Review of run size and escapement estimates for the Nass River system: Chinook, Sockeye, and Coho salmon, and summer-run steelhead. Part 2 – Bayesian capture-recapture model

Carl James Schwarz, Richard Alexander, Charmaine Carr-Harris, Ian Beveridge, Cameron Noble, Chelsea May

Fisheries and Oceans Canada
Science Branch, Pacific Region
417 – 2nd Avenue West
Prince Rupert, BC V8J 1G8

2026

Canadian Technical Report of Fisheries and Aquatic Sciences 3774

Fisheries and Oceans Canada Pêches et Océans Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of
Fisheries and Aquatic Sciences 3774

2026

REVIEW OF RUN SIZE AND ESCAPEMENT ESTIMATES FOR THE NASS RIVER SYSTEM: CHINOOK, SOCKEYE, AND COHO SALMON, AND SUMMER-RUN STEELHEAD. PART 2 – BAYESIAN CAPTURE-RECAPTURE MODEL

by

Carl James Schwarz[1], Richard Alexander[3], Charmaine Carr-Harris[2], Ian Beveridge[3],
Cameron Noble[3], Chelsea May[2]

[1] StatMathComp Consulting by Schwarz, Inc
625 Bentley Rd
Port Moody, BC V3H 3A4

[2] Science Branch – Pacific Region
Fisheries and Oceans Canada
417 - 2nd Avenue West
Prince Rupert, BC V8J 1G8

[3] LGL Limited Environmental Research Associates
9768 Second Street
Sidney, BC, V8L 3Y8

Cat. Fs 97-6/3774E-PDF ISBN 978-0-660-99676-9 ISSN 1488-5379

https://doi.org/10.60825/q1zj-cm92

Correct Citation for this publication:

Schwarz, C.J., Alexander, R., Carr-Harris, C., Beveridge, I., Noble, C., and May, C. 2026. Review of run size and escapement estimates for the Nass River system: Chinook, Sockeye, and Coho salmon, and summer-run steelhead. Part 2 – Bayesian capture-recapture model. Can. Tech. Rep. Fish. Aquat. Sci. 3774: x + 71 p. https://doi.org/10.60825/q1zj-cm92

Table of Contents

List of Tables

List of Figures

List of Appendices

Appendix A: Mathematic description of the Bayesian model

ABSTRACT

Schwarz, C.J., Alexander, R., Carr-Harris, C., Beveridge, I., Noble, C., and May, C. 2026. Review of run size and escapement estimates for the Nass River system: Chinook, Sockeye, and Coho salmon, and summer-run steelhead. Part 2 – Bayesian capture-recapture model. Can. Tech. Rep. Fish. Aquat. Sci. 3774: x + 71 p. https://doi.org/10.60825/q1zj-cm92

This is the second part of a review of the capture-recapture methodology used to estimate abundance of Chinook, Coho, and Sockeye salmon and summer-run steelhead on the Nass River, British Columba.

Part I of the review noted that while the current capture-recapture method gives reasonable point estimates, uncertainty is under reported. Part I of the review recommended a Bayesian model be developed that could account for multiple adjustments.

The largest component of the uncertainty in the run size is from the estimate of initial tagging mortality. These are often based on a small number of studies with small sample sizes. While additional studies would be required to reduce this source of uncertainty, the Bayesian model can provide more realistic estimates of uncertainty around run size estimates than the current capture-recapture methods.

The Bayesian model is described in general terms (and mathematically in the Appendix). It is applied to the 2021 data and the results are compared to the estimates computed using the current methods. Point estimates of the run sizes were very similar but as expected, estimates of uncertainty were slightly larger from the Bayesian model for Chinook, Coho, and steelhead run size estimates and substantially larger for estimates of run size for Sockeye Salmon. Hindcasts back to 1999 showed the same pattern. Further, sex-, length- and time-stratified and unstratified estimates show little difference when realistic uncertainty is incorporated.

RÉSUMÉ

Schwarz, C.J., Alexander, R., Carr-Harris, C., Beveridge, I., Noble, C., and May, C. 2026. Review of run size and escapement estimates for the Nass River system: Chinook, Sockeye, and Coho salmon, and summer-run steelhead. Part 2 – Bayesian capture-recapture model. Can. Tech. Rep. Fish. Aquat. Sci. 3774: x + 71 p. https://doi.org/10.60825/q1zj-cm92

L'estimation de la mortalité initiale liée au marquage est la plus grande composante de l'incertitude de la taille des remontées. Elle est souvent basée sur un petit nombre d'études avec des échantillons de taille réduite. Bien que des études supplémentaires soient nécessaires pour réduire cette source d'incertitude, le modèle bayésien peut fournir des estimations plus réalistes de cette incertitude. Cette nouvelle méthode fournirait des estimations de taille des remontées plus précises que les méthodes actuelles de capture-recapture. Le modèle bayésien est décrit en termes généraux (et mathématiquement dans l'annexe). Il est appliqué aux données de 2021 et les résultats sont comparés aux estimations calculées à l'aide des méthodes actuelles. Les estimations ponctuelles de aille des remontées étaient très similaires, mais comme prévu, les estimations d'incertitude étaient légèrement plus élevées dans le modèle bayésien dans le cas des estimations de taille des remontées de saumons quinnats, cohos et arc-en-ciel, et nettement plus élevées pour les estimations de taille des remontées de saumons rouges. Les rétrospectives remontant à 1999 ont montré la même tendance. En outre, les estimations stratifiées et non stratifiées par sexe, longueur et période montrent peu de différence lorsqu'une incertitude plus réaliste est prise en compte.

1. Introduction

This is the second part of a review of the methodology used to estimate abundance of Chinook, Coho, and Sockeye salmon and summer-run steelhead on the Nass River, British Columbia.

Part I of the review (Schwarz et al. 2025) found that while the current method gives reasonable point estimates of abundance, the uncertainty is under reported and recommended adopting a Bayesian approach to address the following:

1. The uncertainty in the adjustments to the number of fish with tags available upriver of Grease Harbour (GH) due to various losses between Gitwinksihlkw (GW) and GH. The adjustments account for spawning between GW and GH, initial tagging mortality / fall back, tag loss, and the differential harvest of tagged and untagged fish in the Nisg̲a'a Domestic Harvest. Currently, only a point estimate of the number of fish with tags removed between GW and GH is used, and the subsequent analysis (e.g., a Pooled-Petersen estimator) treats the estimated number of fish available with tags just upriver of GH as a known quantity.
2. Several components that affect the number of fish with tags involve multi-year studies (e.g., 2-3 years for the initial tagging mortality, several years of genetic information, and several years of tag loss studies). For the current methods, only data obtained in the particular year are used and no information from other years is "shared" with the current year. A Bayesian hierarchical model can share information across years, particularly if the yearly data are sparse (e.g., 0 fish observed with tag loss in a small sample examined in a particular year).

Two alternative Bayesian models have been created. The first is an *unstratified* model where all fish (above the minimum fork length and therefore not considered "jacks") are treated as a single population. The second is a *stratified* model where fish are stratified by length (e.g., small and large fish). The model for each length stratum is the same as in the unstratified model with minor exceptions. A temporally stratified analysis was also conducted (Coho and Sockeye salmon only) where the week of tagging and/or recovery are used. This was analyzed using previously developed software (*SPAS* and *BTSPAS*).

The Bayesian models are applied to data collected since 1999 and the hind-casts compared to the reported estimates. Some care must be taken with the hind-casts since tagging methods and methodology have changed over the years. Nevertheless, the conclusion is that the Bayesian models produce very similar point estimates to current methods but the estimates of uncertainty from the Bayesian methods are consistently larger than those currently reported.

The code and data for this report are currently stored in a private Github repository that users may access by contacting the authors (email Charmaine.Carr-Harris@dfo-mpo.gc.ca). This report references two Excel workbooks, including:

1. An Excel workbook with all of the data used in these reports; and
2. An Excel workbook with summary information from historical reports.

The report also references several supplemental (web) appendices, listed below. These can be downloaded from a public Git repository.

1. A supplemental report on fitting the unstratified and stratified (by length) Bayesian model to data for any particular year. This includes detailed summary statistics on the prior and posterior distributions, density plots of the prior and posterior distributions, and goodness-of-fit statistics;
2. A supplemental report on fitting the Bayesian model for the four species from 1994 to 2024 (but note that results for 2024 are preliminary) and comparing the results to the previously published estimates. Some caution is needed because past years' published estimates may have used different assumptions (e.g., on handling mortality probabilities) or only had limited data on tag loss. The main purpose of this retrospective analysis is to show that published estimates of uncertainty for the run estimates are under reported;
3. A supplemental report that demonstrates the "automatic stratification" in the stratified model, and shows that when sample sizes are small, the stratified model forces the estimated marked fractions to be close, but when sample sizes are large, each stratum can freely have its own marked fraction;
4. A supplemental report examining the performance of a temporally stratified estimator for Coho and Sockeye salmon. Temporal stratification is currently not feasible for Chinook Salmon or summer-run steelhead due to small numbers of tags applied and recovered by week; and
5. A supplemental report on the impact of adding an additional recovery site (Kwinageese Weir) that was established and has operated annually since 2009 to the current Meziadin Fishway recovery site for Sockeye Salmon.

A mathematical description of the Bayesian model is provided in Appendix A.

2. Unstratified and Stratified Bayesian Model

In this section, we present a general introduction to the parts of the Bayesian capture-recapture model. Because the stratified Bayesian model treats the two strata as independent, the model for each stratum is the same as that of the unstratified model with a few exceptions as noted below. The full mathematical description of the Bayesian model is presented in Appendix A.

2.1 Part I. Modelling the net number of tags just above Grease Harbour available for recapture at upstream locations

Fish are tagged at fishwheels at GW, and at GH for Chinook Salmon and summer-run steelhead) and may be recovered at several upstream locations, including Meziadin

Fishway (MF), Kwinageese Weir (KW), Damdochax Creek (DC), and Cranberry River (CR) (Figure 1), depending on species. For example, currently the Bayesian model only uses recoveries of Sockeye at the Meziadin fishway, but other recovery locations could be incorporated in the future. The number of tagged fish is adjusted for initial tagging mortality, fall back, tag loss (for Coho and Sockeye salmon), spawning between GW and GH (for Chinook Salmon and summer-run steelhead), removals at fishwheels, removals in the Nisg̱a’a Domestic Harvest, and others, described below.

All Chinook Salmon are fully inspected for the presence of a tag and a batch mark, so tag loss is not an issue. Steelhead are currently fin-clipped from fishwheel captures after 1 July as summer-run fish so tag loss is not an issue.

Because Chinook Salmon and steelhead are tagged at both GW and GH, an adjustment for fish tagged at GW that spawn between GW and GH is needed. Because Coho and Sockeye salmon are only tagged at GW, no adjustment for spawning between GW and GH is needed. However, now the estimated run size at GH for Coho and Sockeye salmon also includes spawning escapement between GW and GH.

In Part 1 of the model, the net number of tags available for recapture just upstream of GH is estimated. This modelling is done because some removals (e.g., Nisg̱a’a Domestic Harvest) are thought to differentially affect tagged and untagged fish based on historical assessments (e.g., Link and English 1997).

This first step computes the net number of tagged fish available just upriver of GH as:

GH.tags.net	*= GW.tags + GH.tags*	Tags applied
	- tags.init.mort	Tags not available due to initial mortality/fall back/drop out
	- tags.spawn.GW.GH	Tags not available due to fish that spawn between GW and GH (Chinook and steelhead only)
	- tags.tagloss	Tags not available due to fish that lose tags (Coho and Sockeye salmon only)
	- tags.NH.GW.GH	Tags not available because they are captured in the Nisg̱a’a Domestic harvest between GW and GH
	- tags.GW.fishwheel.harvest	Tags not available because they are removed in the GW fishwheel harvest
	- tags.GH.fishwheel.harvest	Tags not available because they are removed in the GH fishwheel harvest
	- tags.Nisga’a.harvest	Tags not available because they are removed in the Nisga’a fishwheel harvests
	- tags.sport.GW.GH	Tags not available because they are removed in the sport fishery between GW and GH.

The detailed explanation of the computation of these values is explained in Appendix A and summarized below.

2.1.1 Tagged fish lost due to initial handling mortality, fall back or drop out

Radio-tagging studies have provided estimates of initial handling mortality,fall back and drop out. In each study, the number of radio tags that were classified as initial mortality, fall back, and drop out is summarized in Part I of the review (Schwarz et al. 2025).

There are two levels of uncertainty that need to be accounted for when using these results for a particular analysis year. The first is the uncertainty in the initial mortality probability from a specific mortality study. The second is the year-to-year variation in this probability over several mortality studies.

For each mortality study, the observed number of fish that suffered fall back or initial tagging mortality is modelled by a binomial distribution; this estimates the year-specific initial mortality probability, and the Bayesian Markov Chain Monte Carlo (MCMC) procedure provides samples from the posterior that reflect the uncertainty of the estimated probability. For example, mortality studies with larger numbers of fish will tend to have smaller uncertainty about the year-specific initial mortality probability.

Because there is also year-to-year variation in the initial mortality probabilities, a model for this year-to-year variation is a Normal distribution on the *logit* (log-odds) scale. Because only a few radio telemetry studies have occurred to quantify initial mortality for each species, there is also uncertainty about the parameters of the year-to-year process that reflects the number of studies. More studies will lead to better information on the year-to-year parameters.

Finally, for the years in which tagging takes place, we generate the probability for the number of tags lost due to initial mortality, fall back or drop out by first sampling from the year-to-year distribution to get the mortality probability for this analysis year. Then we use a Binomial distribution to model the number of tagged fish lost to mortality/ fall back after tagging.

This hierarchical process (sampling from the year-to-year distribution and then modelling the actual number of tagged fish lost to initial mortality) will incorporate the year-to-year variation in the mortality probability, the uncertainty due to the small number of mortality studies and small sample sizes in each year of the mortality studies, and binomial variation in the actual number of fish that lose tags upstream of GH. The tagging mortality is assumed to be instantaneous so there is no opportunity for these fish to spawn.

2.1.2 Tagged fish lost due to spawning between Gitwinksihlkw and Grease Harbour (Chinook Salmon and steelhead only)

After fish are tagged and released at GW, some fish may spawn between GW and GH. Only those fish tagged at GW are subject to “removal” due to spawning between GW and GH before arriving at GH.

For Chinook Salmon and steelhead, fish are tagged at GW and GH. Only fish tagged at GW are subject to removal due to spawning between GW and GH, and then the estimated run size EXCLUDES the number of fish that spawn between GW and GH. Genetic samples taken at GW provide an estimate of the proportion of the fish captured at GW that spawn between GW and GH through genetic stock identification (GSI) methods. One unresolved issue is dealing with Chinook Salmon from stocks that traditionally spawn below GW (Ksi Hlgin<u>x</u> [Ishkeenickh River]; Ksi Gin<u>g</u>olx [Kincolith River]) but appear at the GW fishwheels (i.e., overshoot their traditional spawning grounds). We are assuming that any fish that appears at the GW fishwheels continues upstream and does not fall back. Current GSI data indicates that between 0.5% and 7.5% (average = 2.7%) of fish at GW belong to stocks that traditionally spawn below GW (Beveridge et al. 2025).

For Coho and Sockeye salmon, fish are only tagged at GW and no adjustment is done to the number of available tags, but the estimated run size then INCLUDES the number of fish that spawn between GW and GH.

We again create a hierarchical structure. First, the actual number of fish in the GSI samples that belong to the stocks that spawn between GW and GH is modelled using a binomial distribution. Second, the year-to-year variation in the probability of spawning is modelled using a Normal distribution on the *logit* scale.

