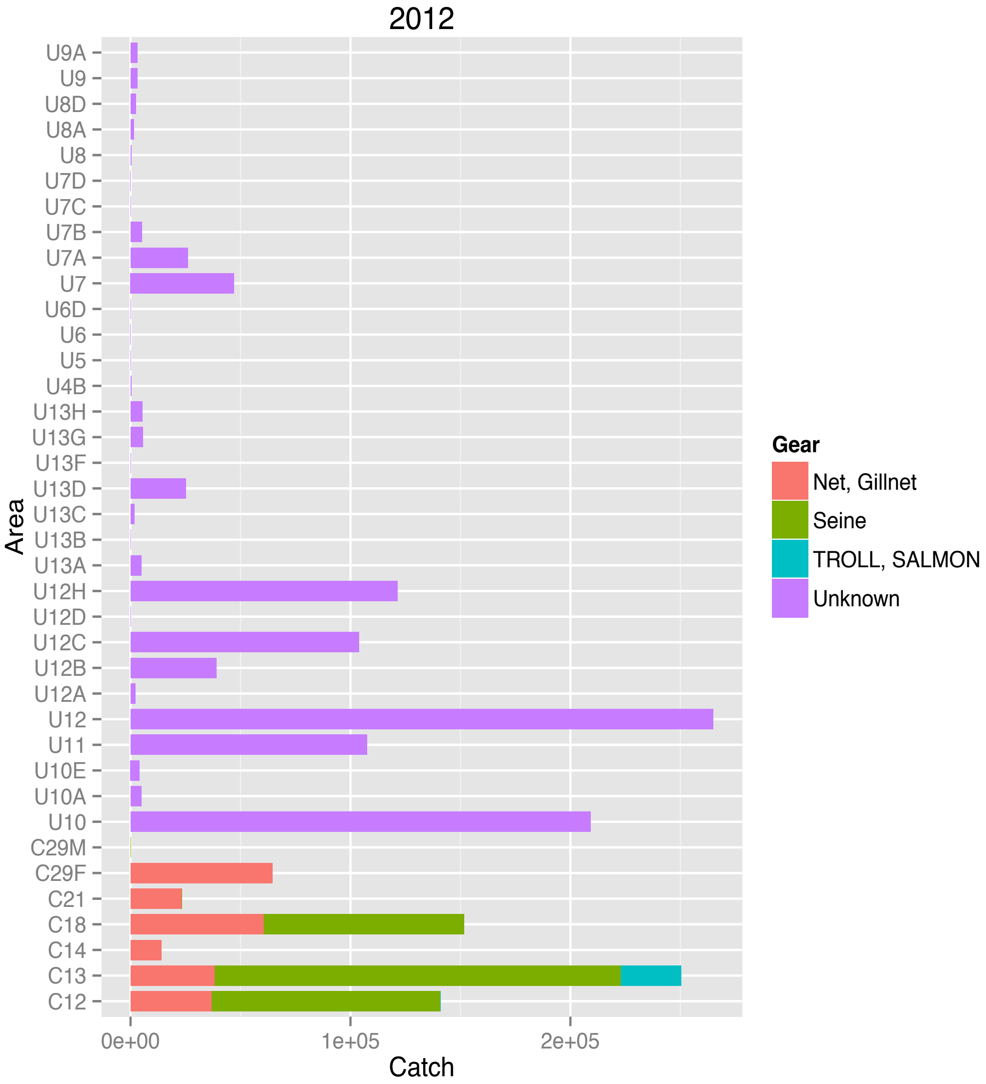


Figure 5: Observed daily catch by area for year 2012.

