**Mission Sockeye Yearlings – CPUE Calculations**

K. Davidson   
*Last update: 14-Mar-2019 @ 12:03 HAPPY PI DAY  
Code in github repository: https://github.com/khdavidson/chum-et-al.git "2017CPUE\_calibrations.R”*

**Doc Outline**

* Objective/overview
  + Underlying calculation overview
* Methods summary to date
* Method 1: Daily discharge
* Method 2: RST volume and sampling event current velocity
* Method 3: Catch tables
* Appendix I: Environmental Data Exploration
* Appendix II: Early methods exploration/early thoughts on ways to correct and display data

**Objective/Overview**

Previous CPUE has been given as the number of fish per run, or as the total number of fish/number of runs (e.g., Townsend et al 2017). However, this does not take into account differences in catch between trap types (RST vs. VT), or changes in the river water volume over the field season. The following document outlines some potential calculations to account for these variations in fishing success. For simplicity, sockeye yearlings are referred to as ‘fish’ from hereon.

Calculations here will be based upon the following three-step process, which may be adjusted to account for trap types or sampling depths in future. For simplicity, the final estimated number of fish is referred to as an ‘index of abundance’ as in Vernon (1966) and Todd (1966). The ‘Index of Fish Abundance’ converts ‘raw’ catch to estimate the abundance of fish passing by Mission at the scale of interest (sampling event, daily, month, year, etc.). Catch-per-unit-effort (CPUE) is the first stage of calculation where ‘fish/m3’ are estimated for each trap. CPUE describes the number of fish for a unit of effort; one unit of effort accounts for both trap type and run length (as shown below).

The following is a detailed look at possible calculations for the index of abundance (IA) for Sockeye yearlings. Each method is designed to increase in complexity and build upon previous methods. Presently, IA calculations are focused on the RST, which will (eventually) be used to ‘standardize’ other trap types (IPT, VT) as the RST is more effective at catching larger yearlings (the IPT and VT were designed to catch fry).

Underlying Calculation

The following three-step process outlines the calculations underlying each iteration of IA method development. Early methods may not use all steps, or may simplify some steps (Table 1).

Equation 1. ***CPUE*** (fish/m3) during a sampling run

= = = 0.027 fish/m3 (in a 900 s run)

Where *n* is the number of fish caught in the trap per run, *VT* is the trap volume calculated using: *AT*,the cross-sectional area of the trap opening in the water, *f* the average flow speed travelled while fishing, and *t* the run time (usually 900 seconds, but varies from 600 to 1320 seconds).

Equation 2. Water volume in Bay (***VB***) during a sampling run

= **⋅** *c* **⋅** *t*  
= (*d* **⋅** ) **⋅** 0.5 **⋅** 900   
= (1.13 **⋅** ) **⋅** 0.5 **⋅** 900   
= 74, 580 m3 (in a 900 s run)

Where *AB* is the cross-sectional area of the Bay calculated using: the depth of the trap, *d*, and one-third of the river width, *w*. *c* is the surface current velocity at time of sampling and *t* is the run time.

Equation 3. Index of abundance (IA)

= CPUE **⋅** VB  
= 0.027 **⋅** 74580   
= 2013.66 fish

**Methods Summary**

The following table summarizes current methods being explored, key parameters and outcomes. As this process is always changing and developing, this table only summarizes the current methods of interest. Previous methods for correcting catch that have been discounted are explained in Appendix II.

**Table 1.** Summary table outlining methods in consideration and the calculation parameters and input data used for each method. Outputs and key results are summarized versions of results given later in the document.

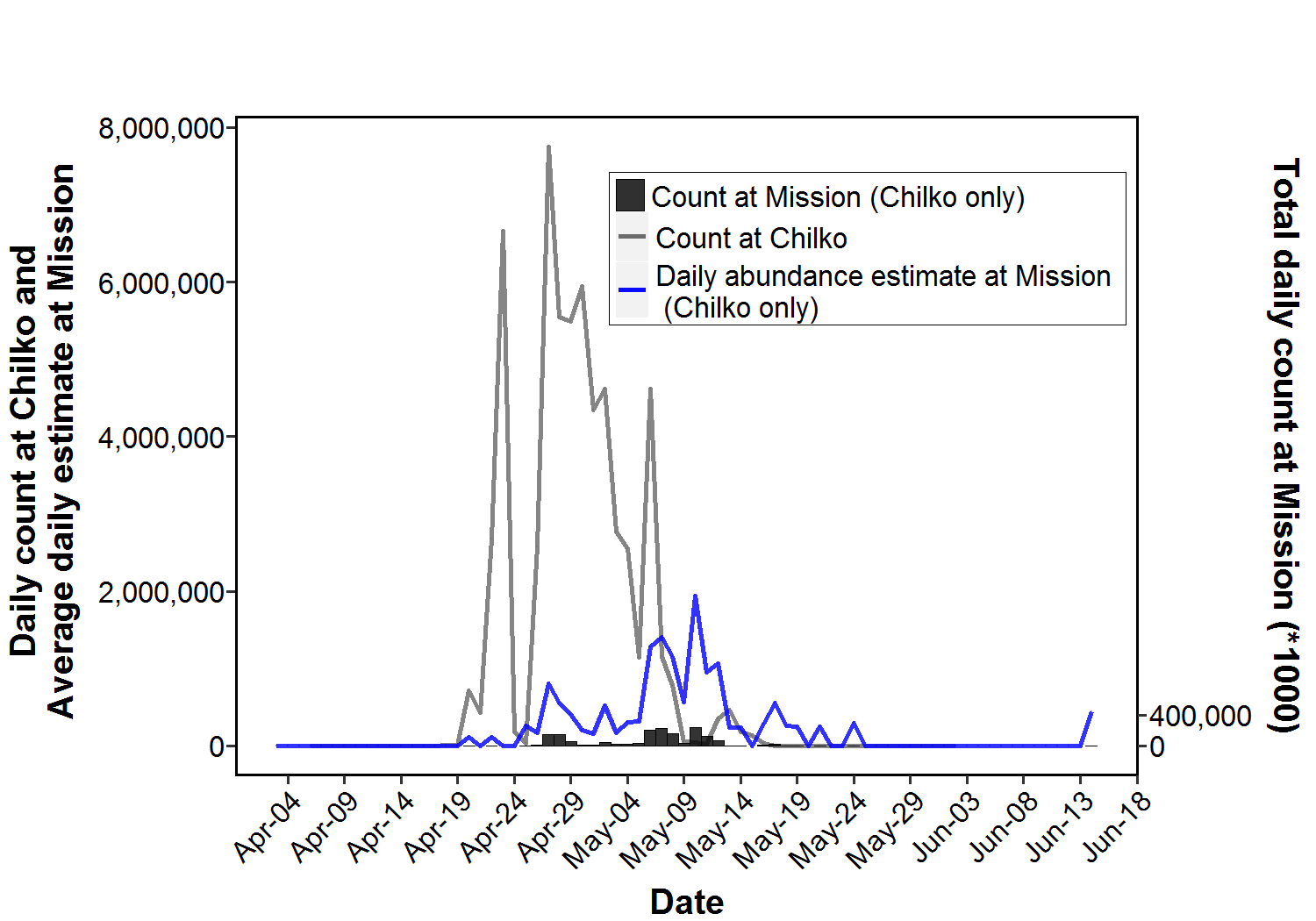
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Method name** | **Input data** | **Calculation Parameters** | **Outputs** | **Key results** | **Good method?** |
| **Method 1** | **⋅** RST catch expanded for sub-sampling  **⋅** RST volume  **⋅** Daily discharge | **⋅** Equation 1  **⋅** Daily average discharge and total day length | **⋅** Fig 1.1. Chilko comparison  **⋅** Fig 1.2. All CUs cumulative run timing  **⋅** (Table 2.) | **⋅** Seasonal effects of discharge confound abundance estimates  **⋅** Cumulative timing curves for all CUs change shape, but timing information is approximately the same | Likely not |
| **Method 2** | **⋅** RST catch expanded for sub-sampling  **⋅** RST volume  **⋅** Surface current (sampling event) | **⋅** Equations 1-3  **⋅** RST depth: 1.13 | **⋅** Fig 2.1. Chilko comparison  **⋅** Fig 2.2. Cumulative run timing comparison  **⋅** Table 2. Cumulative run timing | **⋅** Results seem less confounded by daily discharge patterns than Method 1  **⋅** Run timings and curves are almost identical to Method 1 | Possibly, with some finer adjustments |
| **Method 3** | **⋅** Catch tables using CPUE | **⋅** Equation 1 | **⋅** Table 3.1. Example matrix  **⋅** Table 3.2. Model comparison  **⋅** Table 3.3. NA summary statistics | **⋅** There are a lot of NAs.  **⋅** Bay 6 had slightly more sampling observations (41) than Bays 2 and 11 |  |
| **3.1: *imputeTS*** |  | **⋅** Mean  **⋅** Interpolation - approx  **⋅** Interpolation - spline  **⋅** Interpolation - stine |  |  | **⋅** No  **⋅**  **⋅**  **⋅** |

**Method 1. Daily discharge**

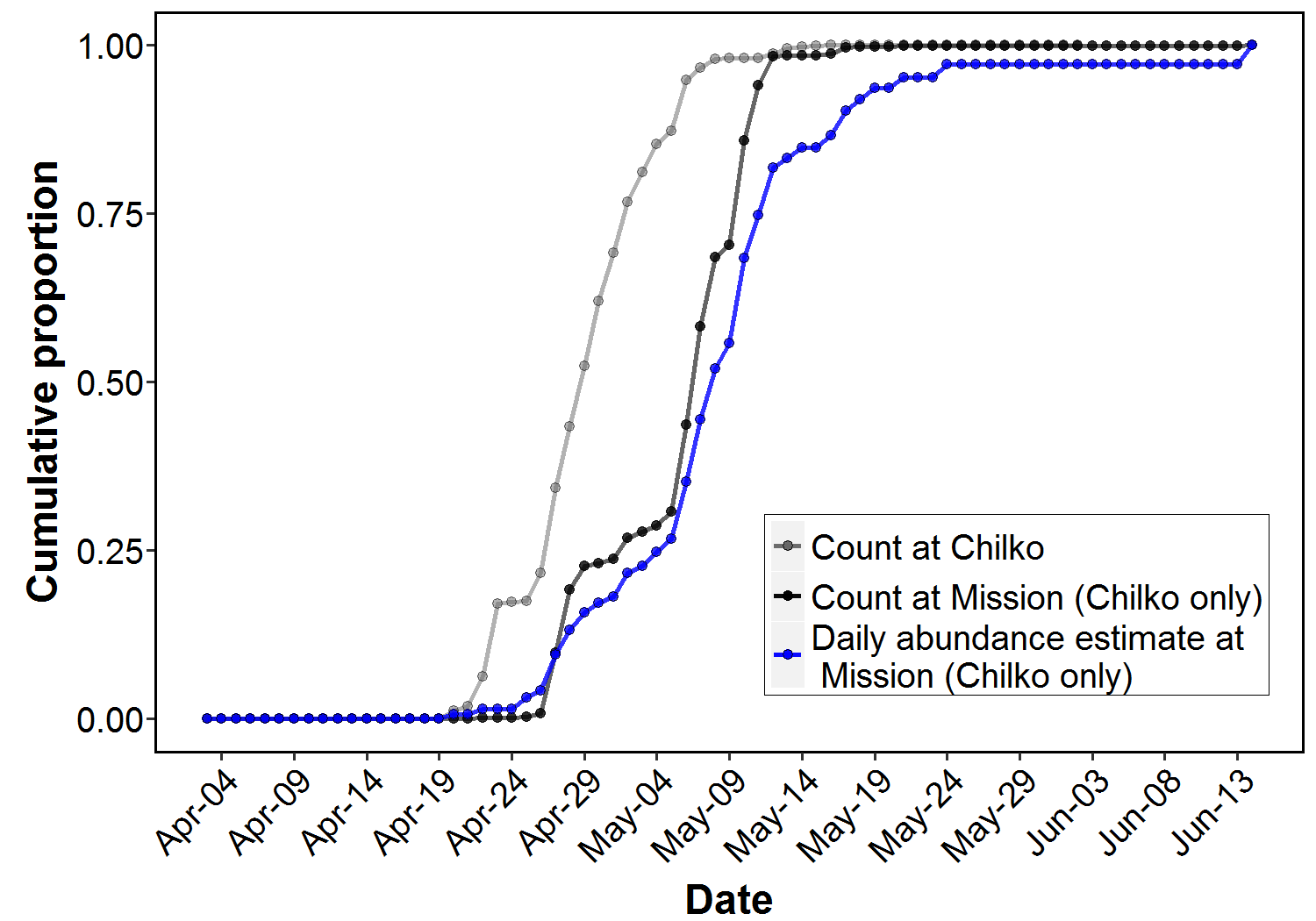
This method uses Equation 1 and considers RST catch (expanded for sub-sampling) only. Using average daily CPUE, the following calculation was done to obtain a rough estimate of daily fish passage per day (1 day = 86,400 sec):

Daily fish passage = daily average CPUE (fish/m3) ⋅ daily average discharge (m3/s) ⋅ 86400 s

Comparing these results to the Chilko counts, the numbers appear reasonable (Figure 1.1a). However, it is important to note that multiplying daily CPUE by daily discharge will confound catch results given that discharge continues to increase throughout the field season (Figure A1c). For example, catch on April 20 and June 14 were both n=1 fish. Discharge was 2400 m3/s on April 20 and 9390 m3/s on June 14, resulting in exaggerated estimates (i.e., relatively taller peaks) later in the season.