The hierarchical model “shares” information across years when sample sizes are small. For example, suppose that in a particular GSI analysis, 1 of 10 fish belonged to the stocks that spawn between GW and GH (i.e., Ksi Sii Aks [Tseax R.] and Ksi Sg̲asginist [Seaskinnish Cr.]), but in other years the sample sizes are much larger and the mean proportion that spawn between GW and GH is around 13%. Then the estimate of 10% (1 fish in 10 fish analyzed) for this study based on small sample size is “shrunk” upwards towards the overall mean of 13%.

There are two cases to consider when modelling the number of fish with tags that spawn between GW and GH. First, if there is no genetic study in the analysis year, the two-step process is followed by first sampling a probability from the year-to-year distribution for the spawning proportion and then modelling the actual number of tagged fish that would spawn between GW and GH using a Binomial distribution. Second, suppose that in a particular year, the mark-recapture study includes genetic analyses. Then we don’t need to sample from the year-to-year variation but instead use the (potentially shrunk) estimates from this study year directly in the Binomial distribution.

2.1.3 Tagged fish that lose tags (Coho and Sockeye salmon only)

Tagged fish that migrate upstream can suffer tag loss which can lead to bias in the capture-recapture estimates unless corrected. As shown in Part I of the review (Schwarz et al 2025), there are two mathematically equivalent ways to deal with tag loss. We have chosen to adjust the number of fish tagged at GW and GH that are available just upriver of GH for tag loss.

Tag loss is not an issue for Chinook Salmon because all fish are carefully inspected for the presence of a tag and the batch secondary mark. Tag loss is not an issue for steelhead because all fish are fin-clipped and it is assumed that the adipose mark will not grow back.

As fish pass the Meziadin Fishway (MF), they are classified as having an observable tag or not. In addition, a sub-sample of fish with no observable tags is taken in several years and fully inspected for the presence of a batch secondary mark.

A hierarchical model is again used. First, the year-to-year variation in tag loss probability is again modelled using a Normal distribution on the *logit* scale. Second, the number of observed fish with tag loss follows a Binomial distribution whose parameters depends on the tag loss probability. The hierarchical model again “shares” information across years when sample sizes are small.

Again, there are two cases to consider when generating the number of fish with tag loss. First, if there is no tag loss study in the analysis year, the two-step process is followed by first sampling a probability from the year-to-year distribution and then modelling the actual number of fish that lose their tag using a binomial distribution. Second, suppose that a particular year includes a tag loss study. Then we do not need to sample from the year-to-year distribution, but use the (potentially shrunk) estimates from this study year directly in the Binomial distribution.

The current Bayesian model only considers tag loss studies conducted at MF and assumes that no additional tag loss occurs upstream of the fishway. Additional upstream locations may provide additional tag loss information (e.g., Kwinageese Weir) and could be combined with tag loss information from MF if no additional tag loss between MF and the upstream location occur. Before "adding" this additional information to the analysis, the marked-fraction at the new locations needs to be compared to that at MF to check the assumption of no additional tag loss.

2.1.4 Tagged fish removed in the Nisg̱a'a Domestic Harvest between GW and GH

The Nisg̱a'a Domestic Harvest consists of two components:

1. the Domestic Fishery (DF) whose catch is estimated using survey methods. Not all tags that are removed in the Domestic Fishery are reported. The estimates and precision of the Domestic Fishery harvest is found using survey methods and estimates are produced yearly.
2. the Individual-Sale Fishery (ISF). Part of the ISF is closely monitored (typically 90%+) as fish are sold and all tags in the monitored fishery are counted. The ISF catch is included in the total catch estimate from the Domestic Fishery.

The number of fish with tags removed in the Nisg̱a'a Domestic Harvest (DF and ISF combined) is estimated based on the tag ratio (proportion of catch that is tagged) in the fully observed ISF. A tag lottery draw is conducted annually to encourage fishers to return tags from the Domestic Fishery.

Compliance information has been recorded since 2000 in three river-strata around and downstream of GH. Notice that in some years, no ISF occurred and that in many years the number of tags returned from the Lower river-stratum (Lax̱galts'ap) is very small in the ISF. As well, in recent years, fewer tag recoveries were returned to the monitors due to Covid-19.

The total catch is also not known with certainty, but full-time catch monitors are present in each of the Nisg̱a'a communities for interviewing fishers on a daily basis. On-water surveys are also conducted to determine weekly fishing effort. The standard error of the estimates of total catch was extracted from catch estimation reports created by the Nisga'a Fish and Wildlife Department (NFWD) which are based on standard survey sampling methodology and assumed to be correct.

A multi-step process is used to estimate the number of fish with tags that were harvested (and removed) between GW and GH.

First, in each river stratum, the observed number of tags in the individual sale fishery is modelled using a Binomial distribution. This provides information on the marked fraction (proportion of fish that are tagged) in each river stratum. Even if the observed number of tags in the individual sale fishery is zero, the prior distribution provides some information on the marked fraction for each stratum. Similarly, if no monitoring is done because there were no sale fisheries, the prior distribution again provides information on the marked fraction.

Second, the actual harvest is modelled based on the estimated harvest and standard error using a Normal distribution.

Third, the modelled marked fraction is used to estimate how many tags are removed in the actual harvest for each river stratum.

Fourth, the modelled harvests are added over the river strata to estimate the total number of fish with tags removed in the harvest.

The use of distributions at each level in the process provides information on the uncertainty in the total number of fish with tags removed in the harvest between GW and GH.

Note that the total catch used to expand the tag-fraction in the ISF ($TotalCatch_h$) may differ from the total catch reported in harvest monitoring due to timing issues. For example, harvest monitoring may measure harvest starting in May, but fishwheels may not start operating until June and so only the harvest from June onwards, when the fishwheels are operating and tagging fish, is used. In most cases, the difference in estimates of catch is small except for Sockeye Salmon in 2000-2006. **In the current Bayesian implementation, the value for the total harvest in river stratum *h* ($TotalCatch_h$) is taken from the harvest reports and not adjusted for timing issues.**

In the unstratified Bayesian model, no further computations are needed. In the stratified model (e.g., by fish length), the allocation of the total number of fish with tags removed in the Nisg̲a'a harvest to the individual strata is based on a sub-sample of fish from the harvest that are measured and classified into the individual strata. We assume a multinomial model for this subsample and then apply the multinomial model to the total number of fish with tags harvested.

2.1 5 Tags lost to the Grease Harbour fishwheel harvest

This is an exact observed value and used directly without any uncertainty.

2.1.6 Tags lost in the Nisga'a sale harvests

These are estimated values, and the estimate and standard error are used to model values using a (discretized) Normal distribution.

2.1.7 Tags lost in the sport harvest between Gitwinksihlkw and Grease Harbour

These are estimated values, and the estimate and standard error are used to model values using a (discretized) Normal distribution.

2.2 Part II. Modelling the marked fraction just upriver of Grease Harbour

The marked fraction (i.e., proportion of the run just upriver of GH that has a tag) is assumed to be constant over the rest of journey upstream (i.e., all tag loss has occurred, all initial mortality has occurred, etc.). Any subsequent removals upstream of GH that operates on tagged and untagged fish equally will not affect the marked fraction and so can be "ignored".

Recoveries could occur at several locations (e.g., Meziadin Fishway, Kwinageese Weir, Damdochax Creek, Cranberry River and others; Figure 1). At each location the number of fish inspected for tags and the number of tags observed is assumed to be recorded exactly without error. The number of tagged fish observed is modelled using a Bailey (1951, 1952) Binomial model that depends on the marked fraction.

In the unstratified Bayesian model, prior information about the marked fraction is modelled using a Beta distribution.

In the stratified Bayesian model, the prior information about the marked fraction in the first fish-stratum is also modelled using a Beta distribution. The marked fraction for the second fish-stratum is allowed to differ from that on the first fish-stratum after transforming to the *logit* scale with the difference (on the logit scale) following a Normal distribution. If the difference is close to 0, this would imply that that marked fraction is very similar in the two fish-strata and stratification would not be needed.

This set up for the stratified model follows the methods and recommendation in Al Amer (2021) with BUGS code similar to that used in their supplementary materials.

It is equivalent to modelling the odds ratio for the marked fraction between the two fish-strata with a prior distribution centered at 1 (or 0 on the *logit* scale) (corresponding to no difference in the marked fraction). If the data in each stratum are strong, the two estimated marked fractions will be different, but in cases where the data are very sparse, the two marked fractions will then be very similar. This is a form of “automated” switching between a fully stratified and fully pooled Petersen estimator. An example of this automated stratification is available in the supplemental information.

2.3 Part III. Modelling the run size at Grease Harbour

The information on the marked fraction and the net tags available just upriver of GH leads to the estimate of the run size at GH as:

$$GH.runsize = \frac{GH.tags.net}{marked.fraction} + GH.fishwheel.harvest.$$

The posterior distributions for the net number of tags available just upriver of GH and for the marked fraction just upriver of GH induce the posterior distribution for the run size just upriver of GH. Because of the multi-step sampling process in the MCMC algorithm, the uncertainty in each step of the process is automatically propagated to provide an estimate of the uncertainty in the GH run size.

Notice that for Coho and Sockeye salmon, this estimated run size just upriver of GH already includes spawning escapement between GW and GH because tags were only applied at GW and no adjustment for escapement between GW and GH is made. If tags were applied at both locations, a similar adjustment could be made as for Chinook Salmon and steelhead.

The run size estimate at GH post fishwheel harvest is available for each fish-stratum, but the run size at GH prior to the fishwheel harvest cannot be stratified (e.g., by length) because the fishwheel harvest values are not available for the individual fish-strata.

2.4 Part IV. Modelling the run size at Gitwinksihlkw

The total run size at GH (i.e., summed over all fish-strata), is then adjusted for sport harvest, Nisg̱a'a harvest, and escapement that occurs between GW and GH. All of these adjustments are only done at the pooled level, i.e., the estimates of run size at GW are not stratified.

For Sockeye and Coho salmon, tagging only occurs at GW. Consequently, the run size estimate at GH already includes escapement between GW and GH and no further adjustments are needed. The run size for Sockeye and Coho salmon at GW is found by adding the estimated sport and Nisg̱a'a harvest between GW and GH, and the fishwheel harvest at GW. The sport and Nisg̱a'a harvest are modelled based on the estimate and standard error for these quantities:

$$GW.runsize = GH.runsize + SpH.GW.GH + NiH.GW.GH + GW.fishwheel.harvest$$

For Chinook Salmon and summer-run steelhead, the run size at GH plus the Nisg̱a'a and sport harvest between GW and GH must be inflated by (1-p(escapement between GW and GH)). We then have:

$$\begin{aligned} GW.runsize = {} & (GH.runsize + SpH.GW.GH + NiH.GW.GH)(1 - p(escapement)) \\ & + GW.fishwheel.harvest \end{aligned}$$

Again, the uncertainty in each of the terms in the above equations is automatically propagated to the estimate of the run size at GW.

2.5 Bayesian model applied to 2021 data

A supplemental document has the (very) granular details on the unstratified and stratified analysis for 2021 (and other years) and should be consulted as needed.

2.5.1 Unstratified model

A summary of the results from an unstratified analysis of the 2021 data for the four species is found in Table 1. The total applied tags at the fishwheels in 2021 ranged from approximately 200 to 9,000 tags.

There were no fishwheel harvests in 2021, so the number of tagged fish removed by the fishwheel harvest is zero. Similarly, there was no demonstration harvest in 2021 and the estimated number of tagged fish removed by the sport fishery between GW and GH was also zero. Tagged fish removed in the Nisg̱a'a harvest downriver of GH ranged from about 4 to about 700 fish, with associated measures of uncertainty.

Because only Chinook Salmon and summer-run steelhead are tagged at both fishwheels, loss of tagged fish to escapement between GW and GH is only relevant for these species. The proportion of fish that pass GW and spawn between GW and GH comes from genetic analysis on the biological samples taken at GW over the year. The posterior distributions from the hierarchical model for the probability of escapement is shown in Figure 2. Details on the hierarchical model to describe the probability of escapement are provided in the technical supplements. Note that this spawning proportion is applied only to the tags applied at GW – we assume that any fish that arrives at GH will not fall back to spawn between GW and GH. As noted previously, we implicitly assume that fish from stocks that traditionally spawn below GW and appear at

GW do not fall back. GSI data indicates that the proportion of the GW run that may belong to “overshooting” stocks is small (on average <3%; Beveridge et al. 2025).

The number of tagged fish lost to initial tagging mortality/fall back was also estimated using a hierarchical model (Figure 3) and applied to fish tagged at both wheels. The uncertainty in the number of tagged fish lost to initial mortality/fall back is large due to the small number of mortality studies conducted and, in some cases, a small number of fish used in the study.

Tag loss is only an issue for Coho and Sockeye salmon. Chinook Salmon are fully inspected for the presence of a tag and the batch mark. Adipose fin clips are used to mark steelhead and are assumed to be permanent. A hierarchical model is fit to years where tag loss studies have been done (a subsample of fish that was initially classified as untagged is re-examined for lost tags) and are summarized in Figure 4. The overall posterior distribution (left-most distribution in Figure 4) is used to estimate the proportion of tagged fish that lose their tags for years not part of the set of studies, e.g., prior to 2012 for Sockeye Salmon. For years in which a study is conducted, the posterior distribution for that study is used. In cases of small sample sizes in a year, information is shared across years. For example, the estimated tag loss probability for Coho Salmon in 2017 where no sub-sampled fish were observed to have lost a tag is based on the entire set of data and so shrinkage in the estimated proportion towards the overall mean takes place.

The estimated total removals of tagged fish until just above GH includes all sources listed above (when applicable) (see Table 1). This is the key difference between the Bayesian model and the past reports where estimates of the number of tagged fish are found using similar methods, but the uncertainty in these removals was not estimated.

This then gives rise to the posterior distribution for the number of tagged fish available just upriver of GH (Table 1). Fish are inspected at MF, KW, DC, and CR (or a subset) and the total number of fish inspected and the number of tags found is available.

The observed data and the prior distribution for the marked fraction gives rise to the estimated marked fraction at GH. The prior distribution for the marked fraction will have an impact when the number of fish inspected and tags recovered is smallish (e.g., the raw marked fraction for steelhead is 7/94 = 0.074 vs. the posterior distribution mean of 0.080).

The Bailey (1951, 1952) model gives the estimated run size just upriver of GH to which is added the fishwheel harvest at GH (Table 1). The estimated uncertainty from the Bayesian model is larger than the estimated uncertainty in the past reports because the latter did not account for the uncertainty in the number of fish with tags that are “removed” prior to arriving at GH based on an estimate of GW tagged fish spawning between GW and GH.

For Coho and Sockeye salmon, the run size estimate at GH includes escapement between GW and GH. For these species, the run size at GW is found by adding the sport harvest between GW and GH, the estimated Nisg̲a’a harvest between GW and GH, and the GW fishwheel harvest to obtain the estimated run size at GW (Table 1).

The uncertainty of the run size at GW will be slightly larger than that at GH because of the uncertainty in these additions.

For Chinook Salmon and summer-run steelhead, the run size estimate at GH excludes escapement between GW and GH. The estimated sport and Nisg̲a'a harvest between GW and GH is added to the estimated run size at GH. This is then expanded by (1-p(escapement))[1]. For example, in 2021, the Chinook Salmon estimated run size at GH is 10.8 (SD 1.1) thousand fish. We add the sport harvest (175 [SD 35] fish) and the Nisg̲a'a harvest (1,014 [SD 182] fish) which gives 12.0 thousand fish, which is expanded by (1-0.108) to give a GW run size of 13.5 (SD 1.3) thousand fish (Table 1). Uncertainty in the sport and Nisg̲a'a harvest and the proportion of fish that escape between GW and GH is automatically propagated to the estimates of run size.