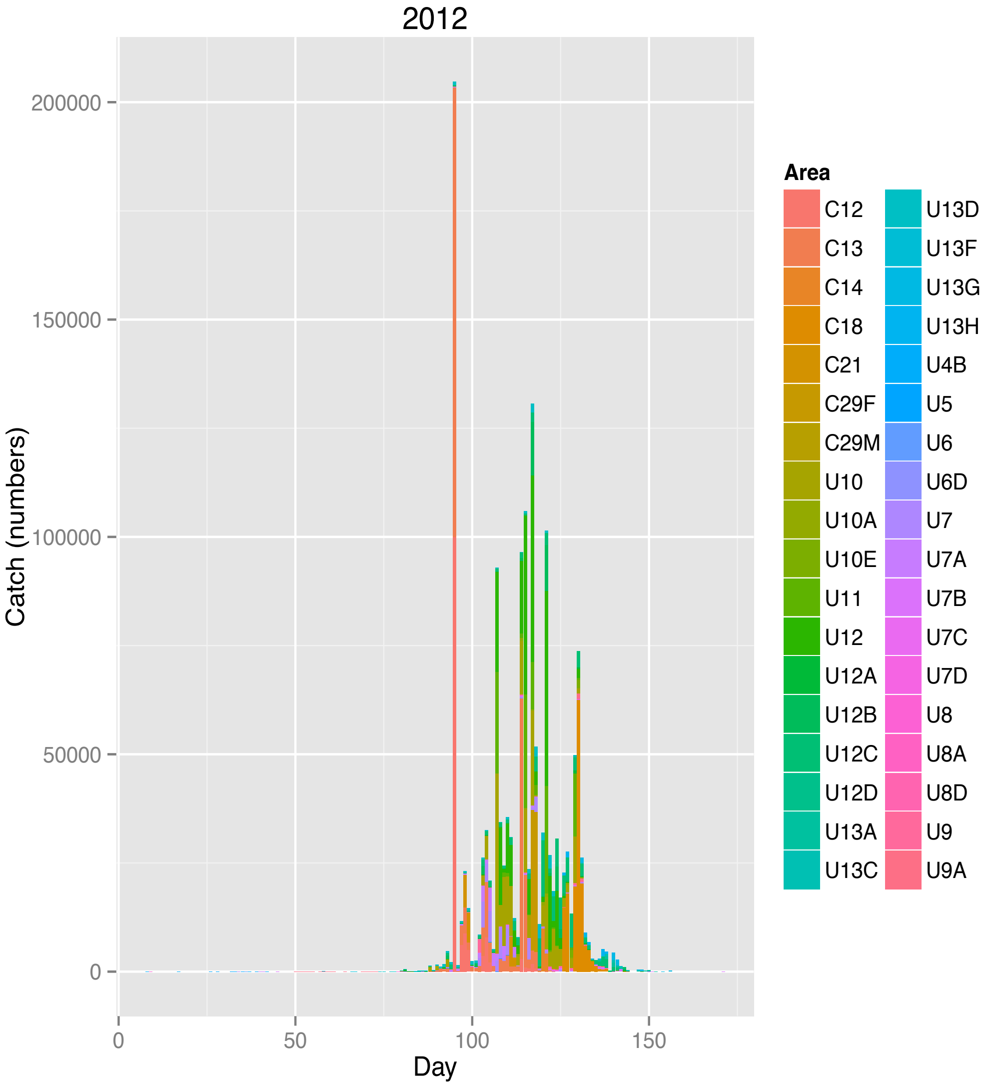


Figure 6: Observed daily CPUE for year 2012.

