Documentation of Fraser River sockeye spawner and recruit data

Production data summary format and production data detailed format

E. Taylor

2022

Contents

1. Introduction

Prior to 1985, the International Pacific Salmon Fisheries Commission (IPSFC) was responsible for the management and assessment of Fraser River sockeye. This included assessment of recruits (return abundance) as well as spawning ground enumerations. In 1985, following the signing of the Pacific Salmon Treaty (PST) responsibility for the enumeration of Fraser River sockeye on the spawning grounds was assumed by Fisheries and Oceans Canada (DFO) while the Pacific Salmon Commission (PSC) continued to assess Fraser River sockeye returns. This split responsibility means that datasets or analyses that rely on inputs from these two distinct biological assessment programs require strong collaboration and effective communication between the two organizations. The stock-recruit dataset, or ***Production Dataset***, is possibly the best example of this kind of dataset.

The Fraser sockeye stock-recruit dataset incorporates core productivity data such as spawner abundances, sex ratios, and spawning success from spawning ground enumeration surveys and juvenile abundances from outmigrant assessment with run size data generated by the PSC's in-season assessment programs; estimates of age-specific return abundances (herein referred to as recruits). This dataset is aggregated into groupings that can be referred to as forecast stock groups or production stock group (herein referred to as production stock groups). This grouping structure consists of 19 major stocks and 6 miscellaneous stocks (see 2ii Stock Groupings).

The ***Production Dataset*** documented here not only incorporates the core data outlined above, but also includes additional data that may be relevant to the interpretation and analysis of these data. The reorganisation of this dataset also aims to streamline the process of generating this important dataset, and ensuring reproducibility, consistency, and accessibility. The resulting data stored on the PSC SharePoint site together with this documentation will be made available to technical experts that contribute to the assessment and management of Fraser River sockeye stocks under the PST.

All data relating to spawner and juvenile assessment is provided by DFO staff within the ***Production Spawner Datafile*** while all data relating to recruit assessment is provided by PSC staff within the ***Production Recruit Datafile***. The process for maintaining and combining these two datafiles is provided below (iii Combined Spawner/Recruit Data), and results in two different versions of the Production Dataset: a summary version and a detailed version. The summary version is similar to the production file that has been distributed in the past and organises the production stock information by brood year. In the detailed version, each row relates to a unique salmon cohort that has the same brood year and the same age when returning to the Fraser River. This format allows providing additional metadata across brood and return years and allows for easy manipulations in Excel or R. The detailed version also includes more recent data (e.g. in-season-based run size estimates, recruit data for younger recruits only) that have been excluded from the summary version due to the lack of metadata to correctly interpret and use these data.

2. Data Definitions

i. Population Aggregation

The population groupings aggregation used in the ***Production Dataset*** relates to the forecast groups, and in this document will be referred to as production groups. Using this grouping structure, the Fraser sockeye complex is made up of 19 major stocks and 7 miscellaneous stocks. A list of these stocks can be found in Table 1 and Table 2. Major stocks are those that have a long timeseries of stock-recruit data and are forecast using models of long-term survival and productivity. Miscellaneous stocks are those that do not have sufficiently reliable stock-recruit data to generate forecasts using stock-recruit models or for which the time series of recruits is very limited. Forecasts for these stocks are in general, based on brood

year escapements and/or long-term observed productivity of proxy stocks. The Chilliwack stock is a recent exception. Although being designated as a miscellaneous stock, this stock has been forecast using a stock-recruit relationship in recent years (DFO 2021).

Production groups differ from other population groupings associated with the Fraser sockeye complex. Table 1 & Table 2 demonstrate the alignment between production groups, conservation units (CUs), and management units (MUs). Although there is some agreement between production groups and conservation units, they show some noticeable differences in population membership for some.

Table 1. Translation of Early Stuart and Early Summer run populations found in the Fraser Sockeye Production Files. See Table 2 for Summer and Late run populations. The number in parentheses in column headings refers to the total number of groups within that population classification. Catch and Racial (C&R) groupings are important for in-season assessments due to similarities in timing, stock identification, and migratory routes.