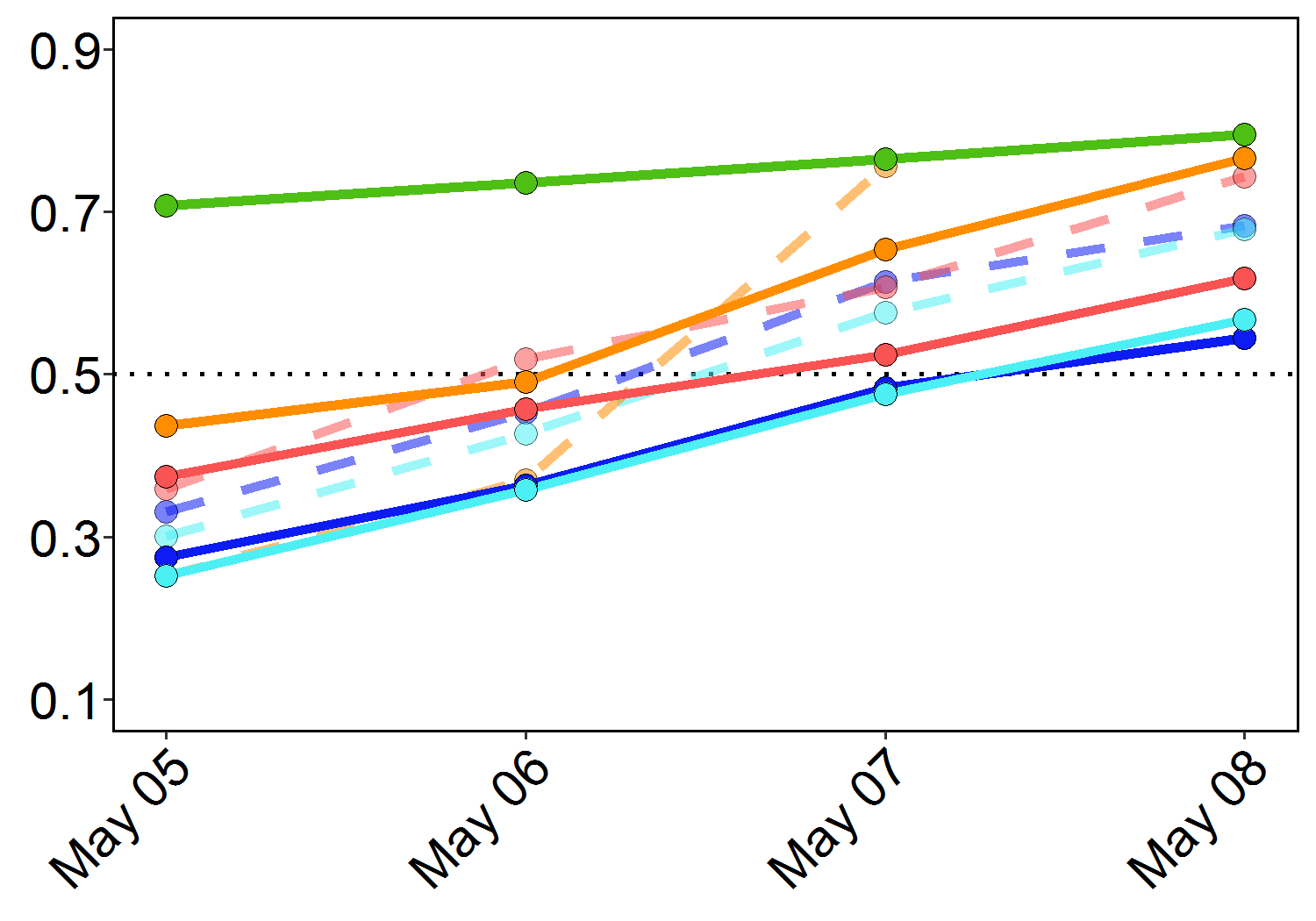
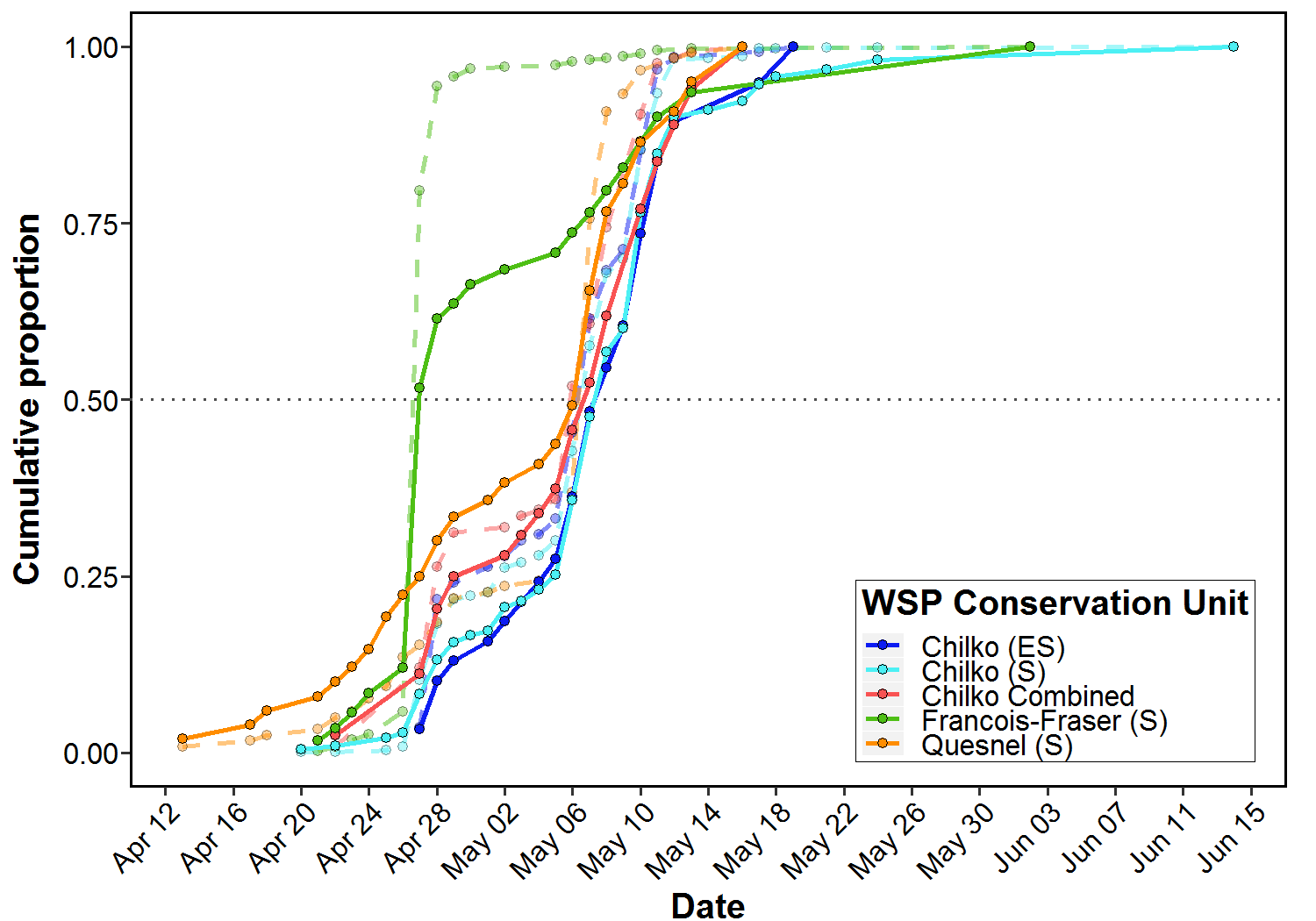
**A**(Method 1)



**B**(Method 1)

**Figure 1.1.** a) Daily count at Chilko (gray line), calculated daily average (blue, increased four-fold) and daily count at Mission (black, increased three-fold) estimates of abundance for Chilko (Chilko (S and ES) and Chilko Combined) yearlings at Mission. b) Cumulative run timing for the same three data series listed previously. All run lengths for the RST only were considered.

As an exercise in comparison, we can compare the cumulative run timing curves for the top 5 Wild Salmon Policy Conservation Units (WSP CUs) between expanded counts and Method 1 abundance estimates (Table 2 and Figure 1.2). Method 1 estimates did not generate very different run timing estimates compared to expanded counts and Method 2 (Table 2), but note that the shape of the curves for Method 1 compared to expanded counts differs considerably (Figures 2.2A and 2.3A).