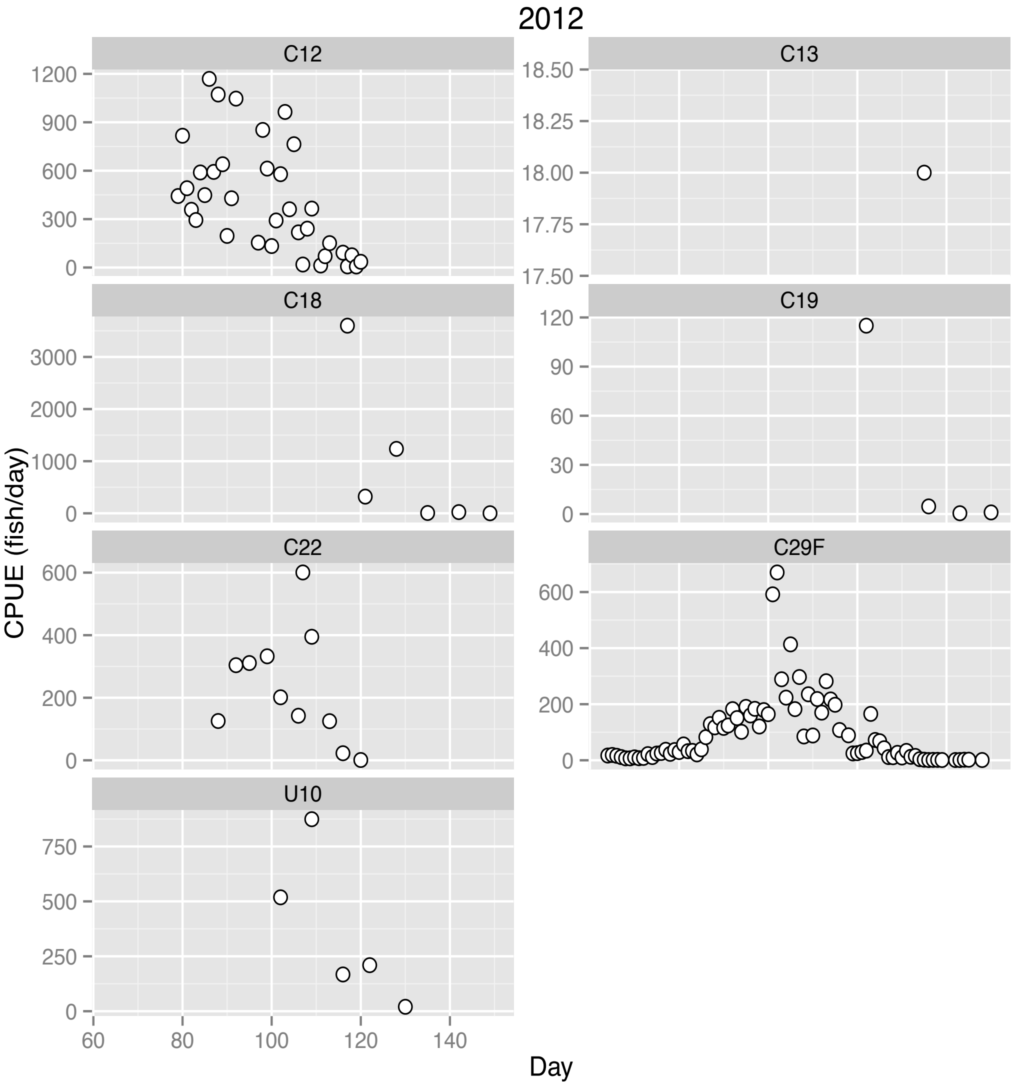


Figure 7: Proportions of stocks present in 30 GSI samples for year 2012.

Figure 8: Parameter estimates (dots) and standard errors (lines) obtained by applying the ChumGEM run reconstruction model to 2012 data.