Baseline (50)	Production Stock Code (27)	Production Group (27)	Conservation Unit	C&R (19)	Management Group (4)
Driftwood-Narrows	1	Early Stuart	Takla Trembleur - Early Stuart	EStu	Early Stuart
Bivouac-Rossette					
Dust-Sinta					
Paula-Felix					
Chilliw_lake	23	Misc (Chilliwack)	Chilliwack - Early Summer	Chwk	Early Summer
DollyVarden_Cr					
Chilliwack_Kok					
Alouette_Kok	18	Pitt	Alouette - Early Summer	PiAC	
Coquitlam_Kok					
Pitt_River			Pitt - Early Summer		
Nadina____	17	Nadina	Fraser - Early Summer	NdBo	
			Nadina-Francois - Early Summer		
			Francois (First Run) Early Summer (extirpated?)		
Bowron_____	4	Bowron	Bowron - Early Summer		
			Indian/Kruger - Early Summer (extirpated?)		
			Upper Fraser - Summer		
Nahatlatch	24	Misc (Nahatlatch)	Nahatlatch - Early Summer	GaNh	
Gates_Creek	16	Gates	Anderson Seton - Early Summer		
			Middle Fraser - Summer		
Taseko	22	Misc (Taseko)	Taseko - Early Summer	Tsko	
Upper_Barriere	14	Upper Barriere (Fennell)	North Barriere - Early Summer	NBar	
Upper_Adams	21	Misc (EShu)	Shuswap - Early Summer	ESTh	
Cayenne____					
Eagle_____					
Seymour___	8	Seymour			
Scotch____	15	Scotch			

Table 2. Translation of Summer and Late run populations found in the Fraser Sockeye Production Files. See Table 1 for Early Stuart and Early Summer run populations. The number in parentheses in column headings refers to the total number of groups

within that population classification. Catch and Racial (C&R) groupings are important for in-season assessments due to similarities in timing, stock identification, and migratory routes.

Baseline (50)	Production Stock Code (27)	Production Group (27)	Conservation Unit	C&R (19)	Management Group (4)
Harrison__	19	Harrison	Harrison River - Summer	Harr	Summer
			Kawkawa - Late		
WidgeonSlough	27	Misc (Widgeon)	Widgeon - Summer	Widg	
Middle_River	2	Late Stuart	Takla Trembleur - Summer	LStu	
Kuzkwa_Creek					
Pinchi_Creek					
Tachie____					
Stellako__	3	Stellako	Francois-Fraser - Summer	Stel	
			Fraser - Early Summer		
			Upper Fraser - Summer		
Chilko_____	7	Chilko	Chilko - Summer	Chil	
Chilko_north					
Chilko_south			Chilko - Early Summer		
Mid_Horsefly	6	Quesnel	Quesnel - Summer	Hfly	
Upper_Horsefly					
McKinley					
Lower_Horsefly			Middle Fraser - Summer		
			Cariboo - Summer		
Mitchell____			Quesnel - Summer	Mtch	
Blue_Lead_Ck					
Wasko-Roaring					
Raft_____	5	Raft	Kamloops - Summer	RaNT	
NorthThompson	26	Misc (N. Thomp. River)			
	25	Misc (N. Thomp. Tribs.)[1]			
Birkenhead_	10	Birkenhead	Lillooet-Harrison - Late	BiBS	Late
			Harrison (D/S) - Late		
Big_Silver	28	Misc (Harrison/Lillooet)			
Lower_Adams	9	Late Shuswap	Shuswap Complex - Late	LShP	
Eagle_late					
LittleRiver-LittleShu					
Lower_Shuswap					
MiddleShuswap					
Portage_Creek	12	Portage	Seton - Late		
			Middle Fraser - Late		
Weaver____	13	Weaver	Harrison (U/S) - Late	WeCu	
Cultus_Lake	11	Cultus	Cultus - Late		

[1]No available baselines for this group. Sockeye from this group are lumped into Raft or North Thompson in-season.