The posterior densities for the run size at GW and GH are shown in Figure 5 which can be used to derive credible intervals or other statistics of interest as shown in the supplemental materials. The posterior density is approximately normally distributed around the posterior mode and so the approximate credible intervals computed as the mean ± 1.96 SD should be a reasonable approximation.

The results from the unstratified Bayesian model can be compared to the estimates obtained using the current non-Bayesian methods (described in Schwarz et al. 2025) and summarized in Table 2. The gross number of tags applied at GW and GH matches (as it must).

The estimated fish with tags removed in the Nisg̲a'a harvest is comparable between the two sets, except for Coho Salmon. In 2021, there were no observed fish with tags returned in the ISF because the Nisg̲a'a Nation did not conduct in-river sale fisheries. The point estimator used in the reports would give a value of 0 for the proportion of tags present in the ISF, so the values for tag loss in Nisga'a fisheries (reported in Tables 1 and 2 are from food fisheries only. The Bayesian method would merge these meagre data with the prior and would "generate" some fish with tags that were removed in the ISF fishery.

Previous reports (prior to 2020) make no adjustment for fish with tags that spawn between GW and GH. This would have affected mainly the Chinook Salmon estimate as the estimated proportion that spawn between GW and GH is considered small (<3%) for summer-run steelhead (Beveridge et al. 2024). The differences in the number of fish with tags removed due to initial mortality is due to the report using probabilities that are slightly different than indicated by the hierarchical model used for initial mortality probabilities. The differences in the number of fish with lost tags is due to the report only using the results from the 2021 subsample, but the Bayesian methods shrink the years' results towards the hierarchical mean because of the small sample sizes (Figure 4). This gives rise to the difference in the net number of tags available just upriver of GH which has only minor effects on the estimated run size. However, the uncertainty in the

---

[1] For example, if there are N_{GW} fish returning at GW and a fraction p escape between GW and GH, then there are $N_{GH} = N_{GW}(1-p)$ fish returning at GH. Conversely, when working backwards, if the run size at GH is N_{GH}, then the run size at GW (after adjusting for escapement) is $N_{GW}=N_{GH}/(1-p)$.

net number of tags available just upriver of GH affects the estimated uncertainty of the run size estimates, especially for Sockeye Salmon.

Why is the increase in uncertainty so large for Sockeye Salmon compared to the other species? The estimated run size at GH for Sockeye Salmon in 2021 is approximately equal to

$$\widehat{N} \cong \frac{mean(Net\ tags\ at\ GH)}{mean(marked\ fraction)} = \frac{6462}{.0169} = 382{,}366$$

which is a close approximation to the actual estimate in Table 1.

Both the numerator and denominator are random variables. The standard deviation of the ratio can be approximated by:

$$SD(\widehat{N}) \cong N\sqrt{\left(\frac{SD(Tag)}{mean(tags)}\right)^2 + \left(\frac{SD(marked\ fraction)}{mean(marked\ fraction)}\right)^2}$$

In the non-Bayesian methodology, the number of tags available just upriver of GH is assumed to be known exactly, and so its variance is zero. The above equation then gives:

$$SD(\widehat{N}_{report}) \cong 382366\sqrt{\left(\frac{.0002}{.017}\right)^2}, = 4{,}498$$

for a RSE of 1.1% which is close to the reported standard error for this estimate.

When the uncertainty is approximate we have:

$$SD(\widehat{N}_{Bayesian}) \cong N\sqrt{\left(\frac{1024}{6462}\right)^2 + \left(\frac{.0002}{.017}\right)^2} = 60{,}924$$

for a RSE of 16% which is close to the uncertainty for the Bayesian estimate.

So, while the relative uncertainty of the number of tags available just upriver of GH is comparable across all species, the relative uncertainty of the marked fraction is much smaller for Sockeye Salmon compared to the other species. The former then swamps the latter leading to a (relatively) much larger increase in the uncertainty for Sockeye Salmon.

The above analysis shows that the key to improving the uncertainty of the Bayesian estimator for the run size lies in both increasing the number of inspected fish (and leading to a more precise estimate of the marked fraction), and also in decreasing the uncertainty in the estimated number of tags available just upriver of GH. A variance decomposition analysis (see multi-year analysis section) shows that the largest contribution to the uncertainty in the estimated number of tags available just upriver of GH comes from the estimate of initial tagging mortality. The initial mortality information is based only on a small number of studies and so it is difficult to estimate the potential year-to-year variation in the initial mortality probability.

2.5.2 Stratified model (Fish length)

The 2021 data were also stratified by fish length. A summary of the results from the stratified analysis of the 2021 data for the four species is found in Table 3.

Table 3 is structured similarly to that from the unstratified analysis except that now data and estimates are separated into two fish length strata for each species, with stratum breaks varying by species (Table 3). The numbers from the two strata may not add exactly to those from the unstratified analysis because some fish that were handled were not measured. In addition, especially for Sockeye Salmon, a sub-sample of tagged and inspected fish was used to allocate the total fish to the two strata. Because the sub-sample was often large (over 1,000 fish), no additional uncertainty has been modelled for this allocation process.

The total number of tagged fish captured in the Nisg̲a’a Harvest upriver of GH was stratified based on a sub-sample of tagged fish. Because this sub-sample was often small, the uncertainty in this allocation process has been included.

The probability of spawning between GW and GH, the probability of initial mortality, and the probability of tag loss were assumed to be the same for both strata, so the hierarchical model from the unstratified analysis was applied to each stratum separately.

The estimated total removals of tagged fish until just above GH includes all sources listed above (when applicable) (see Table 3). This is the key difference between the Bayesian model and the past reports where moment estimates of the number of tagged fish are found using similar methods, but no uncertainty in these removals was accounted for.

This then gives rise to the posterior distribution for the number of tagged fish available just upriver of GH (Table 3). Fish are inspected at Meziadin Fishway, Kwinageese Weir, Damdochax Creek, and Cranberry River (or a subset) and the total number of fish inspected and the number of tags found is available.

The observed data and the prior distribution for the marked fraction gives rise to the estimated marked fraction at GH. In addition, a prior for the difference in the marked fraction between the two strata will “move” the estimated marked fractions closer together when data are sparse. For example, the raw marked fractions for the two strata for Chinook Salmon are 59/256 = 0.230 and 224/1,308 = 0.172 are “moved” closer to reflect the smaller sample size in the first stratum.

Separate estimates of the run size just upriver are found in the same way. These are then combined before the (unstratified) GH fishwheel harvest, the (unstratified) sport harvest between GW and GH, and the (unstratified) Nisg̲a’a harvest between GW and GH are added, and the (unstratified) adjustment for escapement is applied. For this reason, estimates of run size at GW are not available on a stratified basis.

2.5.3 Comparing estimated run sizes between unstratified and stratified model

The estimated run sizes from the two models are compared in Table 4 with a comparison of the density of the posterior distributions shown in Figure 6.

Estimates are comparable between the unstratified and stratified models given the uncertainty in each of the estimates. This is not surprising because the posterior distribution for the difference in the marked fraction between the two strata was very closely centered about 0 (see supplemental material) so that stratification would be expected to have little benefit in reducing the bias in the unstratified estimate relative to the uncertainty in the estimate. Figure 6 illustrates that the density of the posterior distribution for the run size at GW and GH is comparable between the two models. The multi-year study shows that stratification appears to confer little benefit in improving the estimates of the run size at GW and GH (or the SD as well) for Chinook Salmon. The marked fraction in the two strata does not appear to differ substantially over the years of the study (refer to supplemental material), so stratification would not provide any benefit and the "auto-pooling" of the Bayesian methods would automatically move the stratified estimate close to the unstratified estimate (refer to supplemental materials).

3. Bayesian model applied to multiple years

The Bayesian models were applied to the capture-recapture data for the four species from 1994 to 2024 (NFWD 2025). The key purpose of this exercise was to see the impact of the uncertainty in the net number of tags available just upriver of GH that was not accounted for in past analyses. Accounting for the uncertainty in the number of tags available is expected to increase the uncertainty in the estimated run sizes.

Tag types, tag locations, tag loss, removal methods, and other aspects of the studies have changed over the years, so results may not be directly comparable across years. Additionally, the Bayesian model uses more recent data on initial tagging mortality probabilities, tag loss probabilities, escapement probabilities, etc., and again the results from the hind-casting may not be directly comparable to the previously reported results.

Many iterations in the fitting process were used to correct data anomalies that have crept in over the years (e.g., transpositions) and where the data did not match previous reports' summary tables. Further explorations have been performed to investigate the causes, these anomalies generally originate from preliminary estimates soon after the end of the field program that are subsequently refined and updated in the post-season. Full details are provided in the supplemental multi-year reports that use the most updated data releases available for Nass Sockeye (Alexander et al. 2025), Chinook (Beveridge et al. 2025), Coho (NFWD 2025) salmon, and summer-run steelhead (Beveridge et al. 2024). Only the final summary results comparing the estimates and uncertainty from the Bayesian models and the non-Bayesian models are presented here.

3.1 Unstratified model

The gross number of tags applied at GW and GH for the historical re-analysis is shown in Table 5. Refer to description in previous section. The net number of tags just upriver of GH is found for each year and species as described previously. The results for the hierarchical models to estimate the initial tagging mortality probability, escapement between GW and GH probability, and tag loss probabilities were presented in the previous section and will be consistent across all the reanalysis years. These hierarchical models were also used in years prior to the data being collected which will also affect the earlier estimates. A comparison of the net number of tags available just upriver of GH between the Bayesian model and the previous analysis is shown in Figure 7.

The estimated number of fish with tags available just upriver of GH is on average about 12% smaller in the Bayesian model compared to the published reports for Chinook Salmon because of the new correction for escapement between GW and GH. The estimated number of tags just upriver of GH are similar for the non-Bayesian and Bayesian approaches for Coho Salmon (-3% average difference). The Bayesian model estimates about 10% fewer fish with tags just upriver of GH for Sockeye Salmon compared to the published reports due to revised mortality and tag loss probabilities being applied. The estimated number of net tags are similar for the older and newer approaches for steelhead (-5% average difference).

The Bayesian model provides estimates of uncertainty in the net number of tags available just upriver of GH due to the various sources (Figure 8). In years prior to when GSI data are available, the uncertainty in the escapement of Chinook Salmon between GW and GH dominates, but when GSI data are available, the uncertainty is markedly reduced. For the other species, the uncertainty in estimating the initial mortality predominates, followed by uncertainty in the proportion of tag loss for Coho Salmon. All other sources of variation have a relatively small contribution except for the occasional year when something “odd” happened – details are available in the supplemental document (e.g., no compliance monitoring in the ISF during recent years due to no sale fisheries conducted).

The estimated run size at GH produced by the Bayesian unstratified model is compared to the reported non-Bayesian estimates in Figure 9. The estimated run size at GH for Chinook Salmon tends to be smaller than the report estimates because the Bayesian model excludes escapement between GW and GH while the reports did not. Estimates are similar for Coho Salmon. Estimates from the Bayesian model tend to be smaller for Sockeye Salmon because of the different initial mortality and tag loss parameters used – refer to the supplemental information. Estimates for summer-run steelhead are sparse because the number of recoveries is often small, but in general, estimates track each other well. The estimated uncertainty of run size estimates is larger in the Bayesian model as expected.

A similar plot for the estimated run size at Gitwinksihlkw is shown in Figure 10. The estimates for Chinook Salmon now include escapement in both cases and so there is close match. Estimates for Sockeye Salmon are lower in earlier years as noted previously primarily because tag loss was not estimated and accounted for until 2004.

Finally, the estimates of uncertainty from the Bayesian model are compared to the current estimates of uncertainty in Figure 11. The uncertainty from the Bayesian model tends to be larger than the uncertainty from the current methods that did not account for uncertainty in the number of tags available just upriver of GH. Estimates of uncertainty for Chinook Salmon from the Bayesian model tend to be about 1.3x larger than those from the reports; 1.6x larger for Coho Salmon; 4.6x larger for Sockeye Salmon; and 1.2x larger for summer-run steelhead. The supplemental material has further explanations for “odd” years seen in the plot.

3.2 Stratified model (fish length)

A length-stratified Bayesian model was fit to the multiyear dataset for all four species. In some cases, the total number of tags recovered was less than 10 or both strata did not have at least four tags and estimates were not computed. For each year and species, the posterior distribution for the odds-ratio of the marked fraction between the two strata was examined to see if there was evidence of a difference in the marked fractions (which would indicate that stratification is needed).

A comparison of the estimated total run size at GH between the unstratified and stratified estimates is presented in Figure 12. The two estimates track each other quite well over time and there does not appear to be a systematic difference between the stratified and unstratified estimates even in those years where there was evidence of a difference in the marked fraction. The graph comparing run sizes at GW is similar and not shown. This may indicate that the tagging program at GW and GH are working as intended and selecting a constant fraction of the run over the season for tagging.

The comparison of estimates between the unstratified and stratified model is more readily seen in Figure 13. The estimates are very similar for the most part. There does not appear to be a large difference in estimates even when there was evidence that the marked fractions differed in the strata.

Finally, Figure 14 compares the estimated uncertainty (SD of the estimates) between the unstratified and stratified Bayesian estimates. The estimated uncertainties are very similar except for Coho Salmon where it again appears to be driven by a few years.

The key message from the above analysis is that length stratification appears to be of limited usefulness in reducing bias or reducing uncertainty. Some cost savings could be realized by reducing effort in stratifying fish.

4. Temporal stratification

In the previous stratified model (e.g., by length), fish cannot change strata over the course of the study. In temporal stratification, releases, inspections, and recaptures are stratified by the week of release and recapture, and fish released in one temporal stratum may distribute themselves among several recovery strata (weeks of recovery).

For Sockeye Salmon, different coloured tags are applied in different statistical weeks. For Coho Salmon, individually numbered tags are read at the Meziadin Fishway. In both methods, the week the fish was tagged can be identified and the time of recapture or inspection is also noted. Data are too sparse, or tags applied (e.g., fin clips for steelhead) are not conducive to a temporal stratification for other species.

In the current methodology, the number of tags applied at GW is adjusted to give the number available just upriver of GH for each stratum of release using similar methods as described in Part I of the report, i.e., estimated number of fish with tags that are removed due to initial mortality and fall back, Nisg̲a'a and sport harvest and tag loss is computed for each weekly release. No adjustment is made for spawning between GW and GH.

Data are available for Coho Salmon between 2002 and 2011 and for Sockeye Salmon between 2000 and 2024.

There are two common estimation methods for this type of stratified data. The first is the estimator of Darroch (1961) that was generalized by Schwarz and Taylor (1998) and Schwarz (2024). This method was programmed into the Stratified Petersen Analysis System (SPAS) that used to be implemented in a standalone Windows program but is now available as an *R* package (Schwarz 2024). The SPAS estimator does not account for the temporal structure of the data and often requires extensive pooling (e.g., combining rows and/or columns) to deal with non-identifiability concerns. Consult the supplemental materials for more details.

More recently, Bonner and Schwarz (2011) created an estimator that accounts for the temporal nature of the data and provided the Bayesian Time Stratified Petersen Analysis System (BTSPAS; Bonner and Schwarz 2024). Some preprocessing of the data will be needed to address studies that start in the middle of the run. More details are provided in the supplemental materials.