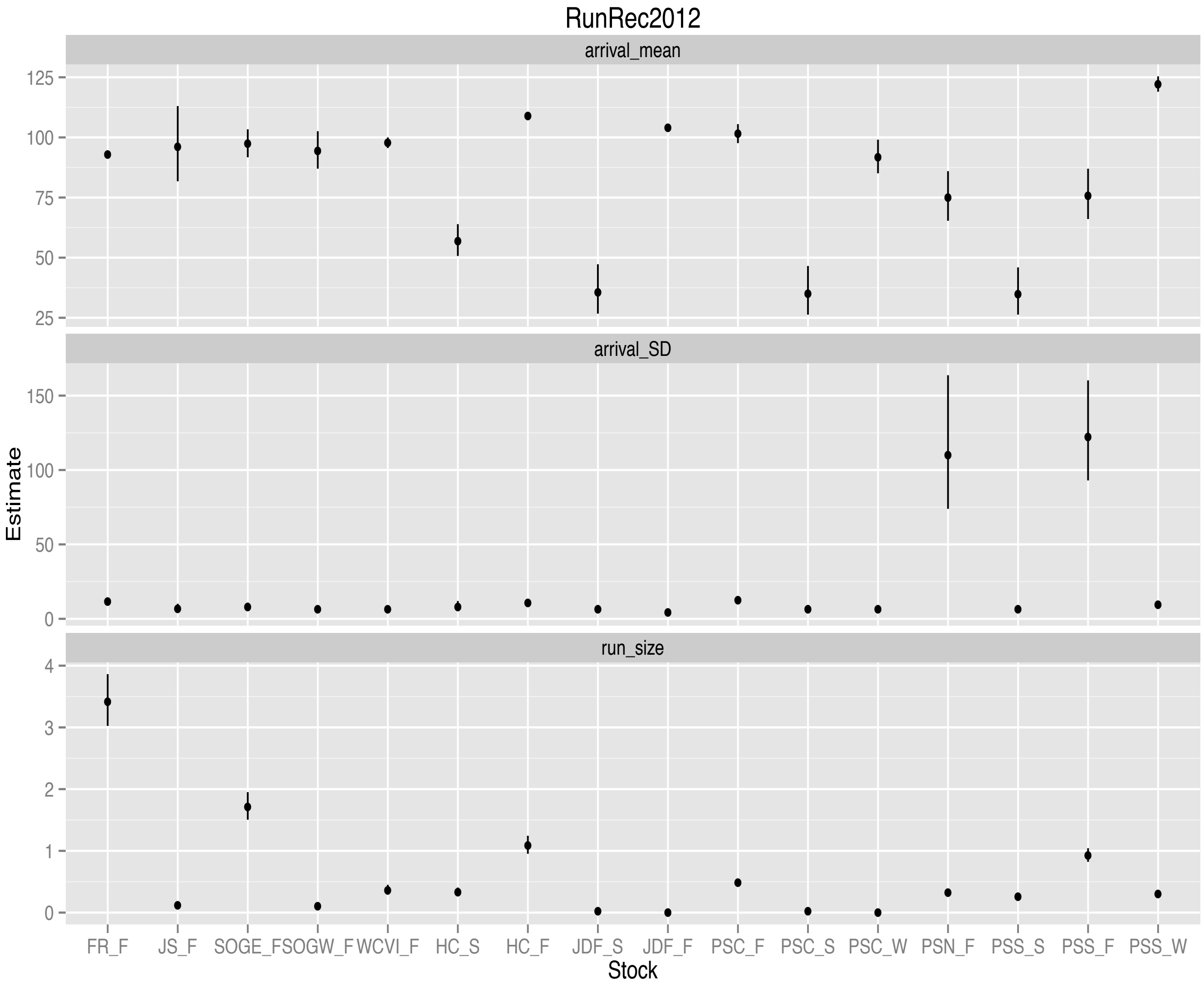


Figure 9: Observed catches (dots) and estimated catches (bars) obtained by applying the ChumGEM run reconstruction model to 2012 data.

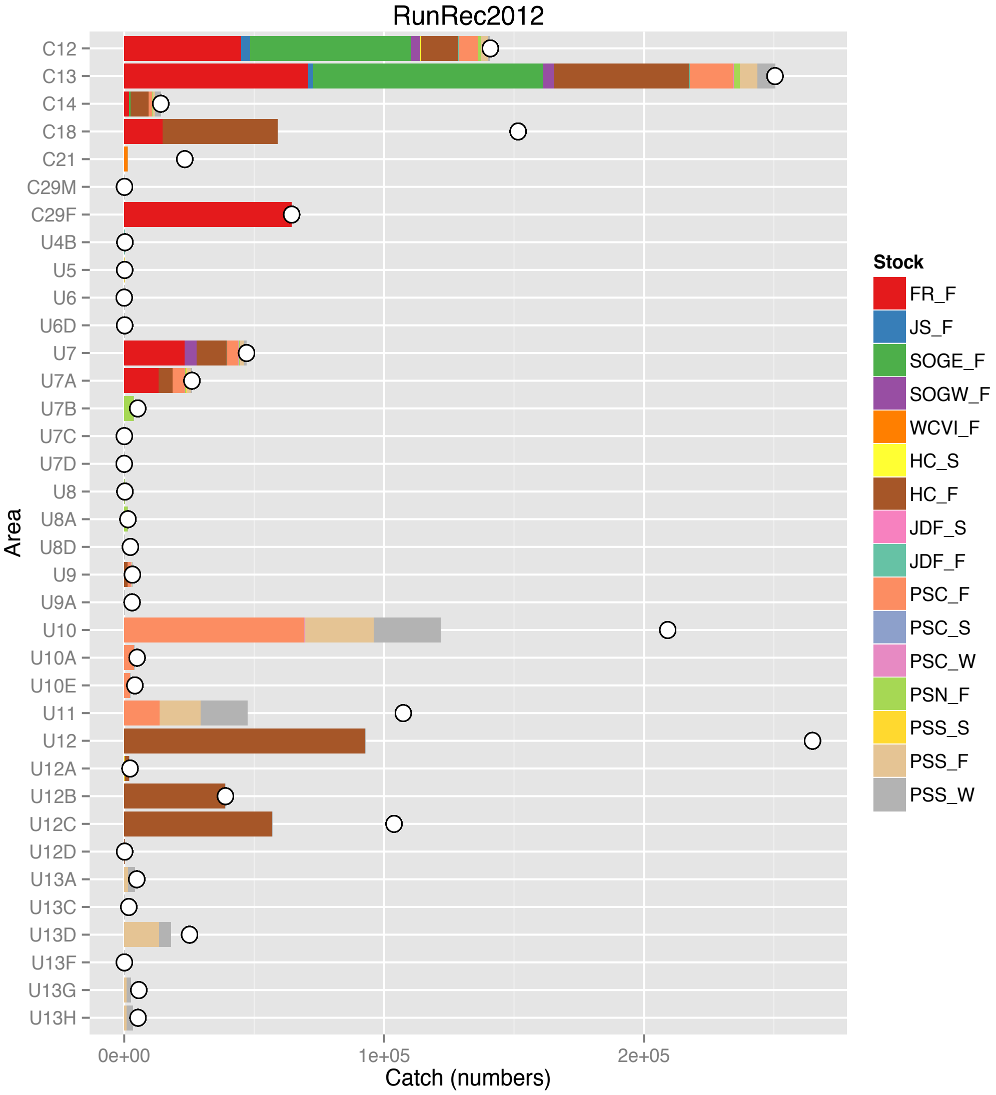


Figure 10: Observed escapement (dots) and estimated escapement (bars) obtained by applying the ChumGEM run reconstruction model to 2012 data.