While the production data are made available at the production group level, the original spawner, juvenile, and adult return data have not been collected at this level. Spawner and juvenile data are collected at a stream or survey level, and adult return data is aggregated based on DNA baseline groups (see first column, *"Baseline (50)"*, in Table 1 & Table 2). The translation table and process used for the creation of the Production Dataset has been confirmed to be functionally identical to the translation used to generate the stock-recruit timeseries used for the 2021 Fraser sockeye forecast process.

ii. Age Notation

The ***Production Dataset*** uses the Gilbert-Rich notation to reference fish age, N_b, where N is the total age of the fish, and b is the fresh-water age (also known at the age at ocean entrance; Figure 1.). For example, an age 4_2 (*"four sub-two"*) spawner has spent two years in freshwater, and two years in the

ocean before returning to spawn (1 year embryonic growth, 1 year freshwater growth and 2 years of ocean/saltwater growth).

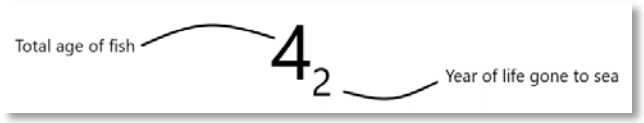

Figure 1. *Gilbert-Rich age notation.*

3. Data Inputs

As noted above, the Fraser sockeye production data is maintained in two separate files that can be easily updated, combined and interpreted; Spawner Data and Recruit Data. The datafiles are joined by code maintained by PSC staff to produce production data files and summaries that can be shared with others.

i. Spawner Data

Maintained by DFO

The spawner data within the production dataset contains Fraser sockeye spawner enumeration data collected through an intensive survey program coordinated by Fisheries and Oceans Canada (DFO) and implemented in partnerships with First Nations. DFO Stock Assessment Division biologists adjust the spawning ground survey data to develop a complete dataset of Fraser sockeye spawning escapement (herein referred to as *Spawner Data*). This dataset is meant to contain the best available estimates of the number of sockeye jacks, males, and females on the spawning grounds, as well as spawner success and effective female spawner estimates.

Spawning ground enumerations are performed using a range of different methods of varying levels of precision that can be broadly categorized as low, medium, or high precision. Historically, low precision methods were used to enumerate stocks with anticipated escapements of less than 25,000, while higher precision methods, such as mark-recapture and fences, were used for higher abundance stocks, escapements greater than 25,000. In 2004, funding limitations resulted in this threshold increasing and now only those stocks with an anticipated return greater than 75,000 are targeted with high precision methods (Grant et al. 2011). Additionally, survey site specific conditions that affect site access and the suitability of different methods to provide accurate estimates also factor into whether or not a high or low precision method is used. High and low precision methods include, but are not limited to; mark recapture, visual counts by ground (foot or boat) or helicopter survey, sonar, and fence counts. Fence and sonar counts are considered to be high precision estimates as nearly all fish are expected to be counted. Visual surveys are low precision as only a fraction of the total number of spawners are able to be visually observed. Total escapement is the result of the expansion of visual counts, with the use of a experimentally determined expansion factor. Work has been ongoing by DFO to improve calibration of visual assessment methods to minimize bias in visually enumerated populations in the Fraser River watershed (Welch and Grant 2018). Mark-recapture estimates are considered to be of a high precision, though less precise than fence or sonar enumeration. Bias in mark-recapture estimates can be experimentally identified and corrected for when producing total spawner abundance estimates.

The spawner data provided by DFO at stream specific resolution are aggregated by PSC staff to the production stock aggregation. Additional modifications to the spawner data include the imputation of spawner data for stocks and years when no enumeration occurred, for example Birkenhead River and

Horsefly River in 2002 and Mitchell River in 2005. In addition, some spawners enumerated by DFO are not included in the production dataset. These spawners are fish with no known stock of origin (e.g. fish spawning at a location without a consistent spawning population) and cannot be assigned with high confidence to individual production stock groups.