The following methods will be used to estimate the run size just upriver of GH (plus spawning between GW and GH) for Coho and Sockeye salmon:

- a Pooled-Petersen (Chapman) estimator
- a Stratified-Petersen estimator using the *R* package *SPAS*
- a Stratified-Petersen estimator using the *R* package *BTSPAS.*

None of the measures of uncertainty for the above estimates account for the uncertainty in the adjustment process (see Section 2.1) and so the reported uncertainty is less than the actual uncertainty as shown in the previous sections of this document.

An average measure of the proportion of tagged fish that are "removed" is computed using the sum of the weekly values. A fourth estimator can often be computed

- a Stratified-Petersen estimator accounting for fallback/loss using the *R* package *BTSPAS*

The supplemental material should be consulted for more granular details.

4.1 Coho Salmon

A comparison of the estimators computed using *SPAS* and the simple pooled-Petersen estimator is show in Figure 15. The SPAS estimate tends to be slightly smaller than the pooled-Petersen estimate, but this appears to be driven by one estimate in 2002. A comparison of the standard errors (Figure 16) shows that the standard errors computed by SPAS tend to be about 1.3x larger than those from a pooled-Petersen estimate. Given the small difference in estimates and increase in uncertainty, temporal stratification does not appear to be worthwhile. Again, note that these reported standard errors do not account for the uncertainty in the adjustment process.

BTSPAS estimates cannot be computed because the temporally stratified data are too sparse on a weekly level. It may be possible to compute the BTSPAS estimates if data are pooled to 2-week strata, but given the small difference in the SPAS weekly estimates to the unstratified estimates the results are unlikely to provide additional insights. Consult the supplemental material for more details.

4.2 Sockeye Salmon

A comparison of the estimators computed using *SPAS* and the simple pooled-Petersen estimator is shown in Figure 15. Both estimators are similar and there is no obvious choice between them. A comparison of the standard errors (Figure 16) shows that the standard errors computed by SPAS tend to be about 1.6x larger than those from a pooled-Petersen. Given the small difference in estimates and increase in uncertainty, temporal stratification does not appear to be worthwhile. Again, note that these reported standard errors do not account for the uncertainty in the adjustment process.

The estimate of run size at GH computed using BTSPAS are compared to the pooled-Petersen estimates in Figure 17. The estimates are similar (except for 2002, which is related to a peculiarity in the data). There is again no overwhelming reason to choose the temporally-stratified estimator. The comparison of the measures of uncertainty shows a large increase in uncertainty (about 5x) when using the BTSPAS estimators (Figure 18). Again, note that these reported standard errors do not account for the uncertainty in the adjustment process.

Given the above results, we do not report on the results from BTSPAS adjusting for fallback, but details are available in the supplemental materials.

In short, there does not appear to be any benefit to temporal stratification for Sockeye Salmon.

5. Adding additional recovery locations

A critical feature of a successful capture-recapture study is the number of tagged fish that are recaptured because precision is basically a function of $\sqrt{marks\ recaptured}$. The number of tagged fish recaptured can be increased by increasing the number of tags applied or increasing the recovery effort or a combination of both.

It seems unlikely that additional tagging effort can be done at the Gitwinksihlkw fishwheels for Chinook or Coho salmon or summer-run steelhead as the majority (>80%) of these fish of these species captured are already tagged. Indeed, the second set of fishwheels at Grease Harbour were added because of an inadequate number of tagged Chinook Salmon and steelhead from the first set of fishwheels (i.e., Gitwinksihlkw). Additional tagging could be done for Sockeye Salmon because not all Sockeye Salmon captured in fishwheels are tagged with only a target of 300 tagged fish/day at the wheels.

It is easier to add additional recovery locations. Ideally, these would be located on the mainstem before individual stocks split off to their respective tributaries. If additional recovery locations are located at terminal streams, a critical assumption is that the fishwheels were operating at roughly the same efficiency for the entire run so that the stock returning to the terminal streams is sampled (and tagged) at roughly the same rate as other stocks. Mark rates will generally be lower for stocks tagged later in season. Chinook Salmon all pass the fishwheels in about three weeks so are believed to have similar mark rates, but Sockeye and Coho salmon have wide distributions in run timing. For Sockeye Salmon, this means early timed fish can have different (higher) mark rates than later-timed fish due to differences in water levels and catchability at the Gitwinksihlkw fishwheels. Coho Salmon and summer-run steelhead will typically be captured at similar fishwheel efficiencies from later July and August, with lower catchability in September (low water).

Modifying the Bayesian analysis for additional recovery locations is trivial. The simplest approach would be to simply add over all recovery locations and use the totals over all recovery locations in the model. Or, additional Bailey (1951, 1952) Binomial components can be added to the model.

Notice (Figure 8) that the largest component in the uncertainty in the net number of tags available just upriver of GH is due to the uncertainty in the initial mortality probability. Adding additional tagging effort at the fishwheels or additional recovery effort upstream will not impact the uncertainty in the initial mortality probability and so it may be a better use of resources to reduce the uncertainty in the initial mortality probability by conductions additional mortalitity studies with increased sample sizes.

5.1 Addition of Kwinageese Weir recoveries for Sockeye and Coho Salmon

At the moment, the Bayesian model only uses the recoveries of fish at the Meziadin Fishway (MF). Additional data have been collected at the Kwinageese Weir (KW) in 2002, 2005, 2006 and every year since 2009 (

Table 6). In most years, the number of fish inspected at Kwinageese Weir is much smaller than the number inspected at Meziadin Fishway and so, a priori, the impact of including these additional fish is expected to be small.

If the data from KW are to be "pooled" with that from MMF, the marked fractions should be similar. The marked fraction at each site is compared in Figure 19. The marked fraction at MF larger in most years (about 1.5x larger on average) than the marked fraction at KW. Dominant Meziadin stocks (Hanna/Tintina) do pass earlier in the season under higher water and ideal capture conditions (i.e., higher marked fraction) than the later-timed Kwinageese stock that passes under lower water and less ideal capture conditions (i.e., lower marked fraction) that may explain some of the differences observed (Figure 20). A similar mark-rate difference has been observed for Coho Salmon at KW where the marked fraction is also larger at MF (1.3x larger than average) from 2010-2025 using anchor tags. These mark rate differences may be due to additional tag loss of spaghetti and anchor tags between MF and KW or differential tagging at the fishwheels for fish inspected at MF and KW.

Tag loss data were also collected at the two locations. At MF, a sub-sample of fish was physically handled and inspected for tag loss, while at KW, most fish were inspected from video review. A comparison of the estimated tag loss proportions is shown in Figure 21. The overall tag loss proportion is about 2x larger at MF than KW. This seems contrary to intuition where sites further upstream from the tagging sites would be expected to have a higher tag loss proportion. However, the difference may be in the methods of observing tag loss from visual images versus physical handling for tag loss.

The estimated run size at GH for Sockeye Salmon was compared when only the MF recoveries were used, and when the KW recoveries were added to the MF recoveries (Figure 22). The estimates are very similar to the estimates based on MF alone being 0.99x that of the estimates based on MF and KW. The estimated SDs are also comparable, and similar findings occur for the estimated run size at GW (not shown but see Schwarz et al. 2025).

This analysis shows some contradictory results:

1. The marked fraction is lower at KW than at MF. This could arise because of additional tag loss between MF and KW and different capture probabilities of the later timed stock; and
2. The estimated cumulative tag loss at KW is lower than the cumulative tag loss at MF. Because KW is further upstream of MF, the reverse would be expected but the video estimates may not truly detect all fish passing that were missing a mark.

The estimates computed using MF and then using the MF and KW show small differences typically less than 2%, with the estimates using data from MF and KW slightly larger than the estimates using data only from MF only. This is expected because adding the KW data with a lower marked fraction will dilute the marked fraction at MF leading to a larger population estimate. However, the addition of the data may be an appropriate indicator of later timed stocks that pass the fishwheels when marked fractions are lower due to lower water levels.

The difference is also very small relative to the uncertainty in the estimates. This result was not unexpected given the very small number of fish inspected at KW relative to inspections at MF.

6. Summary and recommendations

6.1 Accounting for uncertainty in the adjustment to the number of tags available for recapture upriver of Grease Harbour

Part I (Schwarz et al. 2025) of the review found that physical methods (e.g., tag types, placement of fishwheels, studies to estimate tag loss, etc.) used in estimating the run size in the Nass system were generally acceptable. The capture-recapture estimation methods currently used are expected to give reasonable point estimates (i.e., should be close to being unbiased), but the reported measures of uncertainty (the reported standard errors) are likely too small because the current method does not account for uncertainty in the estimated net number of tags available for recapture upriver of Grease Harbour.

The key reason for adopting the Bayesian capture-recapture model is to report more realistic estimates of the uncertainty in the run size. This is particularly true for Sockeye Salmon where the current estimates of uncertainty are too small (<2% for most years) by a factor of around 5x!

Further, the Bayesian model correctly accounts for the removal of tagged fish due to escapement between GW and GH for Chinook Salmon and summer-run steelhead. Notice that estimates of run size at GH from the Bayesian model for Coho and Sockeye salmon include escapement between GW and GH.

The estimates of run size, when the Bayesian methods are hind-casted, track the estimates from the yearly reports quite well. Some caution is needed because estimates from earlier years were recomputed using information just recently collected (e.g., tag loss, genetic studies) and so the two sets of estimates are not a like-on-like comparison. Nevertheless, the Bayesian estimates appear to give similar estimates of run size as the current reports when hind-casted, but, as noted earlier, the estimated uncertainty in the estimates from the Bayesian approach is larger.

6.2 Stratification by length or sex

Heterogeneity in tagging probabilities or recapture probabilities among fish can lead to bias in estimates. Pure heterogeneity (e.g., larger fish tend to have higher tagging probabilities and higher recapture probabilities compared to smaller fish) leads to a negative bias in estimates of abundance (Schwarz and Taylor 1998). Stratification by a fixed characteristic (e.g., length or sex) could reduce bias in estimates.

A length-stratified Bayesian model was also implemented and applied to the past years data. Estimates of abundance were similar between the stratified and unstratified approaches indicating that the effect of pure heterogeneity appears to be negligible. This is particularly true when the increased uncertainty from accounting for the estimation of the number of tags available is incorporated.

One advantage of the Bayesian stratified model is that it is “self-adjusting”. If the estimated mark-fractions in the two strata are similar and based on small sample sizes (e.g., for Chinook Salmon), the Bayesian model will shrink the individual marked fractions towards a common marked fraction. This sharing of information will also reduce the uncertainty compared to a strict stratified estimator.

Consequently, it appears that stratified analyses (e.g., by length) are not needed. Collecting length data has a cost and the cost savings from not stratifying can be used elsewhere. This is true even for Coho and Sockeye salmon where the numbers in each stratum are estimated based on sub-sampling rather than complete enumeration. The sub-sample is typically large enough that the uncertainty from sub-sampling can be ignored. If both a stratified and unstratified analysis are available, it is sensible to choose the estimator with a smaller uncertainty.

We did not examine stratification by sex, but similar methods can be used, and the current Bayesian model can be used unchanged by replacing the two length classes by the two sexes.

6.3 Temporal stratification

Unlike fixed-attribute stratification, temporal stratification is useful when a tagged cohort (week of tagging) are covered over multiple weeks. Again, pure heterogeneity (e.g., a fish earlier in the season has a higher probability of being tagged and recovered) can lead to bias in estimates.

Because fish that are tagged in a cohort can be recovered in several weeks along with fish from another cohort, a temporally stratified estimator needs to be used. There are two current methods for analyzing temporally stratified capture-recapture experiments:

- A general stratified estimator developed by Darroch (1961) and Schwarz and Taylor (1998) and implemented in *SPAS* (Schwarz 2025)
- A specialized temporal stratified estimator developed by Bonner and Schwarz (2011) and implemented in *BTSPAS* (Bonner and Schwarz 2024*).*

The current program uses *SPAS.* However, use of this program is not recommended for temporal stratification because of the need to (arbitrarily) pool rows and columns and its poor performance when sample sizes are small. We recommended that the *BTSPAS* approach be used if a temporal stratification is required because it handles sparse data more ideally than *SPAS,* which shares information across strata. Neither approach can currently handle the uncertainty in the estimated number of tags available just upriver of GH and so the reported measures of uncertainty from BTSPAS will be larger than that of a pooled-Petersen but will still be too small.

However, this above issue is moot given that the temporally stratified estimates are so similar to the non-stratified estimates. As with length stratification, there does not appear to be a good reason for computing the temporally stratified estimators when estimating run size. There may be other reasons where temporally stratified estimates may be required (e.g., in-season estimates) that were not examined in this report.

6.4 Is there a need for stratification?

A pooled-Petersen estimator will be unbiased under one of two common conditions:

1. Homogeneous tagging probabilities
2. Homogeneous recovery probabilities.

The current methodology does extensive testing for equal marked fractions (consistent with homogeneous tagging probabilities) or equal recovery proportions from release groups (consistent with homogeneous recovery probabilities) across strata defined by length, sex, or time of tagging, or time of recovery.

For many cases, there is no evidence of differential marked fractions or recovery proportions indicating that a pooled-Petersen may be suitable. In other cases, there is evidence of a difference across strata, but the sample sizes are large so that quite small differences can be detected. In many cases, the stratified and unstratified estimates are very similar (Figure 12).

The concern about biases introduced by heterogeneity in tagging or recovery probabilities was driven by the apparently very precise estimates of abundance so that small changes due to stratification appear to represent “real” fish. However, the Bayesian model has shown that the reported uncertainties are underestimated (and considerably so for Sockeye Salmon). While stratified estimators can still be computed, the larger uncertainty will likely show that there is no need or a stratified analysis. Consequently, unless data needed for stratification can be collected cheaply, the concerns about heterogeneity appear to be moot. Additionally, for Chinook Salmon, the stratification by length often performs poorly using the current estimation methods because of the small sample sizes in some strata. The Bayesian model is “self” adjusting so that if sample sizes are small, apparent large difference in the marked fractions due to small sample sizes (e.g., a marked fraction of 2 marked fish out of 5 (0.40 marked fraction) appears to be quite different than 100 marked fish out of 500 inspected fish (0.20 marked fraction), but the Bayesian method will recognize that the stratum with small sample sizes should be shrunk to the overall mean automatically. Consequently, the Bayesian model is intermediate between a complete stratification and complete pooling and will automatically deal with the cases of very small sample sizes in one stratum.

6.5 Reducing uncertainty in the estimates

In standard capture-recapture studies, the uncertainty of the estimators is a function of $\sqrt{marks\ recaptured}$. However, in this study, we need to account also for the uncertainty in the number of tagged fish available for recapture above Grease Harbour. This uncertainty can quickly overwhelm the uncertainty induced by the number of tagged fish recaptured as seen for Sockeye Salmon, for which the classical measure of uncertainty is a small fraction of the total uncertainty.

In particular, the largest contribution to the uncertainty in the number of tagged fish available just upriver of GH is the estimation of mortality/fall back of tagged fish after release (Figure 8). For most species, there are only a few initial mortality studies with smallish sample sizes conducted many years ago. It is recommended that additional initial mortality/fall back studies be conducted to reduce this contribution to the overall uncertainty. This is particularly true for Sockeye Salmon.

6.6 Adding additional recovery locations

Adding additional recovery locations is straight forward in the Bayesian model. However, unless the sample sizes are comparable to the other locations, the benefit may be minor. As well, before adding additional locations, a careful examination of the marked fraction and the tag loss probability should be undertaken and reconciled.

Assuming that tagging is roughly homogeneous over the run, there is no technical reason not to include as many recovery locations as possible. However, the current models show that the limiting factor for uncertainty is the estimation of the initial mortality probability, so there may be little improvements in uncertainty when additional recovery locations are included.