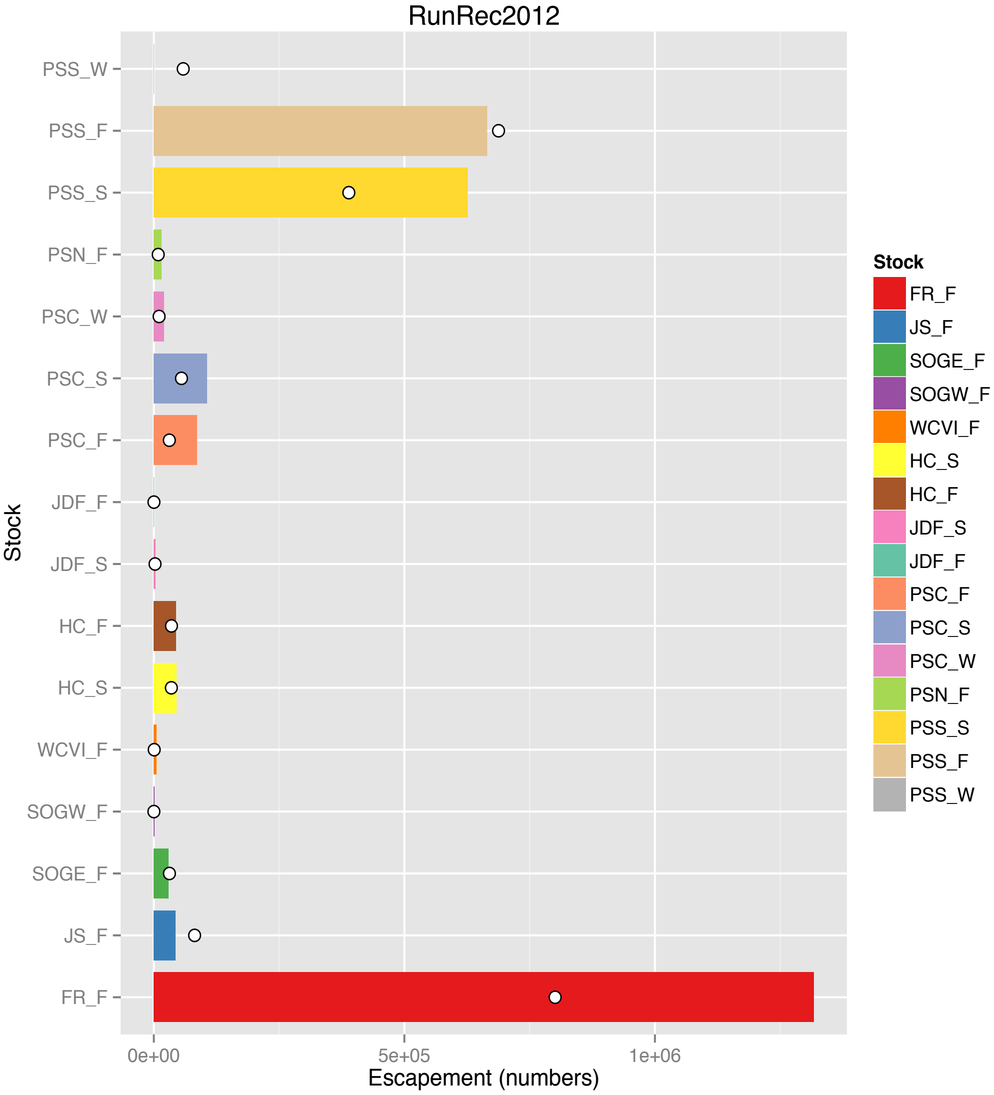


Figure 11: Observed daily CPUE (dots) and estimated daily CPUE (bars) by area obtained by applying the ChumGEM run reconstruction model to 2012 data.