Note: Spawner data provided by DFO are not comparable with the data from the New Salmon Escapement Database System (NuSEDS). NuSEDS houses the total escapements by stream but some of the historic data is still under revision for completeness in the breakout by sex and spawning success and should not be referenced as production data.

ii. Recruit Data

The recruit data within the production dataset contains estimates of the number of Fraser sockeye recruits, or returning mature salmon, estimated by production stock group, return year, and age class (described above in Section 2ii above). Recruits are traditionally based on age-specific spawning ground estimates plus catch (historical run size estimates) but over the years, methods have varied. In recent years, increased natural mortality due to adverse migration conditions in the Fraser River resulted in the need to correct these estimates for en-route losses. Prior to 2009, the corrections were based on difference between estimates (DBEs), i.e. differences between Mission minus upstream catches and the spawning ground estimates, resulting in post-DBE run size estimates. Since 2009, a run size adjustment (RSA) process has been used for the post-season evaluation of different factors influencing en-route mortality and evaluations of biases in the estimates of spawners and catch. Because this process is expert-led and requires various assessments and detailed discussions among participants, the resulting post-RSA run size may not be available for one or potentially more years. In-season run size estimates, derived from abundance estimates at Mission plus seaward catches can provide preliminary run size estimates until post-RSA run size estimates are available.

Post-RSA run size

Post-RSA estimates are derived through the RSA process and rely on spawning ground estimates, catch estimates and environmental conditions affecting survival as well as evidence of biases in spawning ground and catch estimates. The Run Size Adjustment (RSA) can be seen as an adjustment to the total number of recruits in cases where there is evidence that the in-season estimate of the number of fish returning based on Mission escapement plus seaward catch deviates from the post-season estimate of catch plus spawning escapement. Because their reliance on spawning ground estimates, post-RSA estimates are considered to better represent the true number of recruits.

In-season-based run size

In-season-based recruit estimates are derived using Mission escapement estimates and seaward catches in combination with stock identification information. Although, in-season-based estimates are available immediately following the termination of in-season assessment programs, they do not incorporate information provided by spawning ground enumeration surveys except for spawning ground age composition information. Age composition of in-season-based recruit estimates are initially generated using expert judgement and information from test and commercial fisheries, as well as forecast and brood age composition, but are adjusted using spawning ground samples once available in November/December of the return year (see Age Composition section below). Another challenge associated with the in-season-based run size estimates is the stock ID information required to distinguish individual stocks. In-season run size estimates rely on in-season catch samples to derive stock composition. In addition, particular stocks are not strongly differentiated by currently implemented DNA methods (Beacham et al. 2004). The quality of the stock ID information will impact the reliability of the in-season-based run size for individual stocks (see Stock Composition section below).

Age Composition

Within the recruit dataset, the final estimation of age composition relies upon age composition data from scale and otolith samples taken from the spawning grounds as well as from test and commercial fisheries sampled during the return of the run. Where sample sizes allow, spawning ground information is used preferentially over samples from fisheries. In case scale and otolith from the spawning grounds do not meet an expert selected sample size threshold, the age composition of certain groups may be determined by combining spawning ground samples with samples obtained from test and commercial fisheries. If the sample size threshold remains unmet, age composition is "gap-filled" using expert judgement and additional sources of information such as the forecast age composition, age composition of similar production groups, and the age composition of the brood year return.

Stock Composition

In-season-based run size estimates rely on stock identification results from catch samples to derive stock compositions. The estimates improve once the stock-specific in-season estimates are revised based on improved stock ID information from the spawning grounds. The stock composition concerns associated with in-season-based run size estimates mainly relate to stocks that are not strongly differentiated by the current DNA methods (Beacham et al. 2004). To convey the quality of the stock ID used to derive the in-season-based run size estimate, a stock ID quality field has been added to the detailed production file. In-season-based run size estimate should not be used if there are major concerns about the ability to distinguish a stock from other comigrating stocks (e.g. Late Stuart and Stellako)

iii. Production Data for Distribution Files

The spawner dataset and recruit dataset are combined to produce the complete production dataset. The production dataset is output for distribution in two separate formats, the first is known as the ***Summary Format***, the second is the ***Detailed format***.

In the ***Summary Format***, each row provides the production stock information by brood year. This format provides users a general overview of the production dataset and is a simpler form. Some important information relating to sockeye productivity cannot be viewed in this format as information pertaining to particular return years cannot be shown by brood year. For example, adding the number of recruits of different ages within a particular return year must be done by selecting values across multiple rows and columns. In addition, data that requires additional metadata unavailable in the summary file to correctly interpret and use the data have been omitted from the summary, such as the in-season-based run size estimate which requires information on stock ID quality.