In particular, there is little effect on the estimates of run size or on the estimates of uncertainty of the run size when adding additional data from Kwinageese Weir for Sockeye Salmon.

6.7 Dealing with tag loss

In the current methodology, tag loss proportions are estimated only using data for that year. Unfortunately, this leads to estimates of tag loss with poor precision because sample sizes are somewhat small.

The Bayesian model uses a hierarchical model to share information on tag loss across years. This is especially useful in the case of small sample sizes (refer to Figure 4). Currently, 10+ years of data on tag loss have been collected for Coho and Sockeye salmon and it appears that tag loss proportions are relatively stable over time. As well, Figure 8 shows that the uncertainty due to tag loss is now a small portion of the total uncertainty in the number of tagged fish available upriver of GH, especially for Sockeye Salmon.

Currently, tag loss data are collected at each recovery station. Assuming that the tag loss probabilities are not too dissimilar, the data can be aggregated over multiple locations and the aggregated data used.

Consequently, it may not be necessary to collect tag loss data for every year. For example, tag loss studies could be conducted every second year, when a new type of tag is introduced, or new tagging crews are trained in applying tags released. The cost savings could be used to (partially) fund the initial tagging mortality study.

7. References

Alexander, R. F., I. A. Beveridge, K. K. English, C. A. Noble, N. Percival, B. Gonu, T. Angus, and B. Adams. 2024. Genetic (DNA) Stock Composition Estimates for Nass River Sockeye Salmon Returns to the Gitwinksihlkw Fishwheels and Preliminary Northern Boundary Run Reconstructions of Nass River Sockeye Salmon Stocks, 2024. Prepared for the Pacific Salmon Commission (NF-2024-I-18), Vancouver, BC, and the Nisg̱a’a–Canada-BC Treaty’s Joint Fisheries Management Committee, Gitlaxt’aamiks and Prince Rupert, BC, by LGL Limited, Sidney, BC, and Nisg̱a’a Fisheries and Wildlife Department, Gitlaxt’aamiks, BC. Nisg̱a’a Fisheries Report #24 19: xvi + 122 p.

Al Amer, F. M., Thompson, C. G., & Lin, L. (2021). Bayesian Methods for Meta-Analyses of Binary Outcomes: Implementations, Examples, and Impact of Priors. *International Journal of Environmental Research and Public Health*, *18*(7), 3492.

Bailey, N. T. J. 1951. On estimating the size of mobile populations from capture-recapture data. Biometrika 38:293–306.

Bailey, N. T. J. 1952. Improvements in the interpretation of recapture data. Journal of Animal Ecology. 21: 120-127.

Beveridge, I. A., R. F. Alexander, S. C. Kingshott, N. Percival, and B. Stewart. 2024. Nisga’a Treaty Projects: Nass River Steelhead, Trout, and Char Assessments, 2023. Prepared for the BC Ministry of Water, Land, and Resource Stewardship, Smithers, BC, and the Nisga'a–Canada-BC Treaty’s Joint Fisheries Management Committee by LGL Limited, Sidney, BC, and the Nisg̱a'a Fisheries and Wildlife Department, Gitlaxt’aamiks, BC. Nisg̱a'a Fisheries Report #23 16: iv + 29 p.

Beveridge, I. A., R. F. Alexander, S. C. Kingshott, C. T. McCulloch, T. Angus, and B. Adams. 2025. Estimating the abundance of adult Chinook Salmon returning to the upper Nass River, BC, using mark recapture techniques, 2024. Prepared for the Pacific Salmon Commission (NF-2024-I-33), Vancouver, BC, and the Nisg̱a'a-Canada-BC Treaty Fisheries Management Committee, Gitlaxt'aamiks and Prince Rupert, BC, by LGL Limited, Sidney, BC, and the Nisg̱a'a Fisheries and Wildlife Department, Gitlaxt’aamiks, BC. Nisg̱a'a Fisheries Report #24-22: xiii + 126 p.

Bonner, S.J. and C.J. Schwarz. 2011. Smoothing population size estimates for time-stratified mark recapture experiments using Bayesian p-splines. Biometrics, 67, 1498–1507.

Bonner, S.J. and C.J. Schwarz. 2024. BTSPAS: Bayesian Time Stratified Petersen Analysis System. R package version 2024.11.1.

Darroch, J. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. Biometrika*, 48*, 241-260.

Link, M. R., and K. K. English. 1997. The 1994 fishwheel project on the Nass River, BC.

Can. Manuscr. Rep. Fish. Aquat. Sci. 2421: xi + 93 p.

Nisga’a Fisheries and Wildlife Department (NFWD). 2025. Nisg̱a'a Fisheries Program Final report of 2024 projects. Nisg̱a'a Fisheries Report NF24-01.

Schwarz, C.J. 2024. SPAS: Stratified-Petersen Analysis System. R package version 2024.1.31.

Schwarz, C.J., R. Alexander, C. Carr-Harris, I. Beveridge, C. Noble, C. May, and L. Warkentin. 2025. Review of run size estimates for the Nass River system: Chinook, sockeye, and coho salmon, and summer-run steelhead: Part 1 – Review of Methods. Can. Tech. Rep. Fish. Aquat. Sci. 3719: xv + 161p.

Schwarz, C.J. and C.G. Taylor. 1998. The use of the stratified-Petersen estimator in fisheries management: estimating the number of pink salmon (*Oncorhynchus gorbuscha*) that spawn in the Fraser River. Canadian Journal of Fisheries and Aquatic Sciences 55: 281-297.

TABLES

Table 1. Summary of statistics from an unstratified analysis of the 2021 data for the four species. Summary statistics from the posterior distribution are denoted by *mean (sd)*.

Parameter	**Chinook**	**Coho**	**Sockeye**	**Steelhead**
Gross tags applied				
GW tagged	625	2,130	9,116	104
GH tagged	1,646	0	0	89
Gross total tags applied	2,271	2,130	9,116	193
Removals of tagged fish				
GW fishwheel harvest	0	0	0	0
GH fishwheel harvest	0	0	0	0
Demonstration harvest	0 (0)	0 (0)	-0 (0)	-0 (0)
Sport fishery GW to GH	-2 (0)	-0 (0)	0 (0)	-0 (0)
Nisga’a Harvest GW to GH	-113 (109)	-4 (5)	-661 (620)	-6 (6)
Escapement GW to GH	-68 (23)			-4 (6)
Initial mortality	-134 (112)	-184 (118)	-1,191 (785)	-17 (11)
Tag loss		-151 (66)	-803 (177)	
Total tagged fish removals	-317 (159)	-340 (136)	-2,654 (1,024)	-26 (15)
Net tagged fish at GH	1,954 (159)	1,790 (136)	6,462 (1024)	167 (15)
Total inspected	1,564	5,592	273,955	94
Total tags found	283	92	4,639	7
Marked fraction at GH	0.1811 (0.0097)	0.0166 (0.0017)	0.0169 (0.0002)	0.0801 (0.0271)
GH fishwheel harvest	0 (0)	0 (0)	0 (0)	0 (0)
Run size at GH (‘000)	10.8 (1.1)	108.9 (14.0)	381.6 (60.7)	2.4 (1.0)
Sport harvest GW – GH	175 (35)	-0 (0)	-0 (0)	-0 (0)
Nisga’a harvest GW – GH	1,014 (182)	756 (146)	12,450 (1,514)	6 (4)
P(escape) GW- GH	0.108 (0.035)			0.032 (0.033)
GW fishwheel harvest	0 (0)	0 (0)	0 (0)	0 (0)
Run size at GW (‘000)	13.5 (1.3)	109.6 (14.0)	394.1 (60.8)	2.4 (1.0)

Table 2. Estimates (SE) for 2021 from reports using non-Bayesian methodology. Run size uncertainty are not adjusted for uncertainty in adjustments in number of tags available at GH.

Parameter	**Chinook**	**Coho**	**Sockeye**	**Steelhead**
Gross tags applied - total	2,271	2,130	9,116	193
Nisga'a harvest GW to GH	-117	-4	-685	-8
Escapement GW to GH				
Initial mortality	-114	-128	-1,094	-10
Tag Loss	0	-213	-1,176	0
Net tags at GH	1,925	1,785	6,301	175
Total fish inspected	1,564	5,592	282,433	94
Total tags recaptured	283	92	4,719	6
GH run size ('000s)	10.6 (0.5)	107.4 (10.7)	377.1 (2.8)	2.4 (0.8)
GW run size ('000s)	14.7 (0.5)	108.5 (10.7)	392.4 (2.8)	2.5 (0.8)

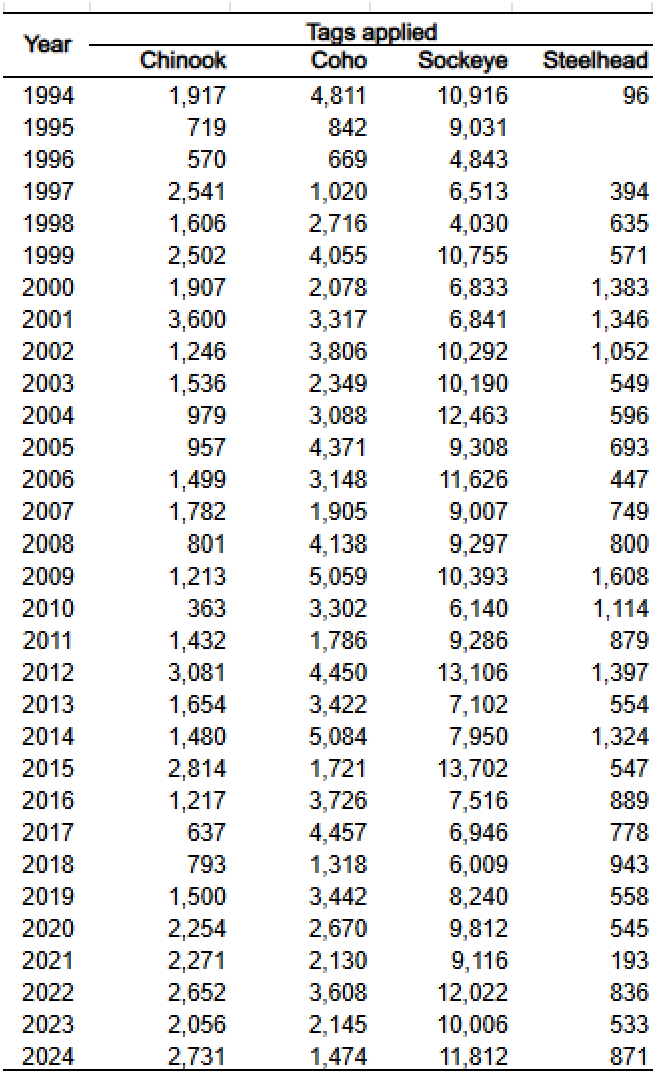
Table 3. Summary of statistics from a stratified (by fish length) analysis of the 2021 data for the four species. Summary statistics from the posterior distribution are denoted by *mean (sd)*. Values for the combined strata are presented in the first stratum columns. Fish lengths provided in millimeters.

Parameter	**Chinook Salmon**		**Coho Salmon**		**Sockeye Salmon**		**Steelhead**	
Stratum (mm NFL)	**500-754**	**754+**	**450-500**	**500+**	**450-580**	**581+**	**450-650**	**650+**
Gross tags applied								
Gross total tags applied	658	1,612	974	1,156	6,851	2,265	277	265
Removals of tagged fish								
GW fishwheel harvest	0	0	0	0	0	0	0	0
GH fishwheel harvest	0	0	0	0	0	0	0	0
Demonstration harvest	-0 (0)	-0 (0)	0 (0)	-0 (0)	0 (0)	-0 (0)	0 (0)	-0 (0)
Sport fishery GW to GH	-2 (2)	0 (0)	0 (0)	0 (0)	-0 (0)	-0 (0)	-0 (0)	0 (0)
Nisga'a Harvest GW to GH	-32 (31)	-80 (77)	-5 (5)	-11 (12)	-496 (463)	-162 (152)	-8 (9)	-8 (8)
Escapement GW to GH	-25 (9)	-43 (15)					-3 (6)	-3 (5)
Initial mortality	-39 (33)	-96 (80)	-84 (55)	-100 (65)	-849 (530)	-281 (176)	-24 (16)	-23 (15)
Tag loss			-70 (31)	-83 (37)	-605 (132)	-200 (45)		
Total tagged fish removals	-98 (46)	-219 (112)	-159 (64)	-194 (76)	-1947 (700)	-642 (231)	-36 (19)	-34 (18)
Net tagged fish at GH	560 (46)	1,393 (112)	815 (64)	962 (76)	4,904 (700)	1,623 (231)	241 (19)	231 (18)
Total inspected	256	1,308	1,660	3,932	170,782	103,173	101	102
Total tags found	59	224	36	56	3,115	1,524	10	10
Marked fraction at GH	0.228 (0.026)	0.172 (0.010)	0.022 (0.004)	0.015 (0.002)	0.018 (0.000)	0.015 (0.000)	0.102 (0.028)	0.098 (0.028)
GH fishwheel harvest	0 (0)		0 (0)		0 (0)		0 (0)	
Run size at GH ('000) (by size class)	2.5 (0.4)	8.2 (0.8)	37.9 (7.4)	63.4 (10.8)	268.9 (38.7)	110 (15.9)	2.6 (0.8)	2.6 (0.8)
Run size at GH ('000) (overall)	10.6 (1.0)		101.3 (14.2)		378.8 (49.2)		5.1 (1.2)	
Sport harvest GW - GH	174 (35)		-0 (0)		-0 (0)		-0 (0)	
Nisga'a harvest GW - GH	1,014 (181)		753 (144)		12,441 (1,526)		6 (4)	
P(escape) GW- GH	0.108 (0.036)						0.030 (0.027)	
GW fishwheel harvest	0 (0)		0 (0)		0 (0)		0 (0)	
Run size at GW ('000)	13.3 (-1.2)		102.1 (-14.2)		391.3 (-49.2)		5.3 (-1.2)	

Table 4. A comparison of the run size estimates at Gitwinksihlkw and Grease Harbour between the unstratified and stratified Bayesian models for 2021. Summary statistics from the posterior distribution are denoted by *mean (sd)*.

Location	**Run size estimate**	**Chinook**	**Coho**	**Sockeye**	**Steelhead**
Grease Harbour	Unstratified ('000)	10.8 (1.1)	108.9 (14.0)	381.6 (60.7)	2.4 (1.0)
	Stratified ('000)	10.6 (1.0)	101.3 (14.2)	378.8 (49.2)	5.1 (1.2)
Gitwinksihlkw	Unstratified ('000)	13.5 (1.3)	109.6 (14.0)	394.1 (60.8)	2.4 (1.0)
	Stratified ('000)	13.3 (1.2)	102.1 (14.2)	391.3 (49.2)	5.3 (1.2)

Table 5.The total number of tags applied at Gitwinksihlkw and Grease Harbour for the re-analysis of historical data using the Bayesian unstratified-model.