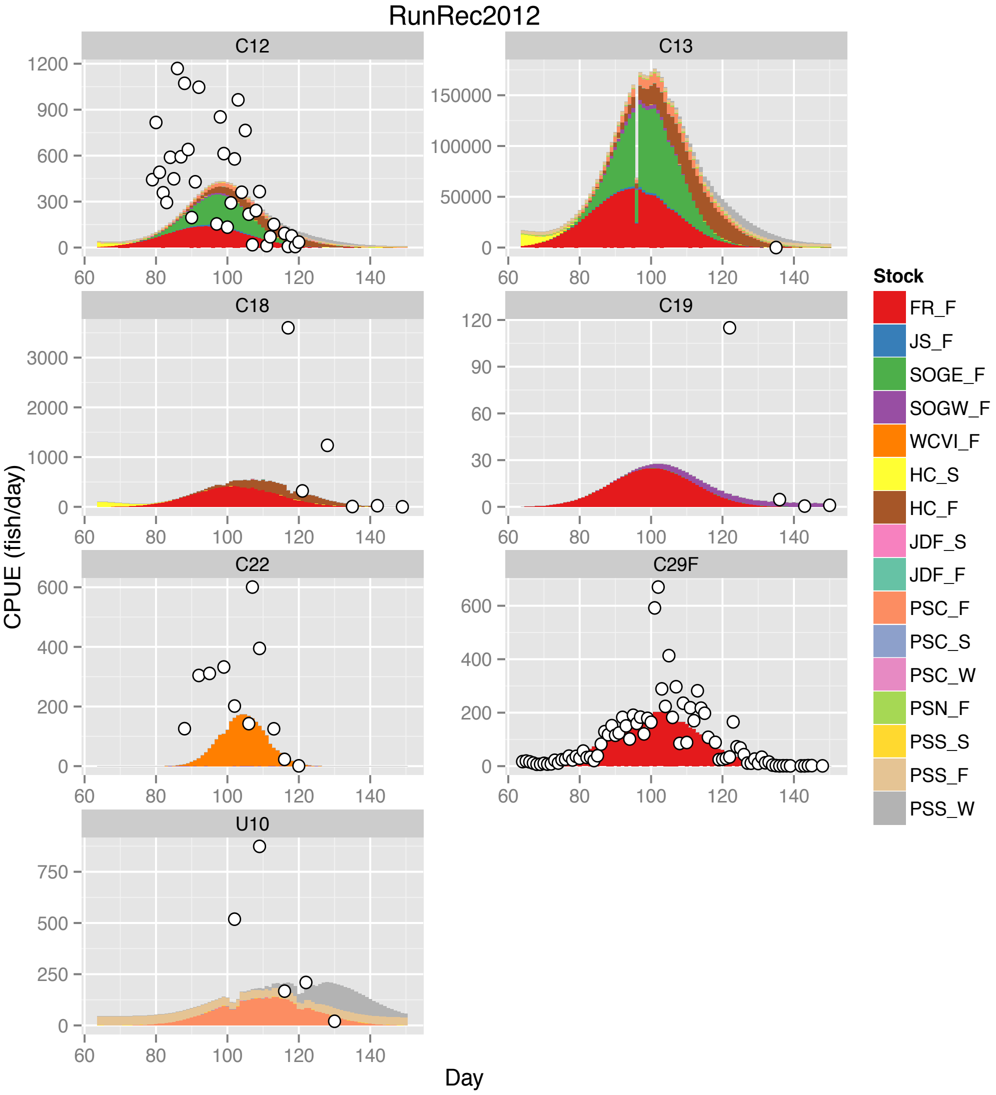
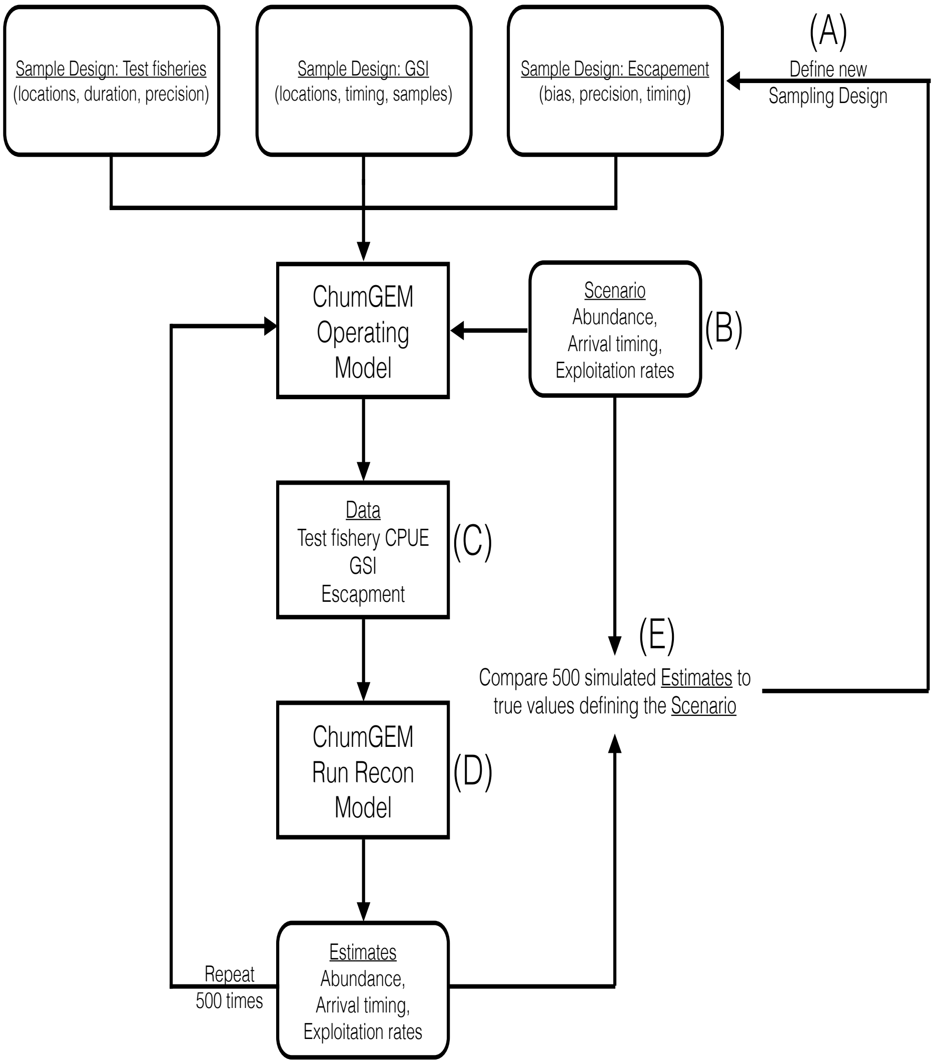