The ***Detailed Format*** contains the most up to date production data including in-season based run size estimates as well as the most recent data for younger recruits with return data for older siblings still outstanding. In this detailed format, each row is a unique cohort of salmon that are from the same production stock, have the same brood year, and had the same age when they returned to the Fraser River. This format allows users, both in Excel and R, to properly manipulate and summarize the production data while retaining detailed metadata or supplemental data associated with different brood and return years. This data format can be thought of as the "long" version as each age class has its own record resulting in many more rows than in the ***Summary Format***. This results in duplication of the spawner data for each recruit age class. Users must be aware of this duplication of the spawner data as seen below in Figure 2.

stock_broodyr_age	production_stock_name	brood_year	total_BY_spawners	total_BY_EFS	total_BY_age1_smolt	total_BY_age2_smolt	total_BY_fry_estimate	return_year	age	num_age_class_return
Chilko_1993_21	Chilko	1993	561865	322298	39721810	129657	NA	1995	21	0
Chilko_1993_31	Chilko	1993	561865	322298	39721810	129657	NA	1996	31	0
Chilko_1993_32	Chilko	1993	561865	322298	39721810	129657	NA	1996	32	18173
Chilko_1993_41	Chilko	1993	561865	322298	39721810	129657	NA	1997	41	0
Chilko_1993_42	Chilko	1993	561865	322298	39721810	129657	NA	1997	42	3380892
Chilko_1993_43	Chilko	1993	561865	322298	39721810	129657	NA	1997	43	208
Chilko_1993_52	Chilko	1993	561865	322298	39721810	129657	NA	1998	52	548667
Chilko_1993_53	Chilko	1993	561865	322298	39721810	129657	NA	1998	53	15931
Chilko_1993_62	Chilko	1993	561865	322298	39721810	129657	NA	1999	62	0
Chilko_1993_63	Chilko	1993	561865	322298	39721810	129657	NA	1999	63	0

Figure 2. Example of Detailed Format table for Chilko production data from brood year 1993. Spawner data (blue) is duplicated for each unique age class (red).

A description of the field present in the production dataset files for distribution can be found in Table 3 (fields in the Production Data Summary Format file), Table 4 (additional fields in the Production Data Detailed Format file).

Table 3 Description of fields in the Production Data Summary Format data file.

Field	Description
production_stock_name	Name of the Fraser River sockeye production stock group. These stock groupings are used to explore stock-recruit relationships and for forecasting future salmon recruits/returns (see 2i Population Aggregation).
production_stock_code	Numeric code assigned to Fraser sockeye forecast stock group.
broodyr	For adult spawner estimates, this is the year of their return to the spawning grounds. For juvenile estimates, this is the year they were propagated or spawned (i.e. the year of the parental return and the fertilisation of the eggs).
total_broodyr_spawners	Total number of spawners estimated to have returned to the spawning grounds in the brood year. Includes both adults and jacks.
total_broodyr_EFS	Effective Female Spawners (EFS): An estimate of female spawner abundance in the brood year (as opposed to the more traditional total of male and female spawners), which is further adjusted downward by the proportion of eggs that were not spawned, as determined by sampling on the spawning grounds.
total_broodyr_age1_smolt	Total number of estimated downstream migrating age-1 smolts. An age-1 smolt has spent one year rearing in freshwater.
total_broodyr_age2_smolt	Total number of estimated downstream migrating age-2 smolts. An age-2 smolt has spent two years rearing in freshwater.
total_broodyr_fry_estimate	Total number of fry estimated the year following the brood year (t + 1). Fry enumeration occurs either as they migrate down the stream where spawning occurred, or once in the rearing lake.
recruits_ageXX	This field represents the estimated total number of recruits of age XX.

Table 4 Description of additional fields in the Detailed Format file.