Year	Tags applied			
	Chinook	Coho	Sockeye	Steelhead
1994	1,917	4,811	10,916	96
1995	719	842	9,031	
1996	570	669	4,843	
1997	2,541	1,020	6,513	394
1998	1,606	2,716	4,030	635
1999	2,502	4,055	10,755	571
2000	1,907	2,078	6,833	1,383
2001	3,600	3,317	6,841	1,346
2002	1,246	3,806	10,292	1,052
2003	1,536	2,349	10,190	549
2004	979	3,088	12,463	596
2005	957	4,371	9,308	693
2006	1,499	3,148	11,626	447
2007	1,782	1,905	9,007	749
2008	801	4,138	9,297	800
2009	1,213	5,059	10,393	1,608
2010	363	3,302	6,140	1,114
2011	1,432	1,786	9,286	879
2012	3,081	4,450	13,106	1,397
2013	1,654	3,422	7,102	554
2014	1,480	5,084	7,950	1,324
2015	2,814	1,721	13,702	547
2016	1,217	3,726	7,516	889
2017	637	4,457	6,946	778
2018	793	1,318	6,009	943
2019	1,500	3,442	8,240	558
2020	2,254	2,670	9,812	545
2021	2,271	2,130	9,116	193
2022	2,652	3,608	12,022	836
2023	2,056	2,145	10,006	533
2024	2,731	1,474	11,812	871

Table 6. Summary of fish inspected and number of tagged Sockeye Salmon at Meziadin Fishway and Kwinageese Weir.

	Meziadin Fishway			Kwinageese Weir		
Year	**Inspected**	**Tags observed**	**Marked fraction**	**Inspected**	**Tags observed**	**Marked fraction**
1994	158,627	4,794	0.030			
1995	205,853	5,139	0.025			
1996	181,840	3,033	0.017			
1997	158,656	3,098	0.020			
1998	163,998	1,875	0.011			
1999	180,350	6,483	0.036			
2000	137,042	2,964	0.022			
2001	116,192	2,982	0.026			
2002	332,442	6,027	0.018	5,891	86	0.015
2003	196,852	4,650	0.024			
2004	140,923	4,417	0.031			
2005	142,833	3,820	0.027	3,186	37	0.012
2006	146,954	4,694	0.032	2,700	123	0.046
2007	104,308	4,082	0.039			
2008	150,396	5,016	0.033			
2009	168,392	4,887	0.029	107	0	0.000
2010	159,120	2,670	0.017	48	0	0.000
2011	167,524	4,213	0.025	10,273	240	0.023
2012	144,923	6,112	0.042	3,688	141	0.038
2013	170,376	3,726	0.022	397	4	0.010
2014	144,920	2,875	0.020	438	3	0.007
2015	185,917	3,859	0.021	7,044	60	0.009
2016	109,868	2,015	0.018	19,797	244	0.012
2017	119,088	2,482	0.021	7,240	71	0.010
2018	96,827	1,676	0.017	290	2	0.007
2019	88,128	2,021	0.023	6,007	69	0.011
2020	126,468	4,011	0.032	3,256	57	0.018
2021	273,955	4,639	0.017	8,478	80	0.009
2022	343,765	6,905	0.020	5,099	78	0.015
2023	408,462	5,637	0.014	8,490	39	0.005
2024	375,357	5,698	0.015	4,667	36	0.008

FIGURES

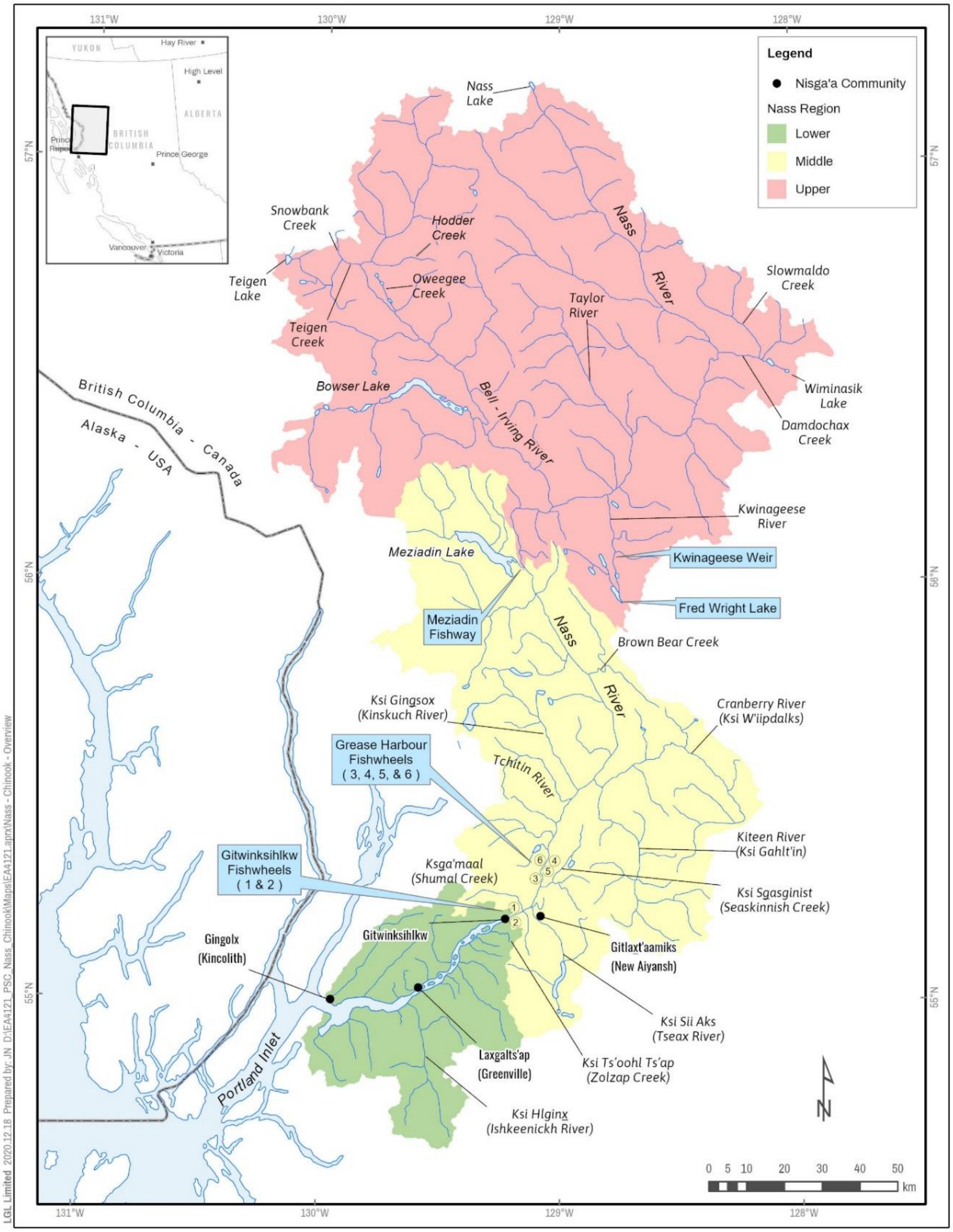


Figure 1. Overview map of the Nass River watershed showing the locations of important stock assessment operations.

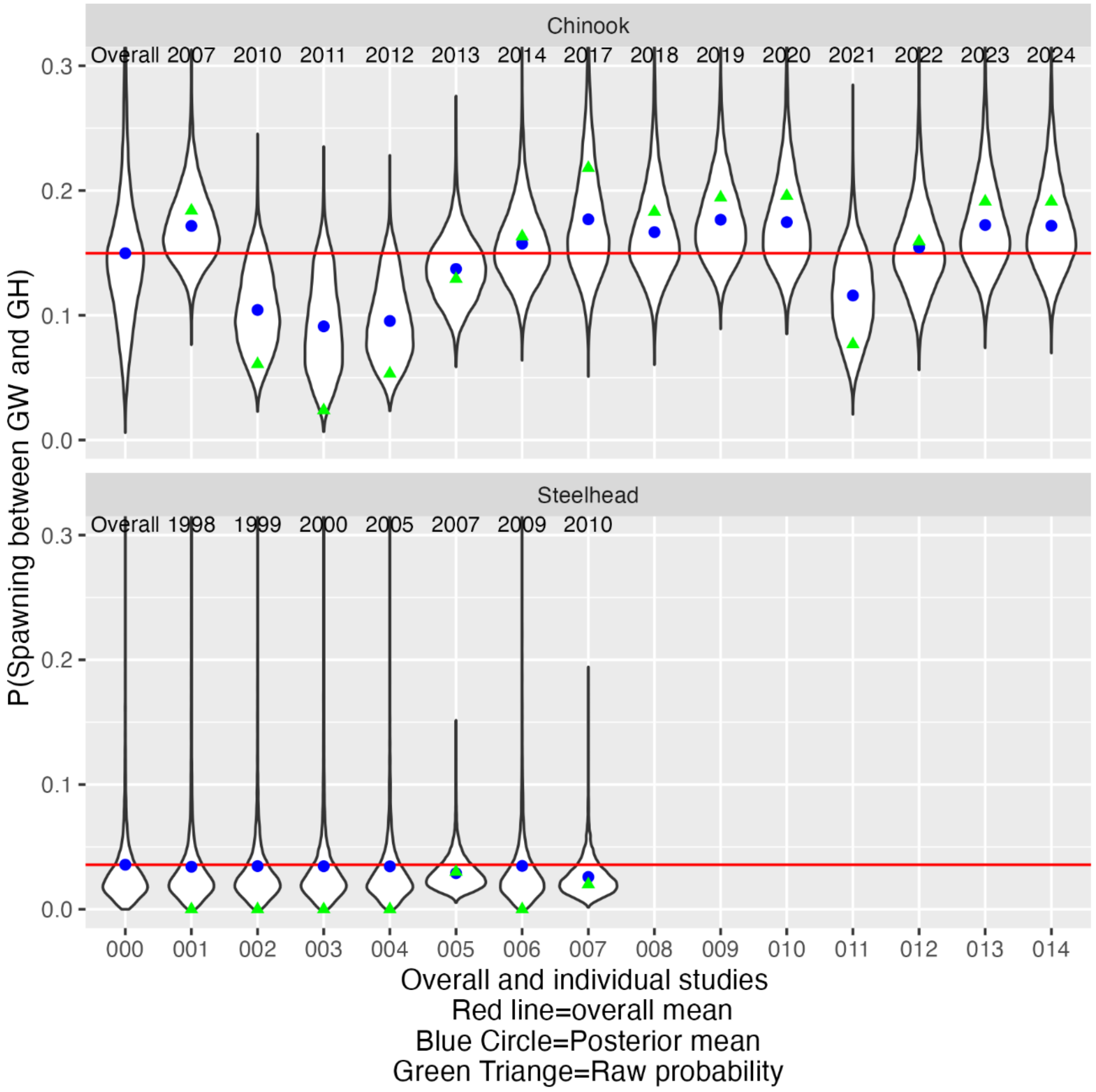


Figure 2. Posterior distribution of the p(spawning) between Gitwinksihlkw and Grease Harbour based on GSI for Chinook Salmon and steelhead. The left-most posterior represents the overall probability of spawning between GW and GH for a random year. Because GSI sample sizes are small in each year, there can be considerable shrinkage towards the overall mean (red line).

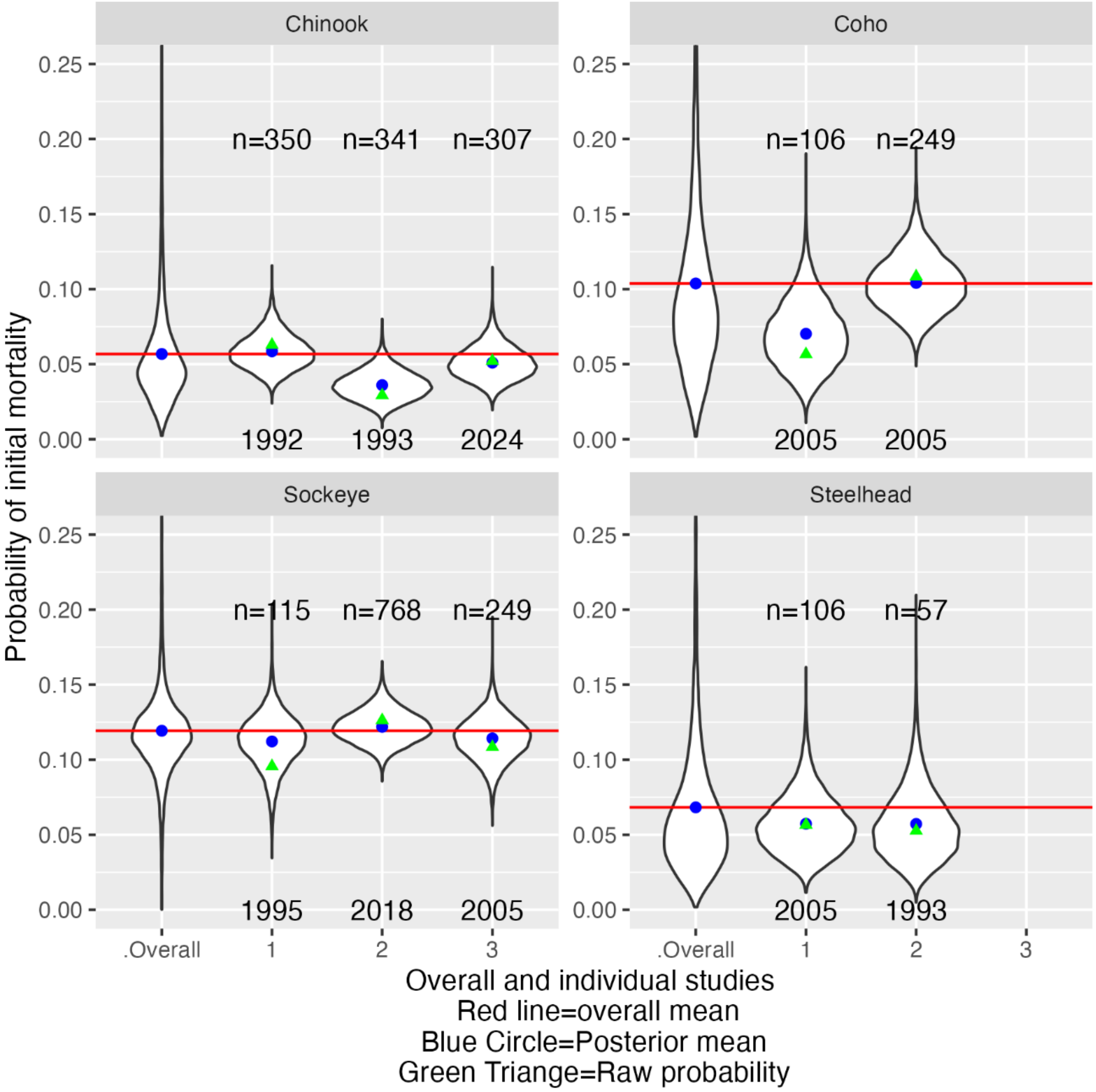


Figure 3. Posterior distributions for the probability of initial mortality or fallback for each species based on a hierarchical model from radio telemetry studies. The studies for Coho Salmon and steelhead are pooled and are identical. The left-most posterior distribution is used to estimate the number of tagged fish that are lost to initial mortality/fall back after tagging.