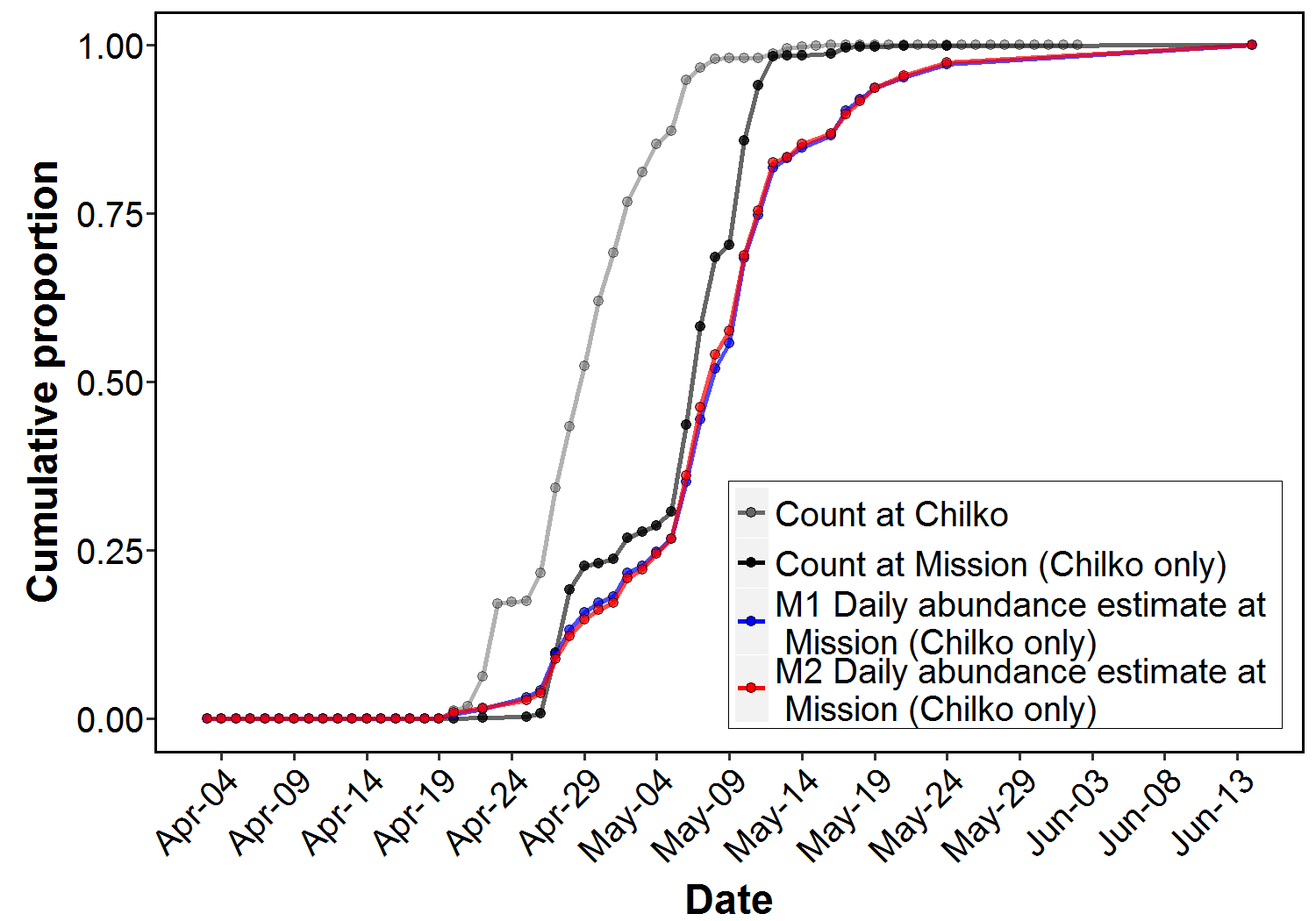
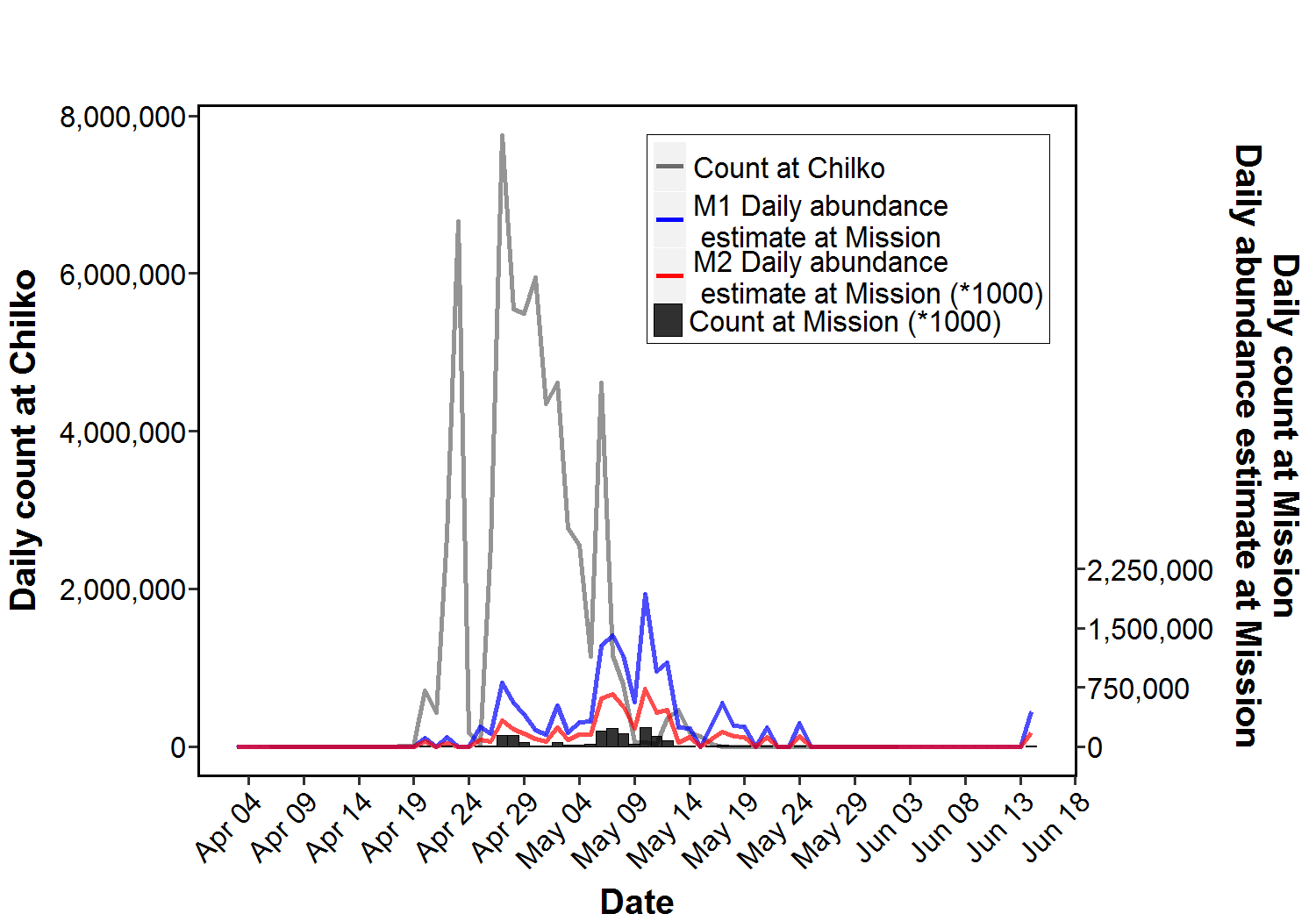
**A**(Method 1)

**B**

**Figure 1.2.** Cumulative migration run timing for the top five most abundant WSP CUs in 2017 using expanded counts (dashed lines) and **Method 1** estimates of abundance (solid line). Gray dotted line indicates 50% migration point. All run lengths for the RST only were considered.

**Method 2. Sampling event current velocity**

While Method 1 accounts for change in daily discharge, it does not allow for finer-scale corrections based on changes in water volume throughout the day (affected by tide) and laterally across the river (Figure A1.1a-c). Therefore, corrections should be made at the sampling event scale, where a sampling event is considered each unique ‘date-run-bay’ combination and the associated surface current velocity. This method uses Equations 1-3 and considers RST catch (expanded for sub-sampling) only.



**B**(Method 2)

**A**(Method 2)

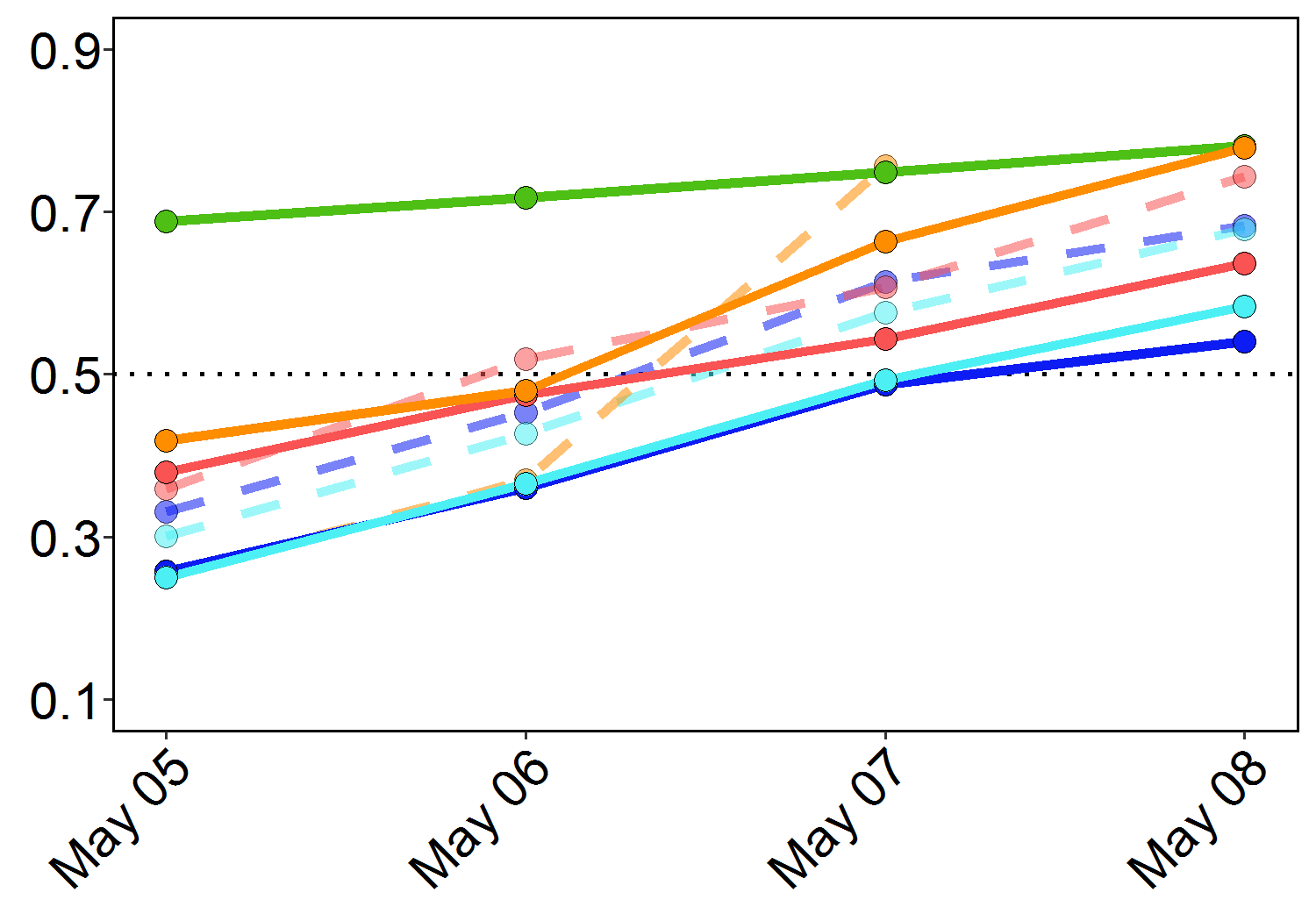
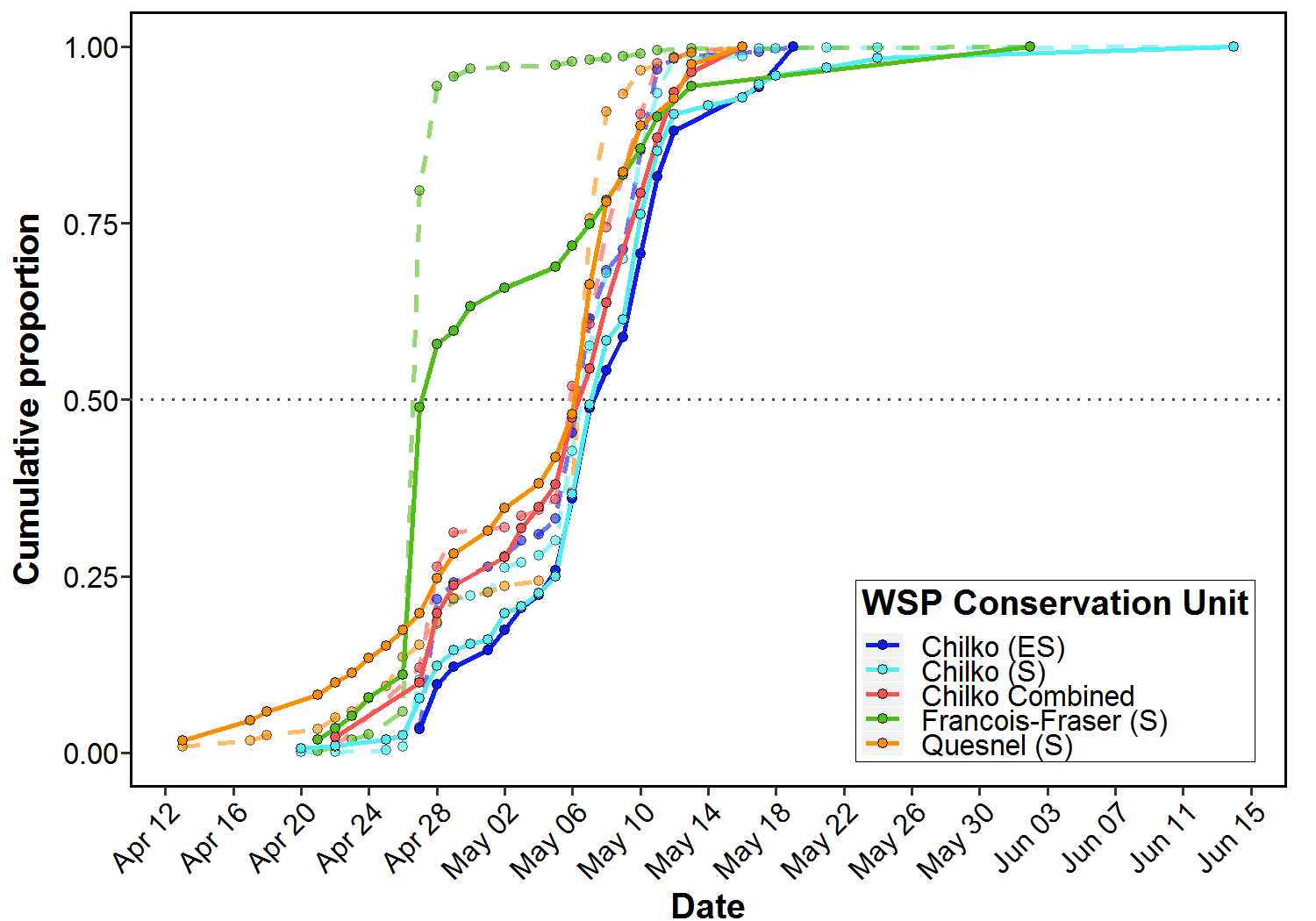
**Figure 2.1.** a) Daily count at Chilko (gray line), calculated daily estimate of abundance (method 1 blue, method 2 red increased four-fold) and daily count at Mission (black, increased four-fold) for Chilko (Chilko S, ES and Chilko Combined) yearlings at Mission. b) Cumulative run timing for the same four data series listed previously. All run lengths for the RST only were considered.