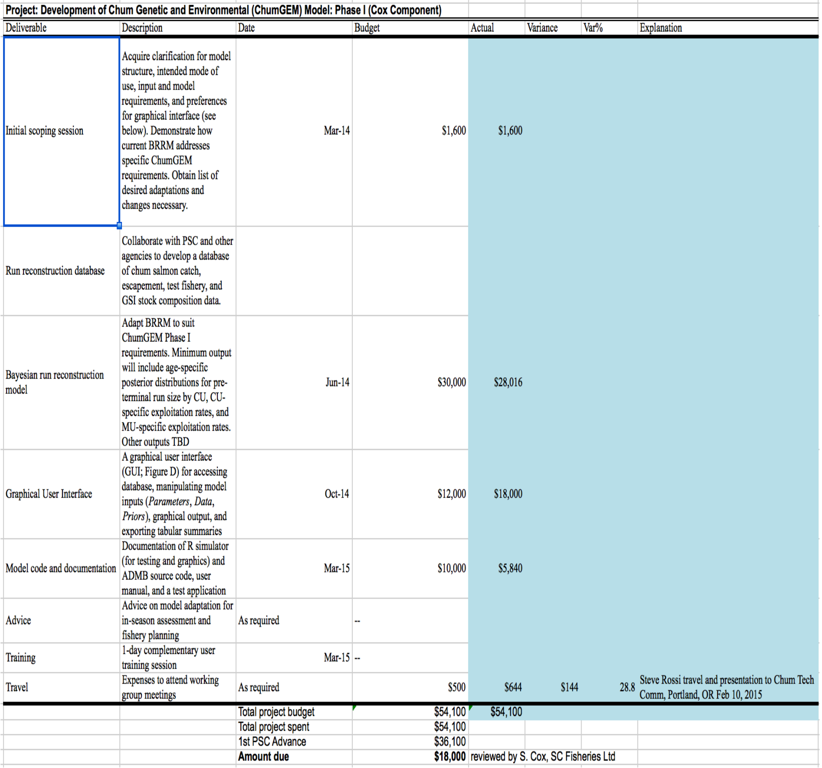