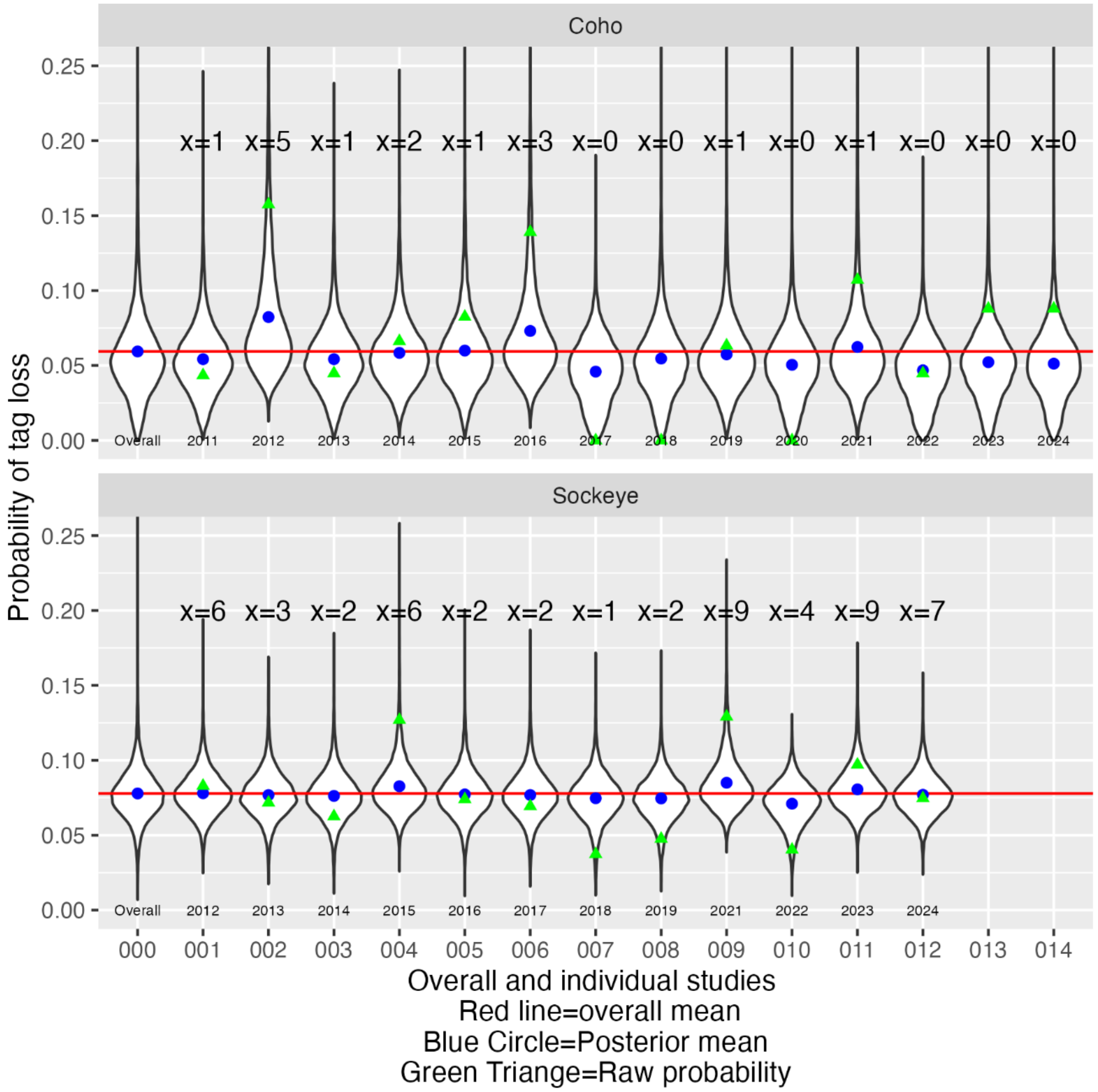


Figure 4. Posterior distributions for the probability of tag loss for each species based on a hierarchical model from recently conducted sub-sampling studies. The left-most posterior distribution is used to estimate the proportion of tagged fish that lose their tags for years not part of the set of studies, e.g., prior to 2012 for sockeye. For years in which a study is conducted, the posterior distribution for that study is used. In cases of small sample sizes, information is shared across years. The x= labels indicated the number of fish observed without a tag.

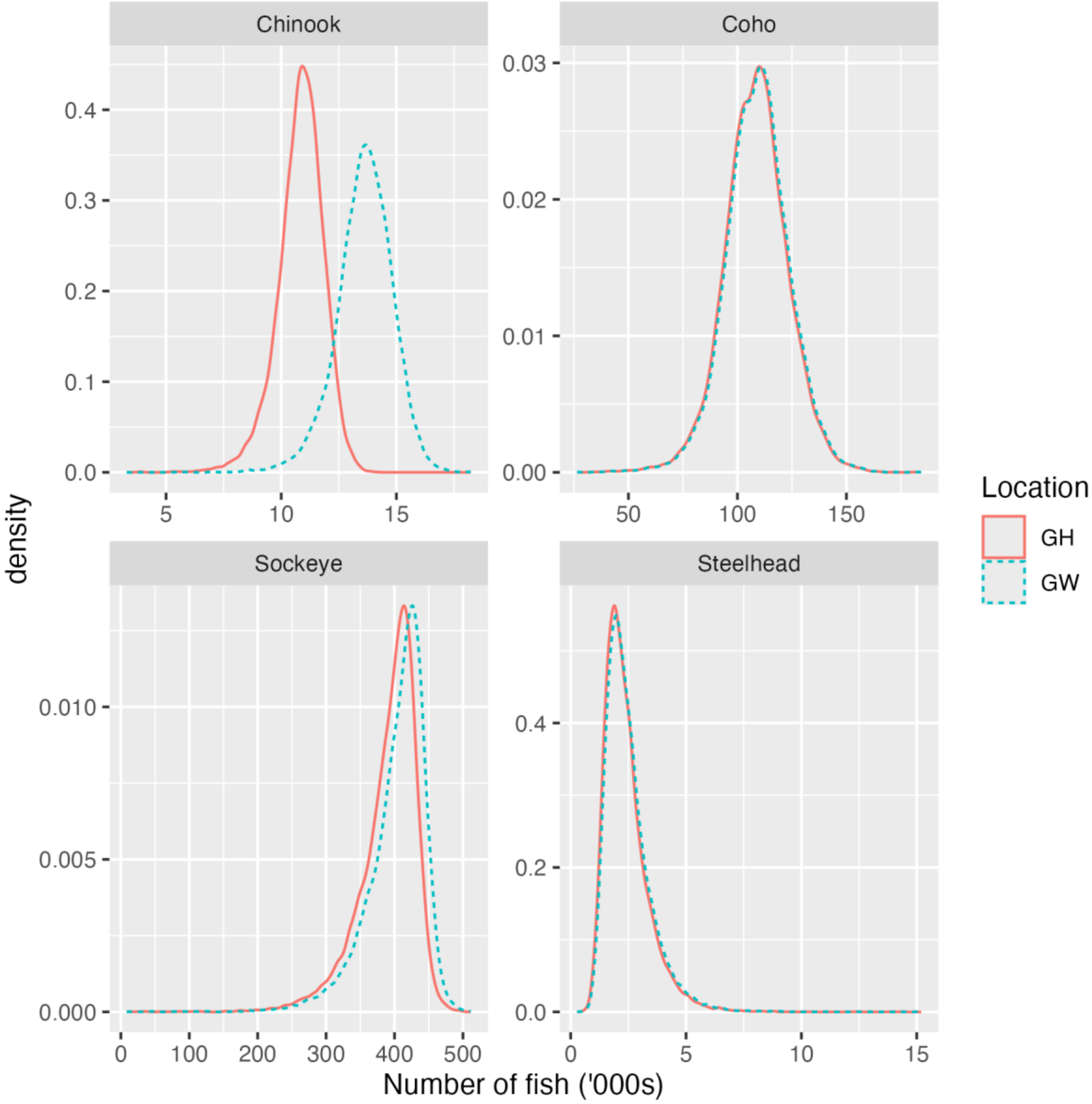


Figure 5. Posterior density for run size at Gitwinksihlkw and Grease Harbour from the unstratified model for 2021.

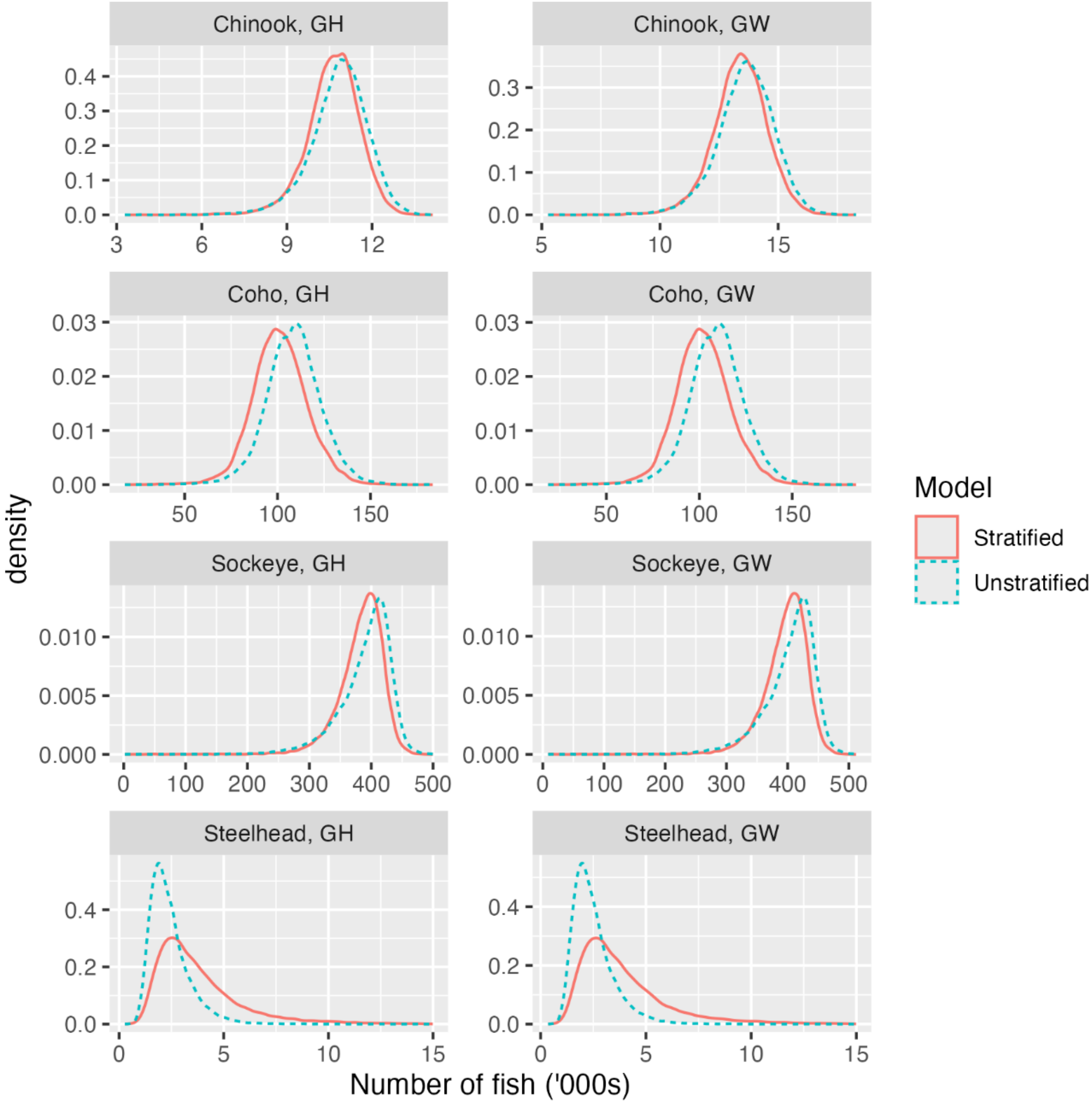


Figure 6. Comparison of the density of the posterior estimates of run size between the unstratified and stratified model for 2021.

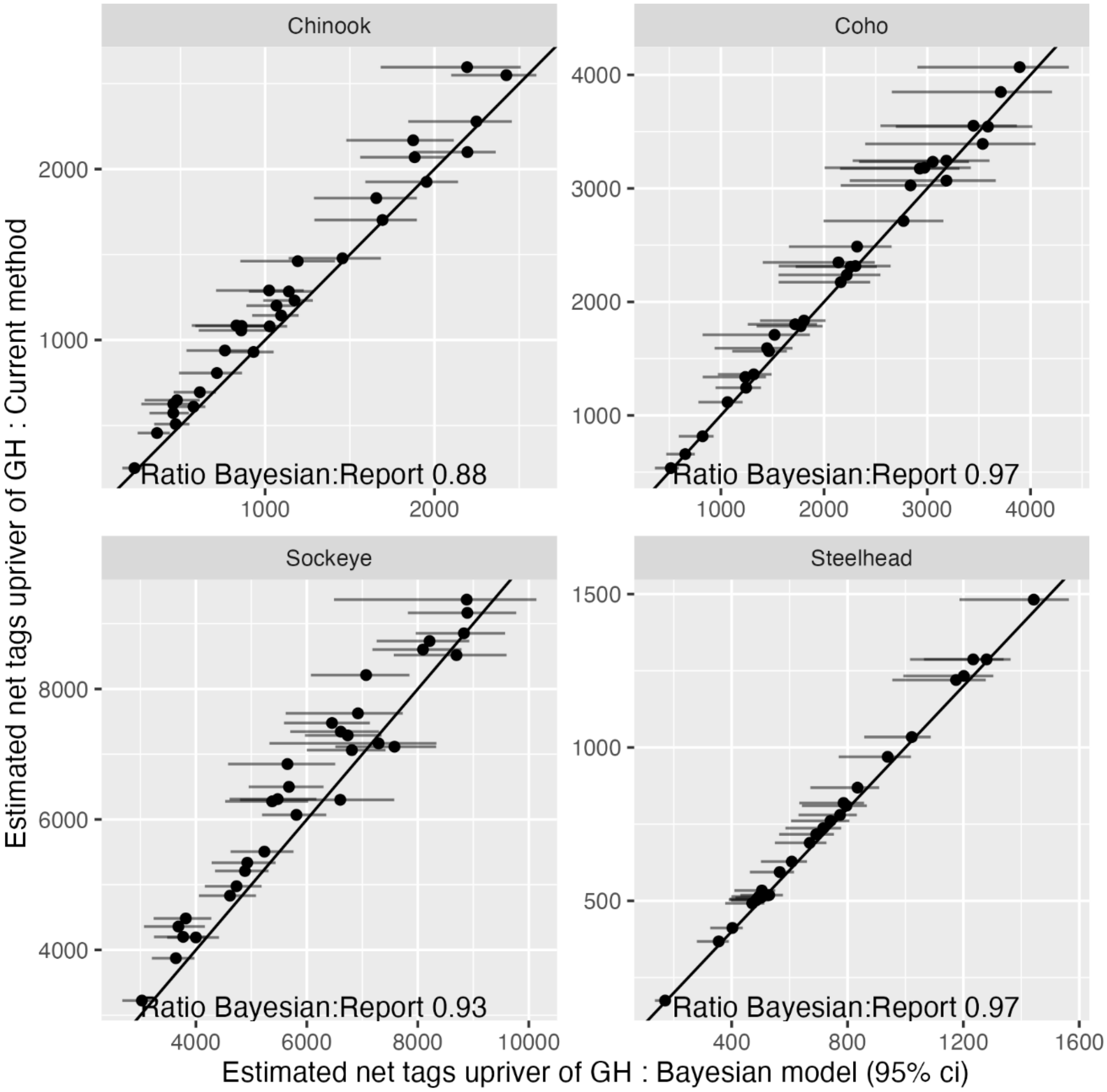


Figure 7. Comparison of the estimated net number of tags available just upriver of Grease Harbour from the Bayesian model and from the published reports. The horizontal line is the 95% credible interval or the number of tags available from the Bayesian model. Reference line is the 1:1 line.