We can also assess the impact of various correction methods on cumulative run timing for top abundance Conservation Units (CUs). Original ‘raw’ catches were expanded for subsampling, and these expanded counts were turned into daily total estimates of abundance for each CU. Expanded counts and estimates of abundance were converted to daily percentages of totals for each CU, and then expressed cumulative proportions of both counts and abundance estimates over time for each CU. Cumulative run timing data for the top 5 most abundant CUs (Chilko (S), Chilko (ES), Chilko Combined, Francois-Fraser (S) and Quesnel (S)) are shown in Figure 2.3 and Table 2). The early ‘hump’ displayed by most CUs around Apr 28-May 6 is less apparent with Method 1, which instead produces a more gradual increase. This is likely reflecting the gradual increase in discharge later in the season.

**Table 2.** Cumulative run timing information for the top five most abundant Wild Salmon Policy Conservation Units (WSP CUs) in 2017. All run lengths for the RST only were considered. Values correspond to Figure 1.2 and 2.2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **WSP CU** | **Method** | **First Observation** | **Last Observation** | **50% Migration Date** | **Peak Migration Date** |
| Chilko (ES) | Expanded counts | Apr 27 | May 19 | May 7 | Apr 28 |
| Method 1 | May 8 | May 10 |
| Method 2 | May 8 | May 7 |
| Chilko (S) | Expanded counts | Apr 20 | Jun 14 | May 7 | May 10 |
| Method 1 | May 8 | May 10 |
| Method 2 | May 8 | May 10 |
| Chilko Combined | Expanded counts | Apr 22 | May 16 | May 6 | May 10 |
| Method 1 | May 7 | May 10 |
| Method 2 | May 7 | May 10 |
| Francois-Fraser (S) | Expanded counts | Apr 21 | Jun 2 | Apr 27 | Apr 27 |
| Method 1 | Apr 27 | Apr 27 |
| Method 2 | Apr 28 | Apr 27 |
| Quesnel (S) | Expanded counts | Apr 13 | May 16 | May 7 | May 7 |
| Method 1 | May 7 | May 7 |
| Method 2 | May 7 | May 7 |

The start, end and peak migration periods do not vary when comparing cumulative expanded counts to cumulative estimates of abundance (Table 2). However, the 50% migration date changes for four of five stocks (Table 2 and Figure 2.3B). Francois-Fraser (S) run timing was almost identical between both methods. For the other four CUs, 50% migration timing was still relatively the same (Table 2), although for all CUs, ‘Method 2’ produced curves that reached 50% slightly later (Figure 2.3B).



**A**(Method 2)

**B**

**Figure 2.2.** Cumulative migration run timing for the top five most abundant WSP CUs in 2017 using expanded counts (dashed lines) and **Method 2** estimates of abundance (solid line). Gray dotted line indicates 50% migration point. All run lengths for the RST only were considered.

**Method 3. Catch tables**

The previous methods do not account for fish abundance in Bays while sampling is not occurring (either when fishing is occurring in another Bay, or over night). Creating catch tables and extrapolating these ‘NA’ values is a more robust way of estimating overall abundance throughout the season (Walters 2003) (Table 3). Eventually, these tables can be made for each population to obtain fine-scale abundance estimates for all WSP Conservation Units of interest.

**Table 3.1.** Example subset of catch table for three sampling events (one ‘cycle’) in all three Bays. Cells are populated with CPUE (fish/m3) per run (Equation 1). NA indicates no fishing occurred during the time interval. ‘USID’ = Unique Sampling ID (date-run-bay-trap-depth).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **USID** | **Interval** | **Bay 2** | **Bay 6** | **Bay 11** |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:12 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:13 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:14 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:15 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:16 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:17 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:18 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:19 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:20 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:21 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:22 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:23 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:24 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:25 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:26 | NA | NA | 0.010184 |
| 20170427-R3-B11-RST-0 | 27/04/2017 8:27 | NA | NA | 0.010184 |
| NA | 27/04/2017 8:28 | NA | NA | NA |
| NA | 27/04/2017 8:29 | NA | NA | NA |
| NA | 27/04/2017 8:30 | NA | NA | NA |
| NA | 27/04/2017 8:31 | NA | NA | NA |
| NA | 27/04/2017 8:32 | NA | NA | NA |
| NA | 27/04/2017 8:33 | NA | NA | NA |
| NA | 27/04/2017 8:34 | NA | NA | NA |
| NA | 27/04/2017 8:35 | NA | NA | NA |
| NA | 27/04/2017 8:36 | NA | NA | NA |
| NA | 27/04/2017 8:37 | NA | NA | NA |
| NA | 27/04/2017 8:38 | NA | NA | NA |
| NA | 27/04/2017 8:39 | NA | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:40 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:41 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:42 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:43 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:44 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:45 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:46 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:47 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:48 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:49 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:50 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:51 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:52 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:53 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:54 | 0.002144 | NA | NA |
| 20170427-R4-B2-RST-0 | 27/04/2017 8:55 | 0.002144 | NA | NA |
| NA | 27/04/2017 8:56 | NA | NA | NA |
| NA | 27/04/2017 8:57 | NA | NA | NA |
| NA | 27/04/2017 8:58 | NA | NA | NA |
| NA | 27/04/2017 8:59 | NA | NA | NA |
| NA | 27/04/2017 9:00 | NA | NA | NA |
| NA | 27/04/2017 9:01 | NA | NA | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:02 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:03 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:04 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:05 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:06 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:07 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:08 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:09 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:10 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:11 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:12 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:13 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:14 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:15 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:16 | NA | 0.060031 | NA |
| 20170427-R5-B6-RST-0 | 27/04/2017 9:17 | NA | 0.060031 | NA |

Filling NA values in the table can be done using several approaches (Walters, 2003):

1. Look forward and backwards in time and perform a trend analysis to infer and extrapolate missing cell values
2. Use spatial statistics from spatially “nearby” cells
3. Use spatial covariates of distribution (e.g., temperature, velocity, daylight) to improve spatial and temporal estimates (essentially fine-tuning #1)
4. Use spatial models of dispersal-migration and local renewal dynamics

Points 1 and 3 are the most realistic places to start with the Mission downstream data. Using spatial statistics (point 2) is likely not applicable due to the small spatial extent of the study system. Point 4 may be of use in the future, particularly if migration dynamics can be incorporated, although local renewal is not a applicable in this case. To note, Walters (2003) outlines that simply taking an average of fished areas and applying this average to empty cells is not appropriate as it makes dangerous assumptions. It assumes that unfished strata behave the same as fished strata (in which case, this exercise is pointless), and also assumes that any given unique unfished strata is equivalent to another unfished strata. Even if segregated by Bay, we would still be assuming that unfished Bay 2 on April 7th is equal to unfished Bay 2 on April 27 (peak of migration).

Moving forward, the most appropriate places to start would be Point 1 (trend analysis) followed by Point 3, and to perform some sensitivity/visual analyses comparing methods. The following sections will outline these filling methods.

Assessing time series structure and autocorrelation prior to infilling

Infilling NA cells cannot be done using simple methods such as linear regression given the time-series nature of the dataset. Using just the base dataset, with known average CPUE per day, we can examine the basic structure of the data prior to deciding upon infilling techniques. Aside from the fact that the data are obviously not normally distributed with equal variance, preliminary analysis also indicates autocorrelation within the data based on summary model fit results from a linear model (*lm*), time-series linear model (*tslm*) and forecasting model (*Arima*) (Table 3.2). Note, although we are not attempting to forecast using these models, fitting an *Arima* model to the data is a useful step to assess whether including the ‘ar’ term improves model fit (indicating autocorrelation in the data). Both *tslm* and *Arima* allow for seasonality and trends in time series data and are therefore expected to provide improved fits compared to *lm*. However, *Arima* does use historical patterns to inform analysis, which means it is best applied to long, stable series.

As indicated by AIC scores in Table 3.2, the Arima model provides the best fit, and the parameter estimate is considerably different from those produced by *lm* and *tslm* models. Collectively, these results indicate autocorrelation within the catch dataset. While autocorrelation is usually implied in most timeseries datasets, particularly those with seasonality or trend components, it is important to determine this before moving forward with infilling methods.

**Table 3.2.** Preliminary model comparisons of three methods to assess autocorrelation within the base dataset. Both tslm() and Arima() use time series data, although based on the structure of Arima() the cpue~date time series is combined into one time series list and is not separately specified as in tslm() (hence the term name “ar1” for the Arima() model, which represents ‘date’).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Term** | | | **AIC** | **ΔAIC** |
| **Intercept** | **date** | **ar1** |
| Arima(timeseries) | 0.00098 | - | 0.5023 | -646.9 | 0 |
| lm(y~x) | 0.1496 | 9.95e-11 | - | -626.5 | 20.4 |
| tslm(y~x) | 0.00130 | -8e-06 | - | -626.5 | 20.4 |

Summarizing missing values

Now that we have an idea of the structure of our base dataset (which will inform infilling), it is useful to conduct an assessment of our catch table (i.e., expanded dataset by 1–minute intervals) and determine the quantity and quality of empty cells to be filled. The *imputeTS* package provides useful summary functions and infilling techniques which will be highlighted in this and the following sections. Table 3.3 outlines summary information regarding the number of NAs for each Bay. Note handling time for *imputeTS* is very slow so the catch has to be analyzed by each Bay separately. This should not be problematic as we are assuming catches in each Bay are independent from each other.

**Table 3.3.** Summary statistics for NA entries in each sampling Bay. Calculated using *imputeTS* package.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Bay 2** | **Bay 6** | **Bay 11** |
| Length of series | 104,756 | 104,756 | 104,756 |
| Number of missing values | 97,546 | 97,505 | 97,506 |
| Difference in missing values | vs. Bay 6: +41 vs. Bay 11: +40 | vs. Bay 2: -41 vs. Bay 11: -1 | vs. Bay 2: -40 vs. Bay 6: +1 |
| Percentage of missing values | 93.12% | 93.08% | 93.08% |
| Longest NA gap1 | 5,658 | 5,659 | 5,661 |
| Most frequent gap size2 | 46  (occurred 45 times) | 47 (occurred 40 times) | 47 (occurred 48 times) |
| Gap size accounting for most NAs | 5,658 (occurred 1 time accounting for 5,658 NAs overall) | 5,659 (occurred 1 time accounting for 5,659 NAs overall) | 5,661 (occurred 1 time accounting for 5,659 NAs overall) |

*1 Series of consecutive NAs  
2 Series of consecutive NA series*

Overall, Bay 2 had 41 more sampling observations than Bay 6 and 40 more than Bay 11 (Bay 6 had the fewest sampling observations). The longest and most frequent NA gaps were fairly similar among Bays. Given these results, for now it appears safe to assume that equal sampling occurred among Bays (the different in sampling effort is relatively small when analyzed at the 1-minute scale). However, in future a sensitivity analysis experiment could be designed where ~40 randomly selected observations are removed from Bay 2 to test whether catch estimates change significantly.

**Method 3.1 Catch tables – Interpolation methods (*imputeTS*)**

The *imputeTS* package contains three key interpolation methods that are applicable to our dataset. Interpolation is done via *approximation*, *spline* and *stine* methods.

* *Look into zoo() more*
* *inputTS seems best*
* *look at plotting mean CPUE/day with inputTS and raw versions – figure out how to get ts into plots – multiseries ggplot() likely best*
* *write up autocorrelation analysis*

**References**

Townsend, M. J., T. E. Cone, S. M. Kalyn, T. R. Whitehouse, C. Neville and J. A. Tadey. 2017. Evaluation of the timing, size, abundance and stock composition of downstream migrating juvenile Sockeye salmon in the lower Fraser River in 2016 (Year 5). A Report to the Pacific Salmon Commission. Fisheries and Oceans Canada, Delta, BC. 52 pp.