Figure 12: Simulation framework for evaluating alternative sampling designs by (A) specifying specific sampling designs for Test Fisheries, GSI, and Escapement monitoring, (B) defining a true operating model scenario representing the true abundance and exploitation dynamics, (C) generating observed test fishery CPUE, GSI, and escapement data, (D) using the data and ChumGEM run reconstruction model to estimate abundance, arrival timing, and exploitation rates, and (E) comparing estimation performance against the true scenario conditions. This approach is repeated for alternative sampling designs to assess trade-offs between assessment performance, program cost, and feasibility.



A Appendix 1: Financial Statement of Expenditures

Figure 13: Planned budget comparison to actual expenditures.



Project: Development of Chum Genetic and Environmental (ChumGEM) Model: Phase I (Cox Component)

Deliverable	Description	Date	Budget	Actual	Variance	Var%	Explanation
Initial scoping session	Acquire clarification for model structure, intended mode of use, input and model requirements, and preferences for graphical interface (see below). Demonstrate how current BRRM addresses specific ChumGEM requirements. Obtain list of desired adaptations and changes necessary.	Mar-14	$1,600	$1,600			
Run reconstruction database	Collaborate with PSC and other agencies to develop a database of chum salmon catch, escapement, test fishery, and GSI stock composition data.						
Bayesian run reconstruction model	Adapt BRRM to suit ChumGEM Phase I requirements. Minimum output will include age-specific posterior distributions for pre-terminal run size by CU, CU-specific exploitation rates, and MU-specific exploitation rates. Other outputs TBD	Jun-14	$30,000	$28,016			
Graphical User Interface	A graphical user interface (GUI; Figure D) for accessing database, manipulating model inputs (*Parameters*, *Data*, *Priors*), graphical output, and exporting tabular summaries	Oct-14	$12,000	$18,000			
Model code and documentation	Documentation of R simulator (for testing and graphics) and ADMB source code, user manual, and a test application	Mar-15	$10,000	$5,840			
Advice	Advice on model adaptation for in-season assessment and fishery planning	As required	--				
Training	1-day complementary user training session	Mar-15	--				
Travel	Expenses to attend working group meetings	As required	$500	$644	$144	28.8	Steve Rossi travel and presentation to Chum Tech Comm, Portland, OR Feb 10, 2015
		Total project budget	$54,100	$54,100			
		Total project spent	$54,100				
		1st PSC Advance	$36,100				
		Amount due	**$18,000**	reviewed by S. Cox, SC Fisheries Ltd			

B Appendix 2: Project Summary

B.1 Project Deliverables

The following table summarizes performance against the seven primary deliverables.

Deliverable	Description	Performance
Initial scoping session	Acquire clarification for model structure, intended mode of use, input and model requirements, and preferences for graphical interface (see below). Demonstrate how current BRRM addresses specific ChumGEM requirements. Obtain list of desired adaptations and changes necessary.	Several scoping, data review, and model development sessions were completed over the course of the project.
Run reconstruction database	Collaborate with PSC and other agencies to develop a database of chum salmon catch, escapement, test fishery, and GSI stock composition data.	We were able to combine a variety of different data types and formats into a single common data structure within the ChumGEM software, while retaining the flexibility for users to maintain their own CSV format data files. Users edit CSV files, which are then IMPORTed in ChumGEM to a common database format. The current ChumGEM distribution holds this database, although it is easily exported.
Bayesian run reconstruction model	Adapt BRRM to suit ChumGEM Phase I requirements. Minimum output will include age-specific posterior distributions for pre-terminal run size by CU, CU-specific exploitation rates, and MU-specific exploitation rates. Other outputs TBD	This deliverable was achieved with two modifications. First, age-specific abundance data were not available, so outputs are by genetic unit, not age-specific. This was agreed with the Chum Tech Committee. Second, the full Bayesian MCMC analysis is possible, but not recommended at this time because of the complexity of the model. The BRRM currently outputs approximations to the posterior distributions as measures of uncertainty.

Graphical User Interface	A graphical user interface (GUI) for accessing database, manipulating model inputs (Parameters, Data, Priors), graphical output, and exporting tabular summaries	This deliverable was achieved exactly as planned.
Model code and Documentation	Documentation of R simulator (for testing and graphics) and ADMB source code, user manual, and a test application	This deliverable was achieved, although there is no simulator, just an estimator. A the term *simulator* was used in error in the original proposal. It was meant to be *estimator*, which is delivered as the BRRM.
Advice	Advice on model adaptation for in-season assessment and fishery planning	We have provided advice while developing and implementing the project, as well as part of interim meetings. Additional advice for adapting this model to the next steps in ChumGEM development were provided in the Recommendations section.
Training	1-day complementary user training session	This was delivered on 7 May 2015. An additional follow-up training day will be provided in August 2015.

B.2 Project Schedule

This project began 31 March 2014 and was scheduled for delivery 31 March 2015. We discussed with Chum Tech Committee that delivery in time for the training workshop would be acceptable. The training occurred on 7 May 2015 as per Chum Tech Committee meeting scheduling. The ChumGEM package was delivered then. Based on feedback obtained during the training, as well as some additional reports following the training, a final revised ChumGEM package was delivered 10 July 2015.

B.3 QA/QC

Progress toward completion was measured via approximately 5 meetings between SC Fisheries and the Chum Tech Committee.

B.4 Monitoring and evaluation

The ChumGEM package was initially tested during the planned training meeting and will be subsequently tested again in August 2015.

B.5 Benefits

It is too early to determine the Benefits coming from this project, but a few might be, e.g., (i) better conceptual understanding of chum salmon fishery and monitoring data, as well as run reconstruction modeling issues and challenges, (ii) a common database and graphics package for visualizing chum salmon fisheries, and (iii) a strong foundation upon which to adapt and build on the ChumGEM vision and strategic plan.