Field	Description
stock_broodyr_age	Only present in the detailed output. A composite key of stock, brood year, and age at return in order to uniquely identify an individual production cohort.
returnyr	This is the year that the recruits returned to the Fraser River watershed to spawn.

total_broodyr_age1_smolt	Total number downstream migrating age-1 smolts associated with one brood year. An age-1 smolt has spent one year rearing in freshwater.
total_broodyr_age2_smolt	Total number of downstream migrating age-2 smolts associated with one brood year. An age-2 smolt has spent two years rearing in freshwater.
total_broodyr_fry_estimate	Total number of fry associated with one brood year, estimated the year following the brood year (t + 1).
num_recruits	This field represents the estimated total number of recruits of age associated with a brood year
hatchery_supplementation	This field provides information on whether the natural population was supplemented through the use of hatchery stock or through a captive broodstock program. Detailed information on the particular supplementation program is not provided within this dataset.
active_spawning_channel	Field for whether there was a spawning channel active for the particular production stock group. In some cases, information on whether a spawning channel was active in a specific year is not available. In these cases, a range of years are documented as "Inconsistent channel use"
major_spawn_event	Description of major events that occurred within the river that would have impacted spawning success and recruitment, e.g., lake fertilization, major landslide event, mine breach, incomplete spawner enumeration, etc.
run_size_type	Identifies the method of recruitment estimation: 1) **historical** estimates based on spawner numbers plus catch, 2) **post-DBE** estimates based on spawner numbers plus catch plus the difference between Mission minus upstream catch and the spawning ground estimates, 3) **post-RSA** estimates based on spawning numbers plus catch adjusted to account for in-river mortality, and/or escapement and catch estimation biases, or 4) **in-season-based** run size estimates, derived from abundance estimates at Mission plus seaward catch in combination with stock ID information.
age_comp_method	The method used in order to derive age composition of the return of the production stock group. Spn. Gr. = Spawning Ground samples, Comb. = Combination of Spawning Ground and Fishery samples, Manual = Expert Judgment informed by similar stocks, forecast age composition, and brood year age composition, Prnt. Stk. = age composition taken from parental stock.
stock_ID_quality	Information on the ability to distinguish a stock from comigrating stocks given the current DNA methods (Beacham et al. 2004): 1) <50% correct classification, 2) 50-80% correct, 3) 80-90% correct, 4) 90-95% correct or 5) >95% correct. In-season-based run size estimate should not be used if there are major concerns about the ability to distinguish a stock
big_bar_flag	A field to flag whether the Big Bar landslide impacted the returning cohort during their upstream migration.
exclude_SR_modelling_spawners	This field acts as a flag to warn users to not use this particular record of spawner or juvenile data for stock-recruit analyses and documents the reason. In most cases, this is either due to human intervention into the productivity of the population through the use of a spawning channel, hatchery supplementation, or due to insufficient or low-quality data.
exclude_SR_modelling_recruits	This field acts as a flag to warn users to not use this particular record of recruit estimates for stock-recruit analyses. Exclusion of recruit estimates is generally less common than the exclusion of spawner estimates.

References

Amended Annex IV of the Treaty between the Government of Canada and the Government of the United States of America concerning Pacific Salmon. Entry into force 1 January 2020.

Beacham, T.D., Lapointe, M., Candy, J.R., McIntosh, B., MacConnachie, C., Tabata, A., Kaukinen, K., Deng, L., Miller, K.M., and Withler, R.E. 2004. Stock Identification of Fraser River Sockeye Salmon Using Microsatellites and Major Histocompatibility Complex Variation. Trans Am Fish Soc **133**(5): 1117–1137. doi:10.1577/t04-001.1.

DFO. 2021. Pre-season run size forecasts for Fraser River Sockeye (*Oncorhynchus nerka*) and Pink (*Oncorhynchus gorbuscha*) salmon in 2021. *In* Can. Sci. Adv. Sec. Sci. Resp.

Grant, S.C.H., Macdonald, B.L., Cone, T.E., Holt, C.A., Cass, A.J., Porszt, E.J., and Pon, L.B. 2011. Evaluation of Uncertainty in Fraser Sockeye (*Oncorhynchus nerka*) Wild Salmon Policy Status using Abundance and Trends in Abundance Metrics. Canadian Science Advisory Secretariat, Research Document **87**: 191 p.

Welch, P., and Grant, A. 2018. Calibration of Visual Assessment Methods for Fraser River Sockeye Salmon (*Oncorhynchus nerka*) - Year 9. Report to the Southern Boundary Restoration and Enhancement Fund (May): 10–27.