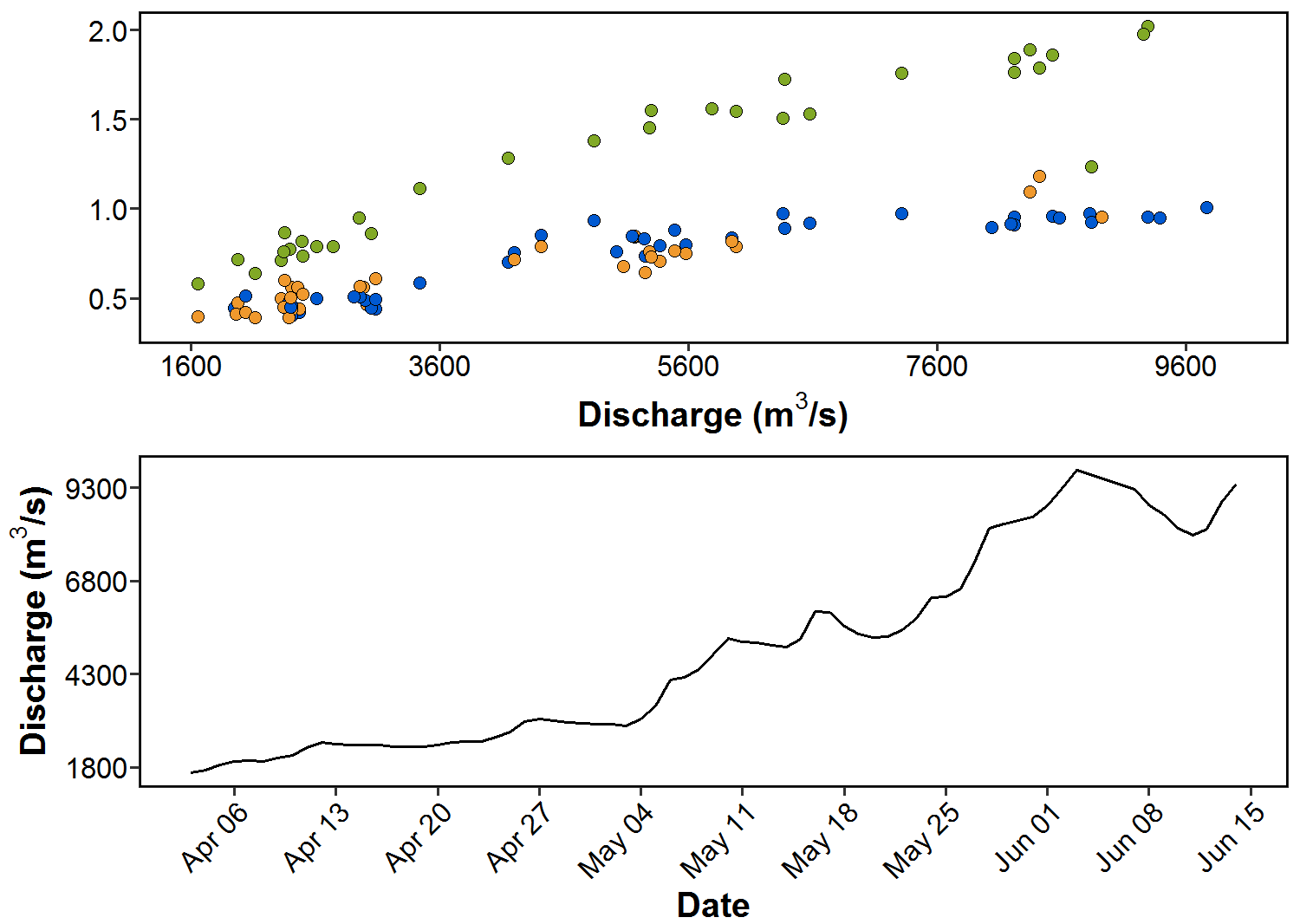
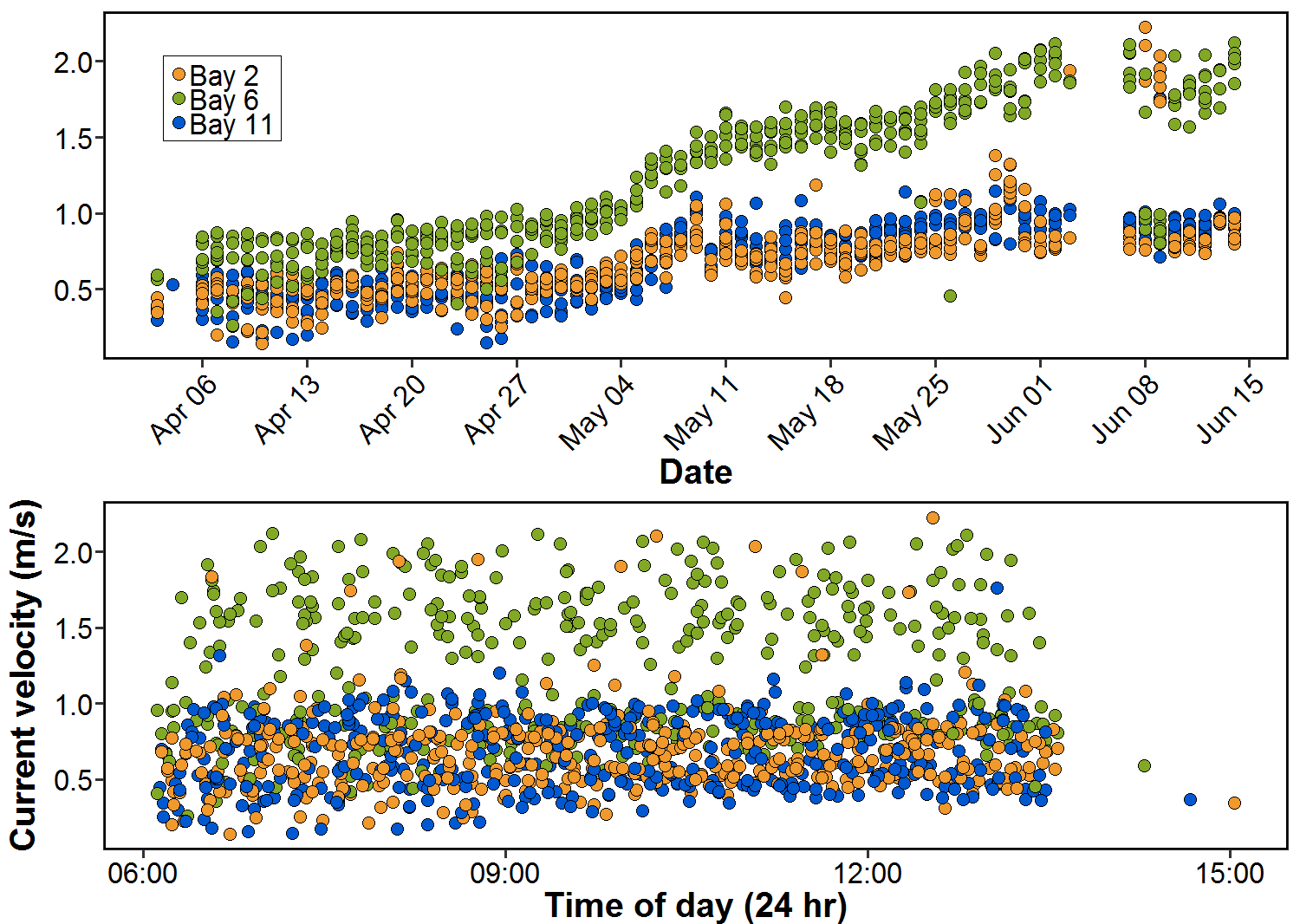
Walters, C. 2003. Folly and fantasy in the analysis of spatial catch rate data. Canadian Journal of Fisheries and Aquatic Sciences. 60: 1433-1436.

**Appendix I: Environmental Data Exploration**

Water flow over time and space

As documented throughout the field seasons, preliminary analysis confirms that water current velocity (m/s) varies significantly across the river (One-way ANOVA, *F* = 323.4, *df* = 2, *p* < 0.001). Current velocity is significantly higher in Bay 6 (Tukey post-hoc, both *p* < 0.001), but there is no significant difference in current velocity between Bays 2 and 11 (Tukey post-hoc, *p* = 0.94). For future analysis using current velocity, Bays 2 and 11 could potentially be grouped for simplicity if needed.

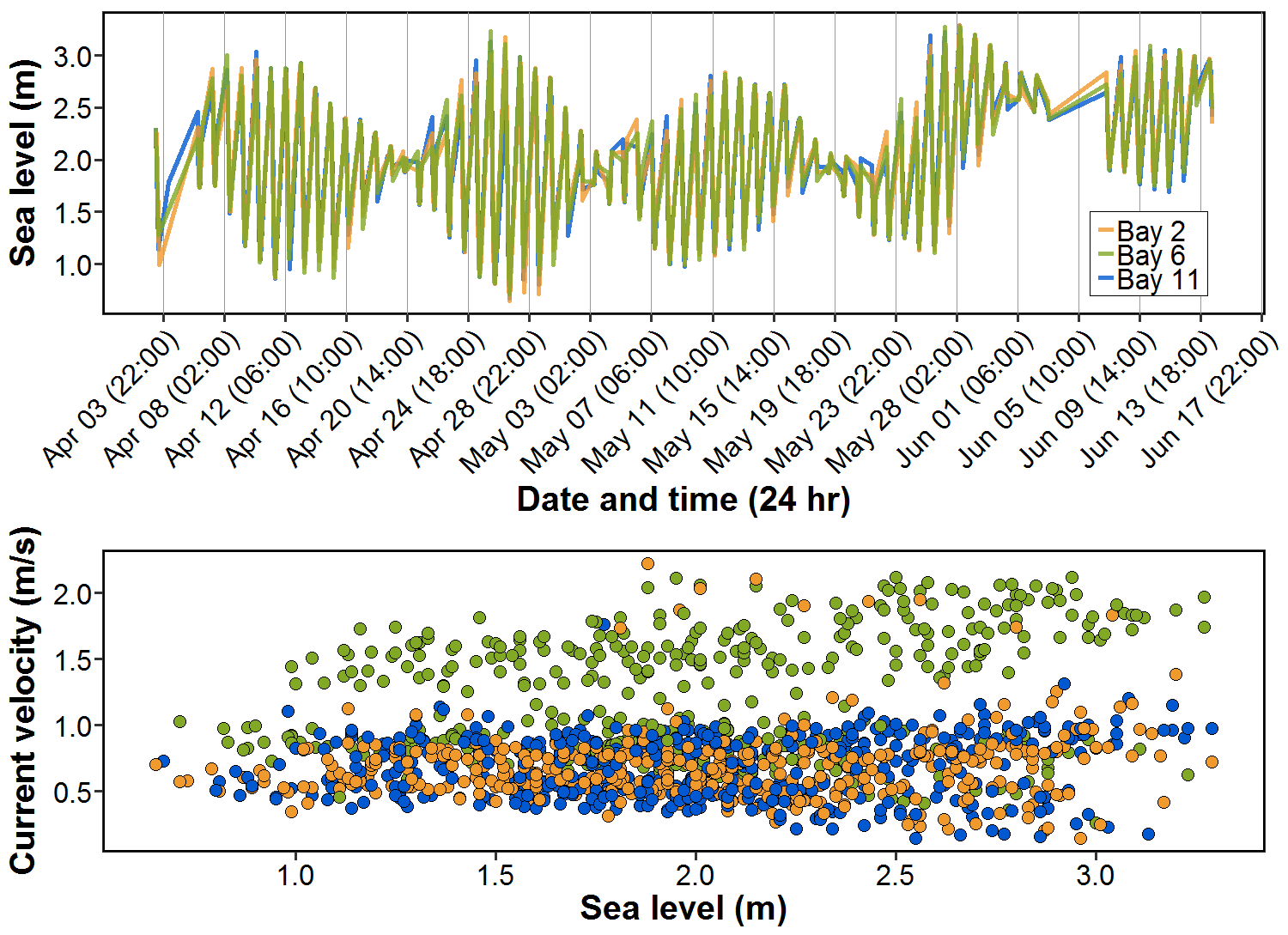
Current velocities also increases over time (intra-annually) with discharge. As the season progresses, the difference between current velocity in the centre (Bay 6) and edges (Bays 2 and 11) increases disproportionately (Figure 1A and C). The daily discharge data was measured at Hope (station #08MF005) by Environment Canada (<https://wateroffice.ec.gc.ca>) (Figure 2 for comparison). For these calculations, discharge data included in the “2017 Mission RST – Main File SOCKEYE” obtained in-season were used as cleaned 2017 data are not yet available.



**Figure A1.1.** Change in current velocity (m/s) a) throughout the year (2017), b) over the course of a sampling day, and c) with daily discharge (m3/s) for each Bay. Discharge at Hope is given for comparison (d). All runs and traps considered.

Relationship with sea level

Bay current velocity will also vary as a function of the date and time of sampling due to changes in tidal currents. Sea level tidal data were obtained from the Fisheries and Oceans Canada Canadian Tides and Water Levels Data Archive (<http://www.isdm-gdsi.gc.ca/isdm-gdsi/twl-mne/index-eng.htm>) for the New Westminster station (#7654), as 2017 stage height data from Environment Canada is not yet available online. Tidal data are supplied at a 1-min resolution, and it is assumed these data represent tidal height at the Mission downstream sampling location. Over the field season, sea level varies fairly consistently by sampling Bay (Figure 2A), and there does not appear to be a relationship between sea level and Bay velocity (Figure 2B).



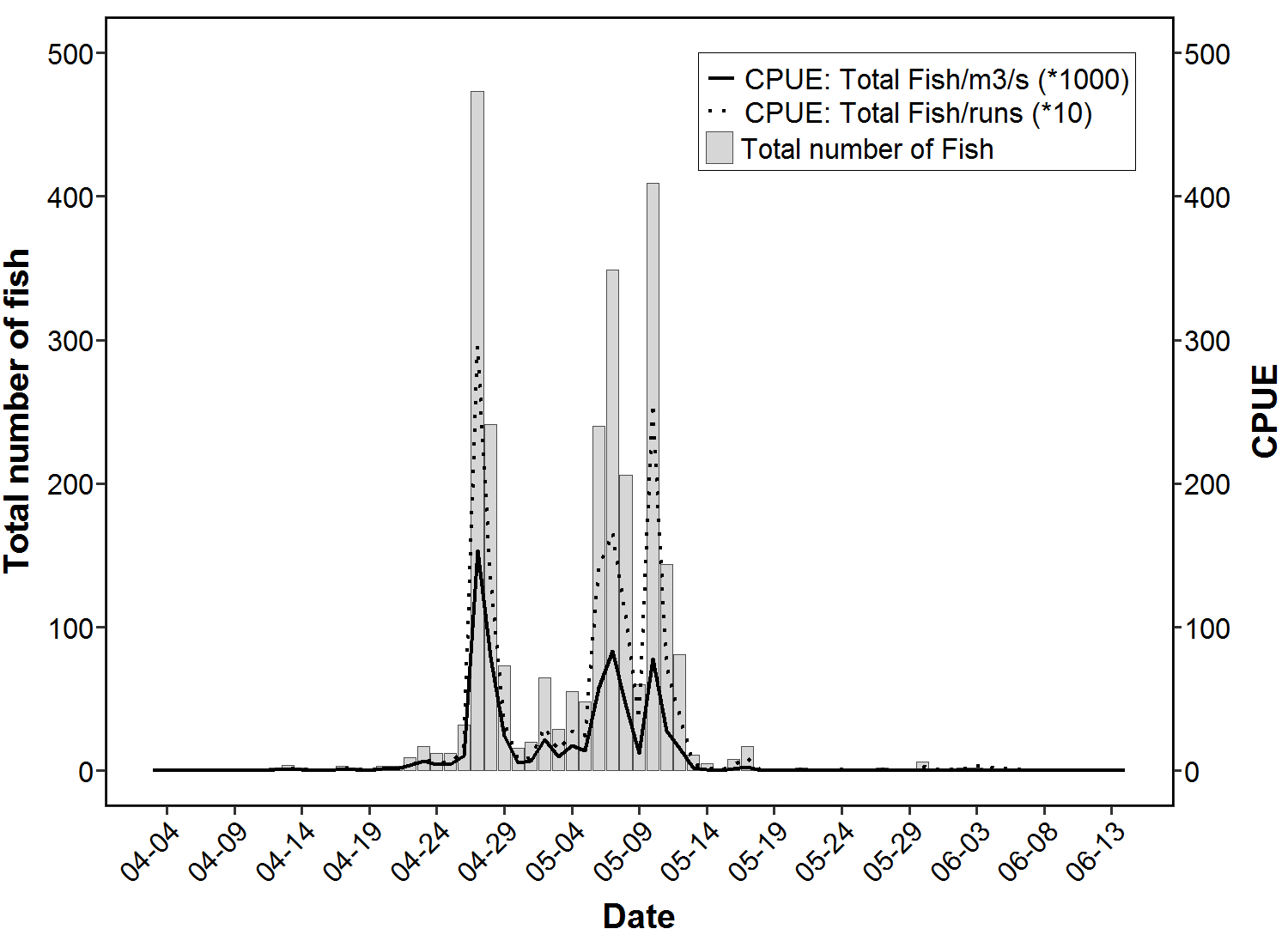
**A**

**B**

**Figure A1.2.** a) Change in sea level (m) over the course of the field season, and b) current velocity (m/s) as a function of sea level (m) for each sampling Bay.

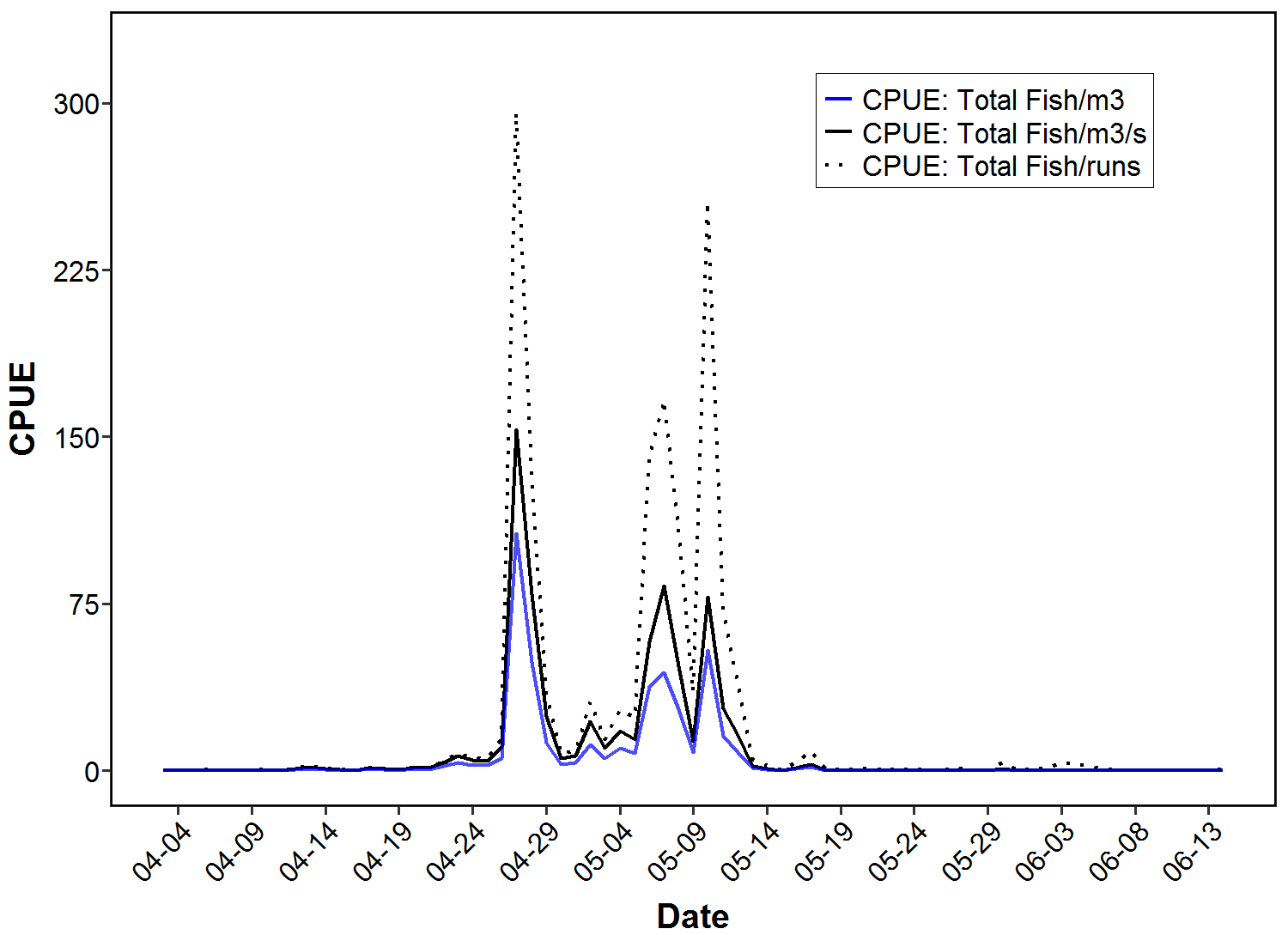
**Appendix II: Early CPUE Calculations**

Starting with a coarse ‘day’ scale, we can divide the total number of sockeye fish by the daily discharge for an estimate of the number of fish/m3•s-1 and compare visually to the previous CPUE calculation (Figure 1).



**Figure 1.** Total number of fish (gray bars), original CPUE increased 1 order of magnitude (fish/run; dotted line) and new CPUE increased 3 orders of magnitude (fish/m3•s-1 water; black line) for each sampling day in 2017. Only 15 minute run times were included. All traps considered.

However, this method of calculating CPUE by discharge does not include the sampling effort associated with the previous method (number of runs per day). By including a second level of calculation, dividing again by the daily total of seconds of sampling, we obtain a CPUE estimate that is calibrated for both river discharge over time, and varied sampling effort (given as seconds of fishing) among days. This also gives a potentially more intuitive unit of effort: daily number of fish per cubic-metre of water (i.e., density of fish; Figure 2). This also allows for the inclusion of all runs (n = 4026), and does not have to reduce the dataset to only runs of 15 minute lengths (n = 3996). While the 15-minute runs represent 99% of the total number of runs in 2017, this could vary among years. However, number of fish/m3 could be misleading as it suggests fish are distributed equally (any given cubic-metre of water is equivalent), when we know fish are typically more concentrated at the surface (1 m3 of water at the surface is not equivalent to 1 m3 at 20 ft when it comes to fish vertical distribution).

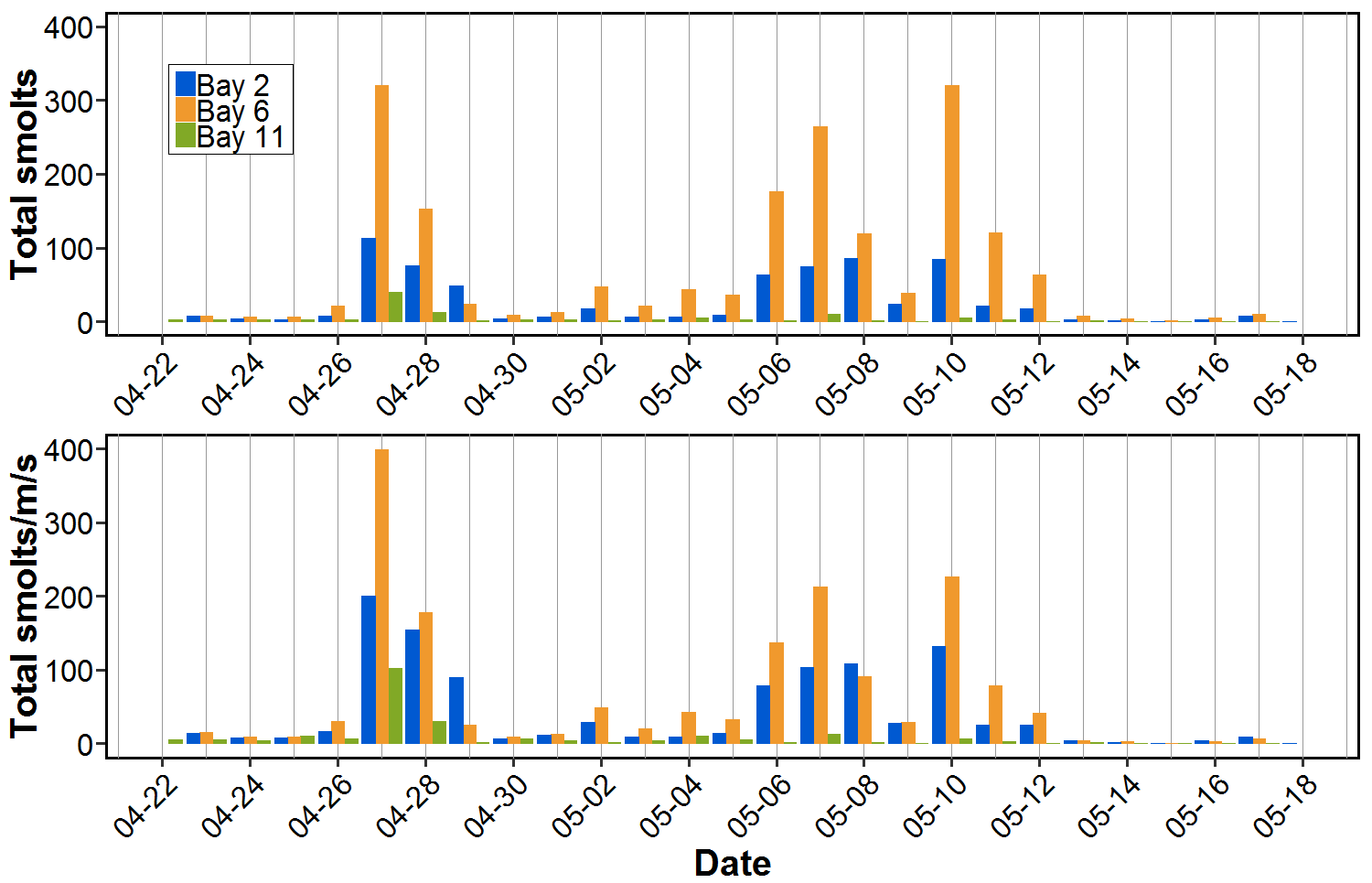


**Figure 2.** Original CPUE increased 1 order of magnitude (fish/run; dotted line), CPUE accounting for daily discharge increased 3 orders of magnitude (fish/m3•s-1 water; black line), and CPUE accounting for daily discharge and sampling time increased 7 orders of magnitude (fish/m3 water; blue line) for each sampling day in 2017. All traps and run times were considered.

Although it is helpful to have one value for total catch, corrected for discharge and sampling time (seconds), the result is very small numbers that have to be scaled up several orders of magnitude to give meaningful CPUE estimates (see Figure 2 figure caption). It may be possible to convert discharge into a relatively smaller number (m3/min, or m3/hr) and calibrate by total sampling time (minutes or hours, respectively), but this moves away from the standard m3/s discharge measure which may not be ideal.

CPUE corrected for velocity

Given the disproportionate change in Bay current velocity (due to discharge) over time, CPUE corrections using flow should be done on a ‘date-run’ scale to incorporate these changes in velocity over time. Dividing run catch by current velocity gives an estimate of number of fish/m⋅s-1, but alters the abundance estimates, particularly among Bays (e.g., note catch differences among Bays on April 27 and 28; Figure 4A and B). However, the index of ‘number of fish/m⋅s-1’ is not as intuitive as the previous index of ‘number of fish/m3⋅s-1’ or ‘number of fish/m3’.



**A**

**B**

**Figure 4.** a) Total ‘raw’ catch (number of fish) in each bay and b) Total fish calibrated for average daily Bay current velocity (number of fish/m•s-1). All runs and traps considered.

Considering Bay Differences When Calculating CPUE

Accounting for spatial and temporal changes in water volume and current is important when considering the relative fishing effort throughout the field season. Therefore, it would seem most effective to obtain an estimate for the cross-sectional area of the river in order to relate daily discharge to current velocity at each sampling event. Area could be acquired in two ways.

First, if depth data are available, we can calculate the cross-sectional area using width (possibly obtained from Google Earth?) and average depth of the river (assuming width does not change considerably over time). Resolution of cross-sectional area measurements would depend upon the resolution of depth data available. Using those area estimates for the entire river cross-section, we can calculate the approximate area of each sampling Bay (river is assumed to be rectangular as in Townsend et al. 2017). Bay area can be multiplied by Bay current at each sampling event for an estimate of Bay discharge at each sampling event.

A second, more derived method could be constructed if depth data are not available. Using average daily discharge and average daily current velocity (average of all Bay current velocities in a given day), we can calculate the average daily cross-sectional area. Using average daily area, we can then calculate the average daily area for each Bay, and use this to then obtain average daily Bay discharge. However, this method is more derived and potentially problematic, as it relies on derived, rather than measured, data.

One final consideration is that discharge at Hope is quite removed from discharge at Mission. Using Hope discharge and Mission cross-sectional area may not be appropriate. Ideally, area information could be available from both Mission and Hope in order to conduct a preliminary sensitivity analysis before proceeding (perhaps river area is approximately the same at both stations).

Should this method be of interest, future steps would include:

* Obtaining cross-sectional area information from:
  + Measured river depth estimates (depth sounder on boat, marine charts, EC measurements)
  + Measured river width (Google Earth, marine charts, EC)
  + Derived calculations from collected data
* Investigating whether cross-sectional area is available for both Mission and Hope
  + Comparing these areas to ensure that relating Hope discharge to Mission cross-sectional area is not problematic
* Investigating if a derived ‘daily average area’ from measured data is reasonable
* If both methods are do-able, compare both methods for estimating cross-sectional area (and subsequently Bay discharge) within the 2017 dataset
  + This can serve as a baseline for other sampling years where some information may not be available

**Method A1. RST, Depth=1.13, Not expanded for subsampling**

Elaborating on the previous idea of using Bay cross-sectional area to calculate volume of water per Bay, it is also important to account for ‘water volume fished’ of each trap type (RST, IPT, VT) when comparing catch among years.

To account for both ‘water volume fished’ and Bay cross-sectional area, ‘raw’ catch per run per day (RST only) was divided by the total RST water volume fished for each run to give an estimate of CPUE (number of fish/m3). For the purposes of this exercise, Bay cross-sectional area was assumed to be equal among Bays, and was calculated using the RST fishing depth (1.13 m) and Bay width. Based on Google Earth measurements, the river width at the PSC sonar station is approximately 440 m; for the purposes of the following calculations we assume each Bay is equal width (~ 146m).

Bay cross-sectional area was multiplied by the current velocity measured at each sampling event to obtain an estimate of discharge (m3/s) for each Bay per second. The number of fish/m3, discharge in Bay per second (m3/s), and run time (sec) were multiplied together for an index of abundance (IA). A sampling event is considered each unique date-time-run-bay combination. An example calculation is given below:

1. CPUE (fish/m3) during one sampling run

= = = 0.027 fish/m3 (in a 900 s run)

Where *catch* is the number of fish caught in the RST trap per run, *A* is the cross-sectional area of the trap opening in the water, *f* is the average flow speed travelled while fishing, and *t* is the run time (usually 900 seconds, but varies from 600 to 1320 seconds).

1. Water volume in Bay during a sampling run

= Bay cross-sectional area **⋅** Current velocity **⋅** Run time  
= (depth **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= (1.13 m **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= 74, 580 m3 (in a 900 s run)

Where the *river width* is 440 m based on measurements taken in Google Earth at the Pacific Salmon Commission sonar array. *Current velocity* here is an arbitrary value and will vary based on run.

1. Index of abundance (IA)

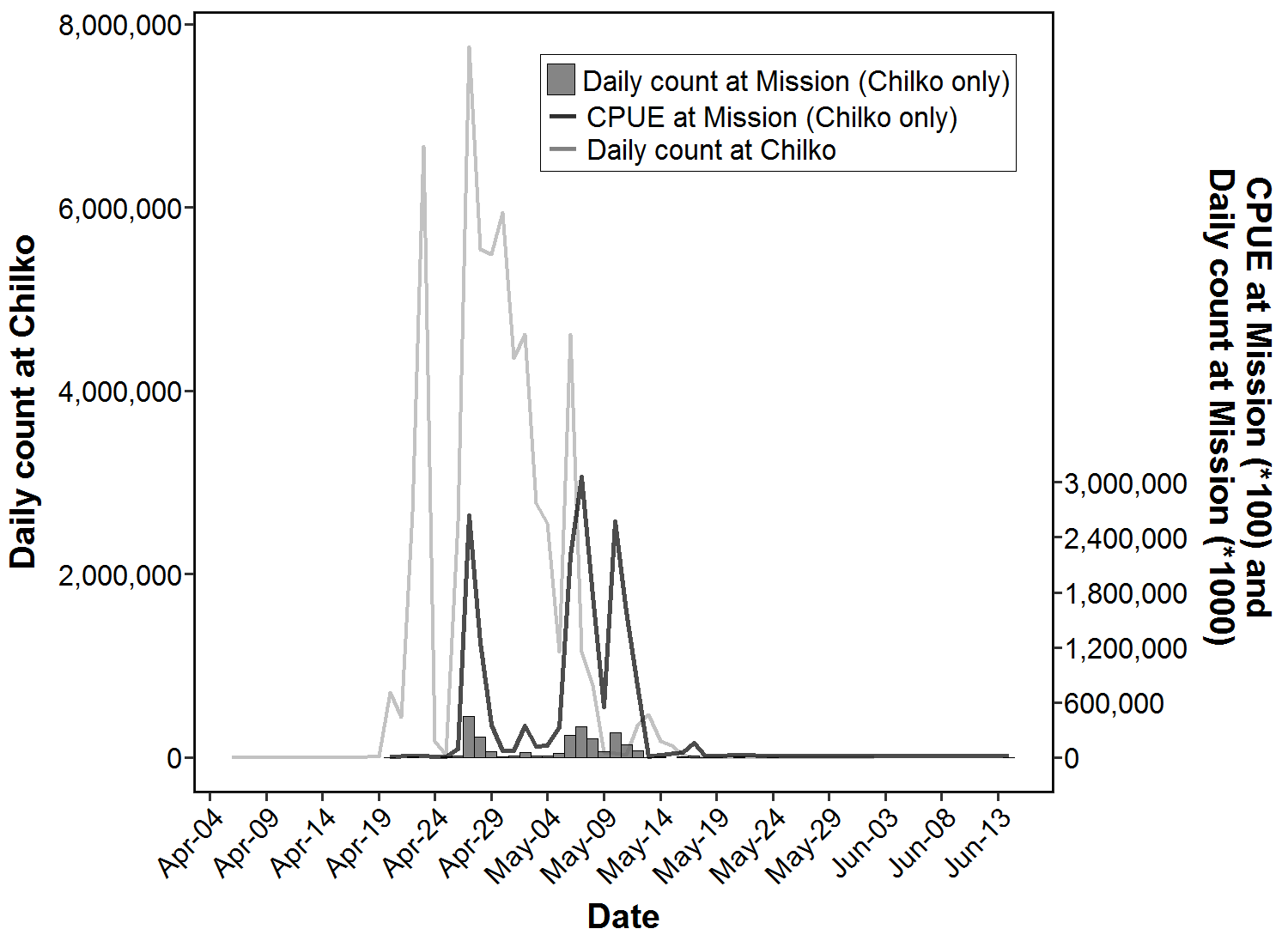
= CPUE **⋅** Water volume in Bay during a sampling run (m3/s)   
= 0.027 **⋅** 74580 m3  
= 2013.66 fish

Where *Total run time* will usually be 900 seconds, but may range from 600 to 1320 seconds.

In this example, only 50 fish were sampled when there were potentially over 2000 fish passing through that Bay at that sampling event.

Confirmation With Chilko Fence

To confirm that this IA calculation does not produce unrealistic estimates, we can compare the IA of Chilko fish at Mission to the daily count at Chilko (Figure 3).



**Figure 3.** Daily count at Chilko (dotted line), daily count at Mission (bars), and calculated IA (black line) for Chilko yearlings. Chilko yearlings included Chilko (S), Chilko (ES) and Chilko Combination CUs. All run lengths for the RST only were considered.

Based on visual assessment, these CPUE estimates appear reasonable in comparison to the Chilko fence counts. However, considering the Bay cross-sectional area to extend only to the depth of the RST trap may be problematic when comparing among trap types. As well, this CPUE adjustment only accounts for sampled fish and is not applied to the sub-sampling expansions done for each CU (see Townsend et al 2017).

**A3. RST and VT, Depth=1.13 and 1.01m, expanded for subsampling**

A similar method involves applying the same calculation as Method 1, but for both the RST and VT. In each calculation, the Bay depth is the trap depth (RST = 1.13, VT = 1.01). The RST calculation is the same as previously. This calculates an area based on the ‘minimum’ fishing volume, but does not include possible fish abundance outside that ‘slice’ of trap area.

1. CPUE (fish/m3) during one sampling run

RST: = = = 0.027 fish/m3 (in a 900 s run)

VT: = = = 0.012 fish/m3 (in a 900 s run)

Where *catch* is the number of fish caught in the RST trap per run, *A* is the cross-sectional area of the trap opening in the water, *f* is the average flow speed travelled while fishing, and *t* is the run time (usually 900 seconds, but varies from 600 to 1320 seconds). VT calculations will be repeated for each catch at depth.

1. Volume of water in Bay during one sampling run

RST:

= Bay cross-sectional area **⋅** Current velocity **⋅** Run time  
= (depth **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= (1.13 m **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= 74, 580 m3 (in a 900 s run)

VT (all depths):

= Bay cross-sectional area **⋅** Current velocity **⋅** Run time  
= (depth **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= (1.01 m **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= 66, 660 m3 (in a 900 s run)

1. IA for each trap

RST:

= Number of fish/m3 **⋅** Water volume in Bay during a sampling run (m3/s)   
= 0.027 **⋅** 74580 m3  
= 2013.66 fish

VT-*x* (repeated for each catch at depth):

= CPUE **⋅** Water volume in Bay during a sampling run (m3/s)   
= 0.012 **⋅** 66660 m3  
= 799.92 fish

However, this method assumes surface current is representative of current at depth.

**A4. RST and VT, Depth=1.13 and 1.01, using current proportions (Vernon 1966), expanded for subsampling**

Using the same method as #3, we can extrapolate current at depth from the vertical current profile in Vernon (Figure 9 and Table 12; 1966). The current at 6 ft and 12 ft is approximately 98% and 93.5%, respectively, of the surface current (top 2 ft). Therefore, we can scale our daily surface current readings to infer current at fishing depths of 6 ft (1.8 m) and 12 ft (3.7 m). This only changes Step 2 of the above calculation.

2) Volume of water in Bay during one sampling run

VT-0:

= Bay cross-sectional area **⋅** Current velocity at surface **⋅** Run time  
= (depth **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= (1.01 m **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= 66, 660 m3 (in a 900 s run)

VT-6:

= Bay cross-sectional area **⋅** 98% of current velocity at surface **⋅** Run time  
= (depth **⋅** ) **⋅** 0.49 m/s **⋅** 900 s  
= (1.01 m **⋅** ) **⋅** 0.49 m/s **⋅** 900 s  
= 65, 326.8 m3 (in a 900 s run)

VT-12:

= Bay cross-sectional area **⋅** 93.5% of current velocity at surface **⋅** Run time  
= (depth **⋅** ) **⋅** 0.4675 m/s **⋅** 900 s  
= (1.01 m **⋅** ) **⋅** 0.4675 m/s **⋅** 900 s  
= 62, 327.1 m3 (in a 900 s run)

This method assumes that current profiles measured at Mission in March-April 1966 represent present-day current profiles. In 1966, surface current from April 2 – June 2 ranged from 0.32 to 1.22 m/s. In 2017, surface current from April 3 – June 14 ranged from 0.14 to 2.23 m/s.

**A5. RST and VT, Depth=1.13 and 1.01m, using current proportions (Vernon 1966), RST expansion factor, expanded for subsampling**

As mentioned previously, the RST is the ‘gold standard’ by which to convert the VT (and eventually, the IPT). The simplest way of doing this is to determine the factor by which the RST CPUE is higher than the VT. This requires paired sampling events where the RST and VT-0 fish concurrently, and both traps are successful at catching fish. Unfortunately, only n = 7 of these sampling events occurred, and in those 7 events, the surface VT trap only caught 1 or 2 fish (Table 1).

**Table 1.** CPUE (fish/m3) for sampling events where the RST and surface VT both caught fish, and the factor difference in CPUE between traps. All runs considered.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sampling event (date-run-bay)** | **RST CPUE** | **VT-0 CPUE** | **Factor difference** |
| 20170422-R7-B11 | 0.000536 | 0.00122 | 0.44 |
| 20170427-R1-B2 | 0.0118 | 0.00122 | 9.66 |
| 20170427-R2-B6 | 0.0509 | 0.00122 | 41.70 |
| 20170501-R6-B6 | 0.000536 | 0.00122 | 0.44 |
| 20170507-R4-B6 | 0.0327 | 0.00244 | 13.40 |
| 20170507-R9-B2 | 0.00750 | 0.00122 | 6.15 |
| 20170508-R6-B6 | 0.0257 | 0.00244 | 10.50 |

There is considerable variation in the CPUE difference between trap types. On average the RST caught 11.8 (± 14.1) times more fish than the surface VT (call this the ‘RST factor’). Using this as a preliminary indication of the efficiency of the RST trap, we can multiply CPUE at 1.8 and 3.7 m by 11.8 to estimate ‘imagined’ RST catch at 1.8 and 3.7 m. From there, the Bay volume can be calculated. As we are implying it is the RST fishing at all depths, we use the RST depth to calculate Bay volume.

1. CPUE (fish/m3) during one sampling run (converted to ‘imagined’ RST-1.8)

VT: = = = 0.018 fish/m3 (in a 900 s run)

VT-1.8 CPUE **⋅** RST factor  
= 0.018 fish/m3 **⋅** 11.8  
= 0.2124 fish/m3

Where *catch* is the number of fish caught in the VT-1.8 trap per run, *A* is the cross-sectional area of the trap opening in the water, *f* is the average flow speed travelled while fishing, and *t* is the run time (usually 900 seconds, but varies from 600 to 1320 seconds). *RST factor* is always 11.8.

1. Volume of water in Bay during one sampling run of the ‘imagined’ RST-1.8

‘Imagined’ RST-1.8:  
= Bay cross-sectional area **⋅** Current velocity **⋅** Run time  
= (depth **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= (1.13 m **⋅** ) **⋅** 0.5 m/s **⋅** 900 s  
= 74, 580 m3 (in a 900 s run)

1. IA for ‘imagined’ RST-1.8

‘Imagined’ RST-1.8:  
= Number of fish/m3 **⋅** Water volume in Bay during a sampling run (m3/s)   
= 0.2124 **⋅** 74580 m3  
= 15840.8 fish

**Comparing Methods A3, A4 and A5**

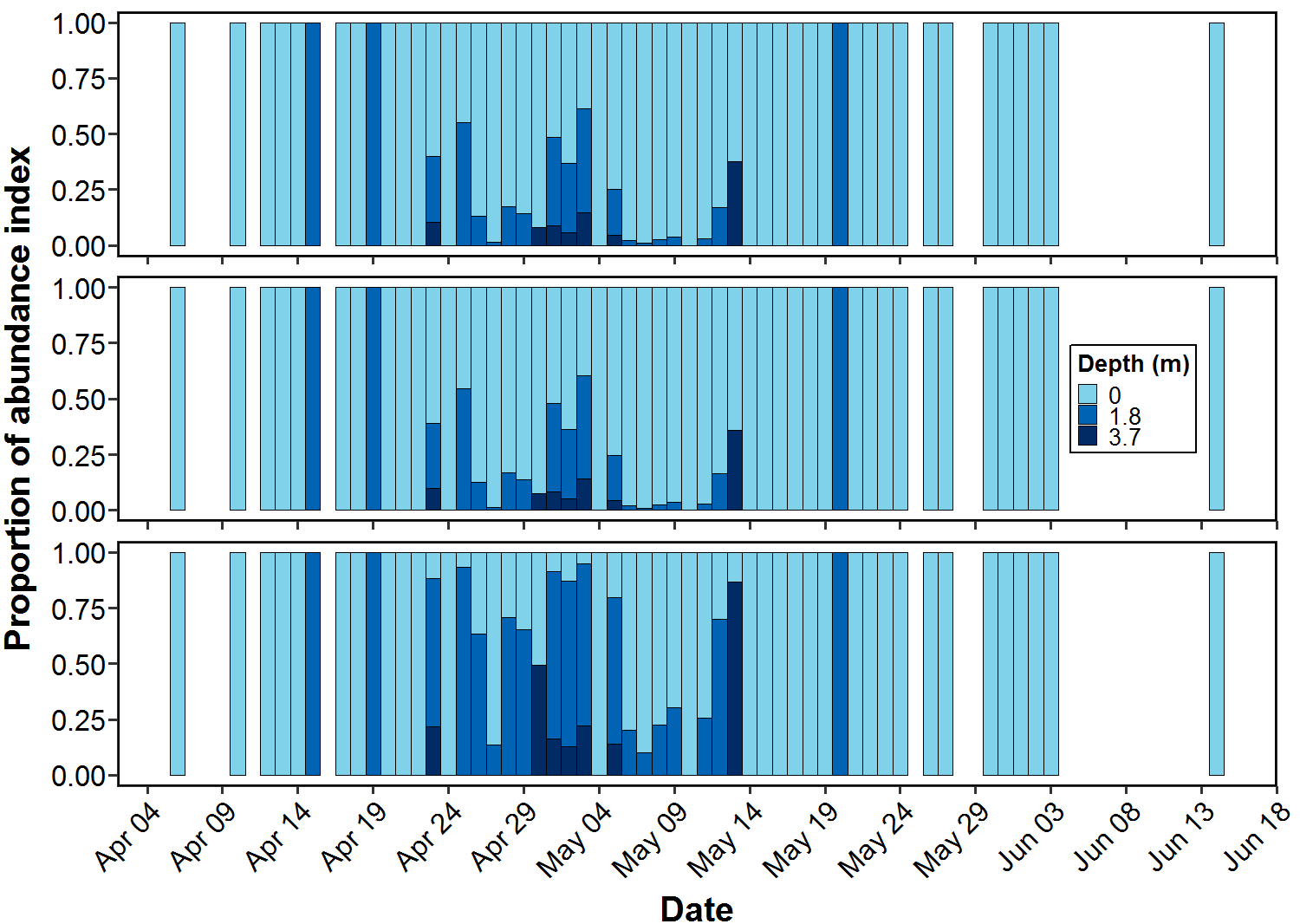
At this point, all methods have the following assumptions/caveats:

* The average flow speed (*f* ) of 0.99 m/s is representative of all runs (x̄ ± SD: 0.99 ± 0.09 m/s)
* The river width does not change significantly intra- or inter-annually, and all Bays are equal width
* The Bay cross-sectional area is only represented by the depth of the trap (estimates are not yet expanded for un-sampled Bay depths)
* Fish distribution is homogenous throughout the Bay cross-sectional area, and there is an equal chance of catching a fish at any given moment and/or location within the Bay
* It is assumed that the RST operates independently from the surface VT when both are fishing simultaneously (see ‘Effort Summary’ below)
* Current velocity will capture both fine- and coarse-scale changes in water flow, for example:
  + Discharge intra- (Figure 1A and C) and inter-annually,
  + Tidal height within and among days,
  + Differences in discharge among Bays (Figure 1C)
* When only surface current velocity is used, we assume it represents current at depth. When the surface currently velocity was scaled to account for depth, we assume that profiles in Vernon (1966) represent present-day current profiles
* All methods rely on current velocity, which may not be available in other years. Within 2017, n=135 runs did not have current values

When comparing the total number of fish caught at each depth, we see a slight overall decrease for IA estimates at 1.8 and 3.7 m (Table 1). However, IA estimates changed by such small amounts that there was no visible difference when summarized as daily totals or by proportion of daily total at each depth (Example shown in Table 1 and Figure 5A and B). In contrast, applying the factor difference between VT and RST traps changes IA proportions considerably at depth (Figure 5C).

**Table 2.** IA for each sampling depth using surface current compared to the proportional decrease in surface current (based on Vernon, 1966). An example of the change in CPUE between these methods is given for April 23. All run lengths for RST and VT were considered. Calculations correspond to Figure 5B.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Depth (m)** | **IA surface current** | **IA proportional current** | **Example: April 23 with surface current** | **Example: April 23 with proportional current** |
| 0 | 193, 191 | 193, 191 | 555.6 | 555.6 |
| 1.8 | 13, 098 | 12, 836 | 274.6 | 269.2 |
| 3.7 | 1, 781 | 1, 666 | 94.68 | 88.4 |



**A**

**B**

**C**

**Method A3 parameters:**   
- RST and VT trap depths (1.13 and 1.01 m)  
- Surface current applied to all depths

**Method A4 parameters:**   
- RST and VT trap depths (1.13 and 1.01 m)  
- 98% surface current for 1.8 m  
- 93.5% surface current for 3.7 m

**Method A5 parameters:**   
- RST and VT trap depths (1.13 and 1.01 m)  
- 98% surface current for 1.8 m  
- 93.5% surface current for 3.7 m  
- VT fish/m3 scaled to ‘imaged RST fish/m3’ at 1.8 m and 3.7 m by factor of 11.8

**Figure 5.** Proportion of abundance index at each sampled depth (0, 1.8 and 3.7 m) using the methods listed above: a) Method A3, b) Method A4, and c) Method A5. All run lengths for VT and RST considered.

Effort Summary

However, none of these methods account for unequal sampling effort at the surface compared to at depth. Assuming the RST and VT-0 operate independently, this further increases effort at the surface compared to at depth. Therefore, when the RST and VT-0 are operating concurrently this is considered 2 sampling events. The total volume of water fished (which accounts for each run time) at the surface was approximately 7 times higher than at 1.8 and 3.7 m (Table 3).

**Table 3.** Total catch, number of sampling events and the resulting average catch for each depth interval. All run times for the RST and VT were considered.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Depth (m)** | **Total catch\*** | **Number of sampling events**† | **Total length of sampling events (hrs)**† | **Total volume fished (m3)** |
| 0 | 2404 | 2008 | 599.0 | 3 783 104.5 |
| 1.8 | 90 | 665 | 170.3 | 557 684.5 |
| 3.7 | 11 | 652 | 164.0 | 537 266.1 |

\* *Surface effort will be even higher if the IPT is considered* † *Surface catch reduces to n = 10 when the RST is excluded*