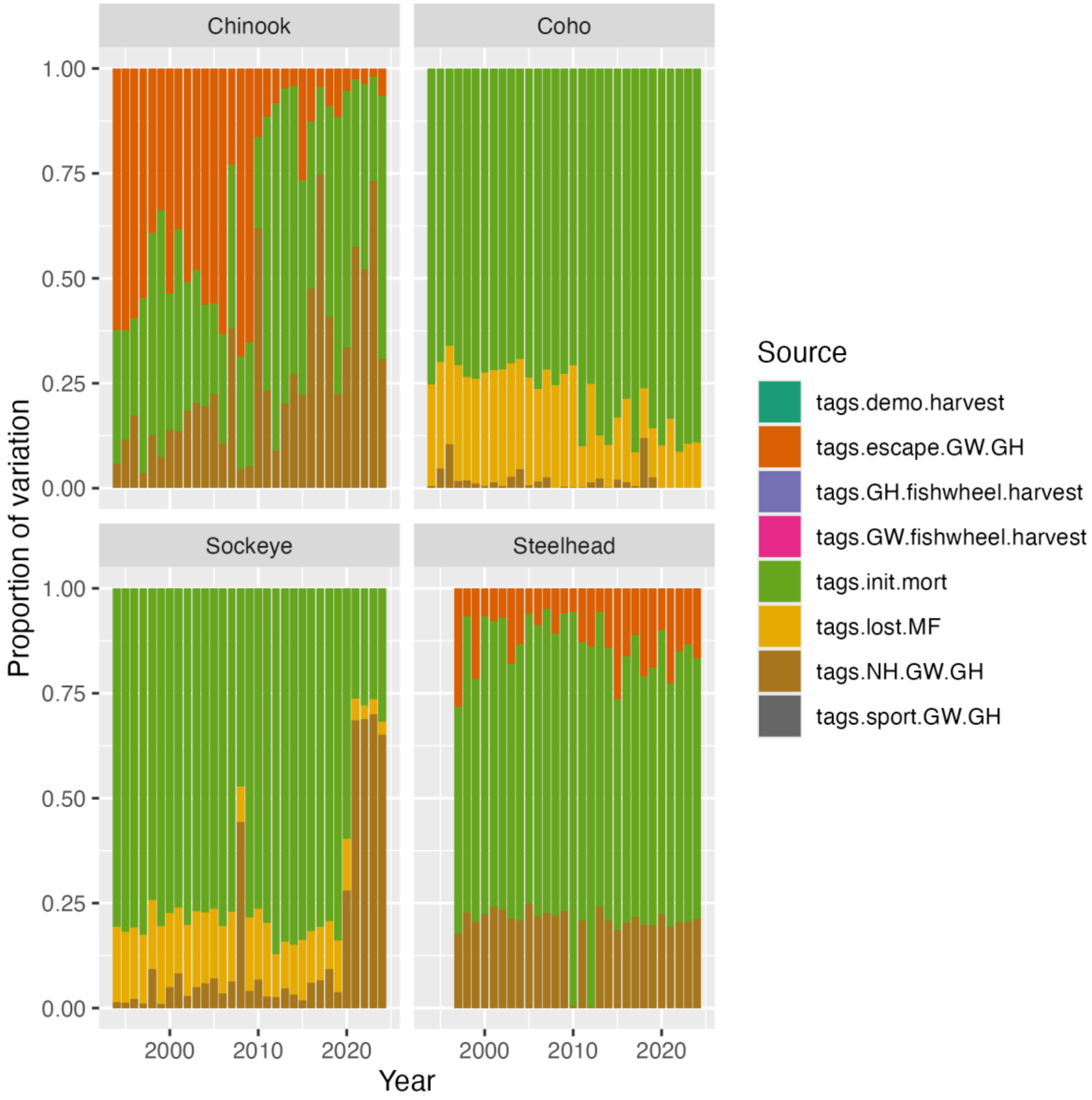


Figure 8. Decomposition of variation in the estimated net number of tags available just upriver of Grease Harbour in the Bayesian unstratified-model.

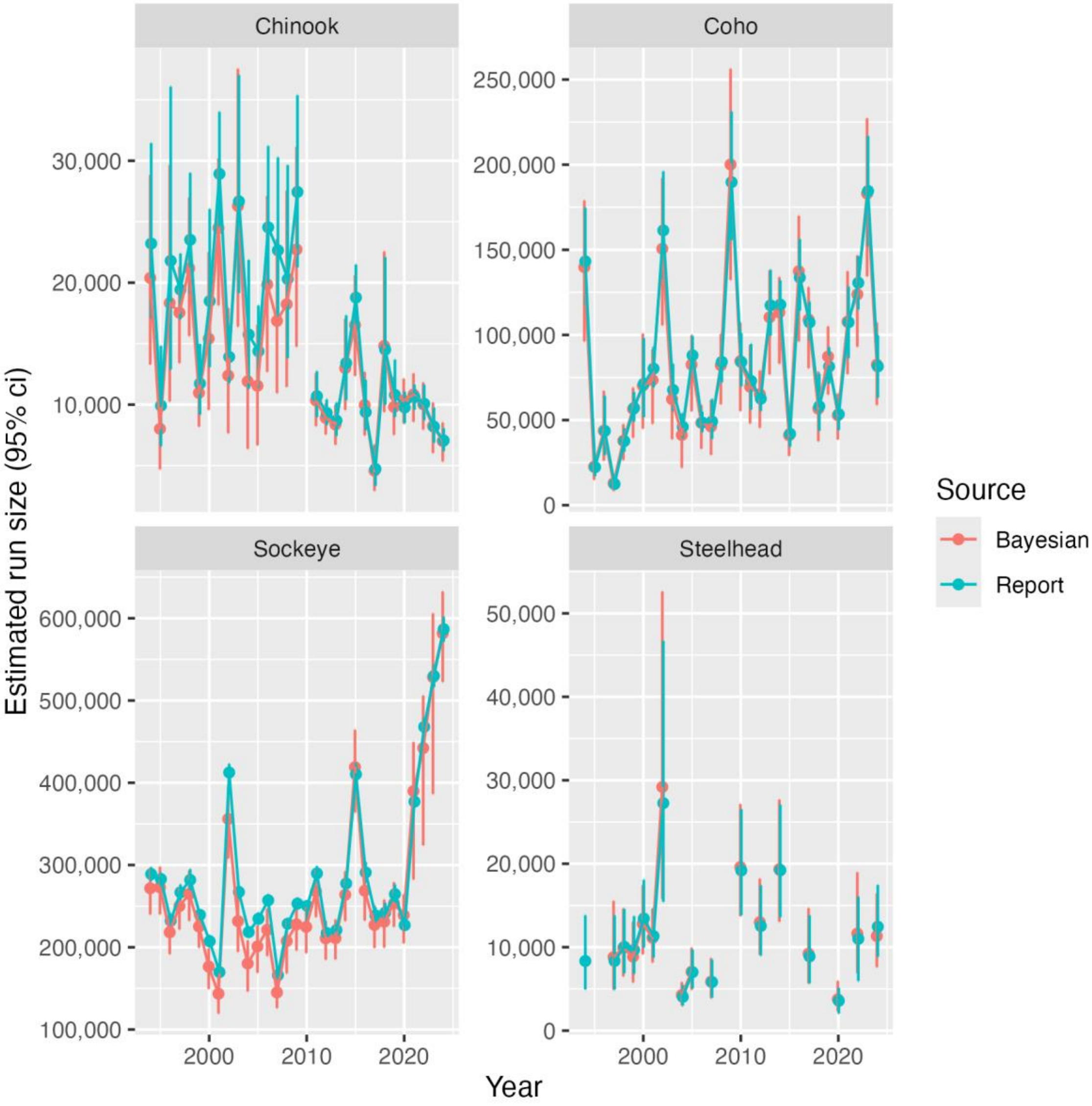


Figure 9. Comparison of estimated run sizes at Grease Harbour between the Bayesian unstratified model and values in reports. Cases where the total number of tags returned is 10 or less are not shown (mainly affects results for steelhead).

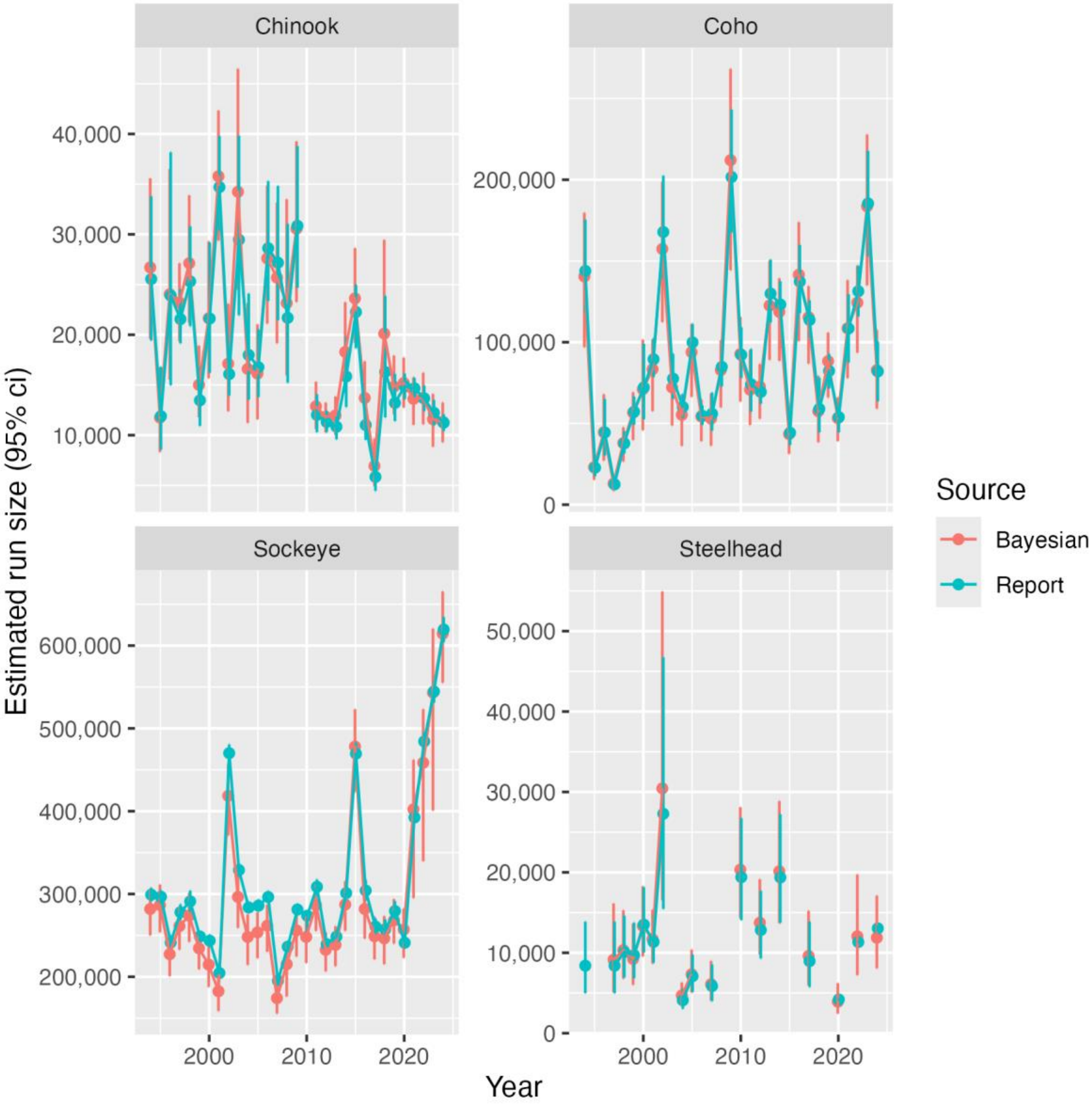


Figure 10. Comparison of estimated run sizes at Gitwinksihlkw between the Bayesian unstratified model and values in reports. Cases where the total number of tags returned is 10 or less are not shown (mainly affects results for steelhead).

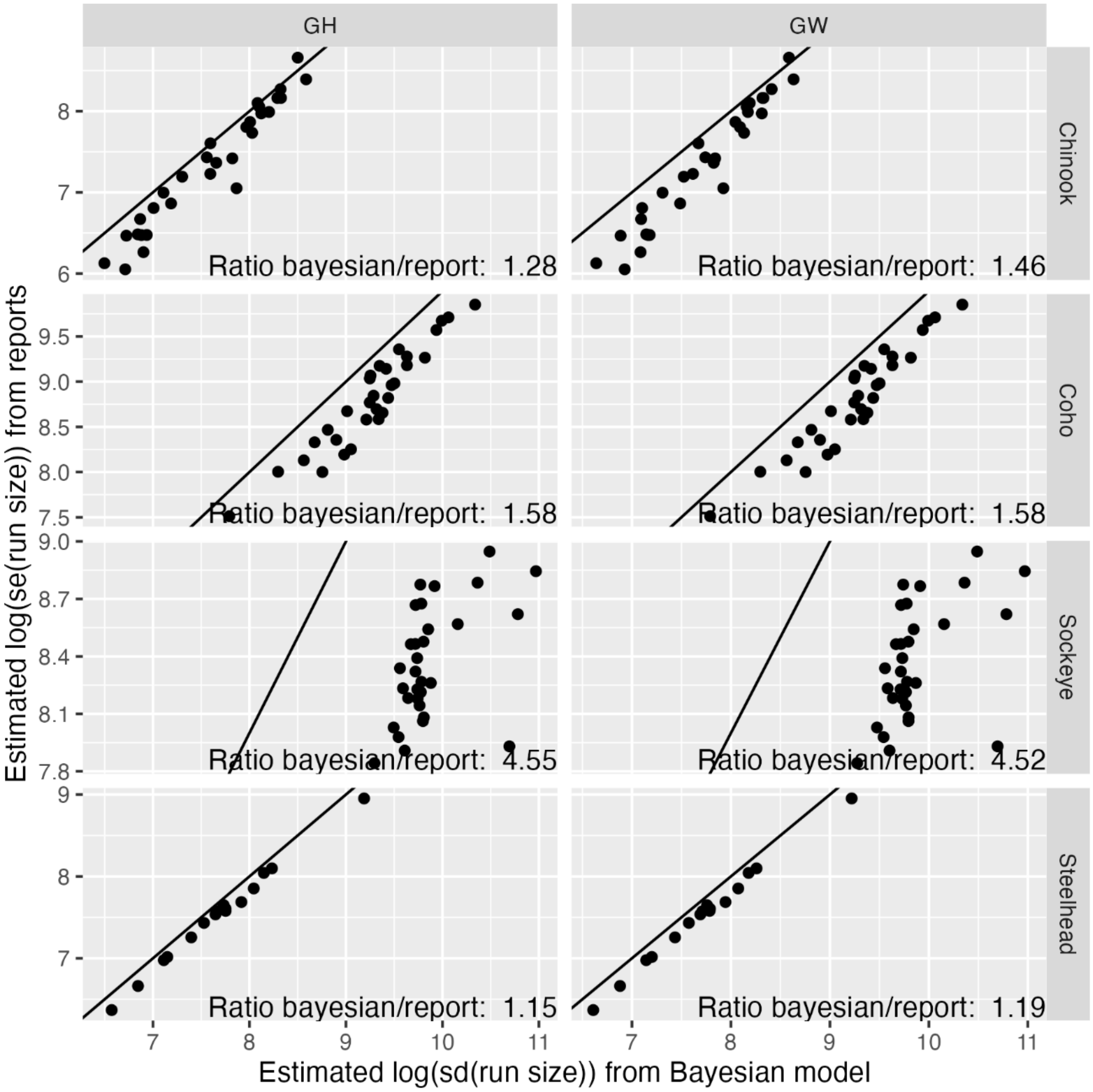


Figure 11. Comparison of the estimated uncertainty (sd from the Bayesian models; se from the reports). Note that estimates are on the log-scale because of the wide range of values. The geometric mean ratio of the standard deviations from the Bayesian models to the standard errors in the reports is shown. Reference line is the 1:1 line.

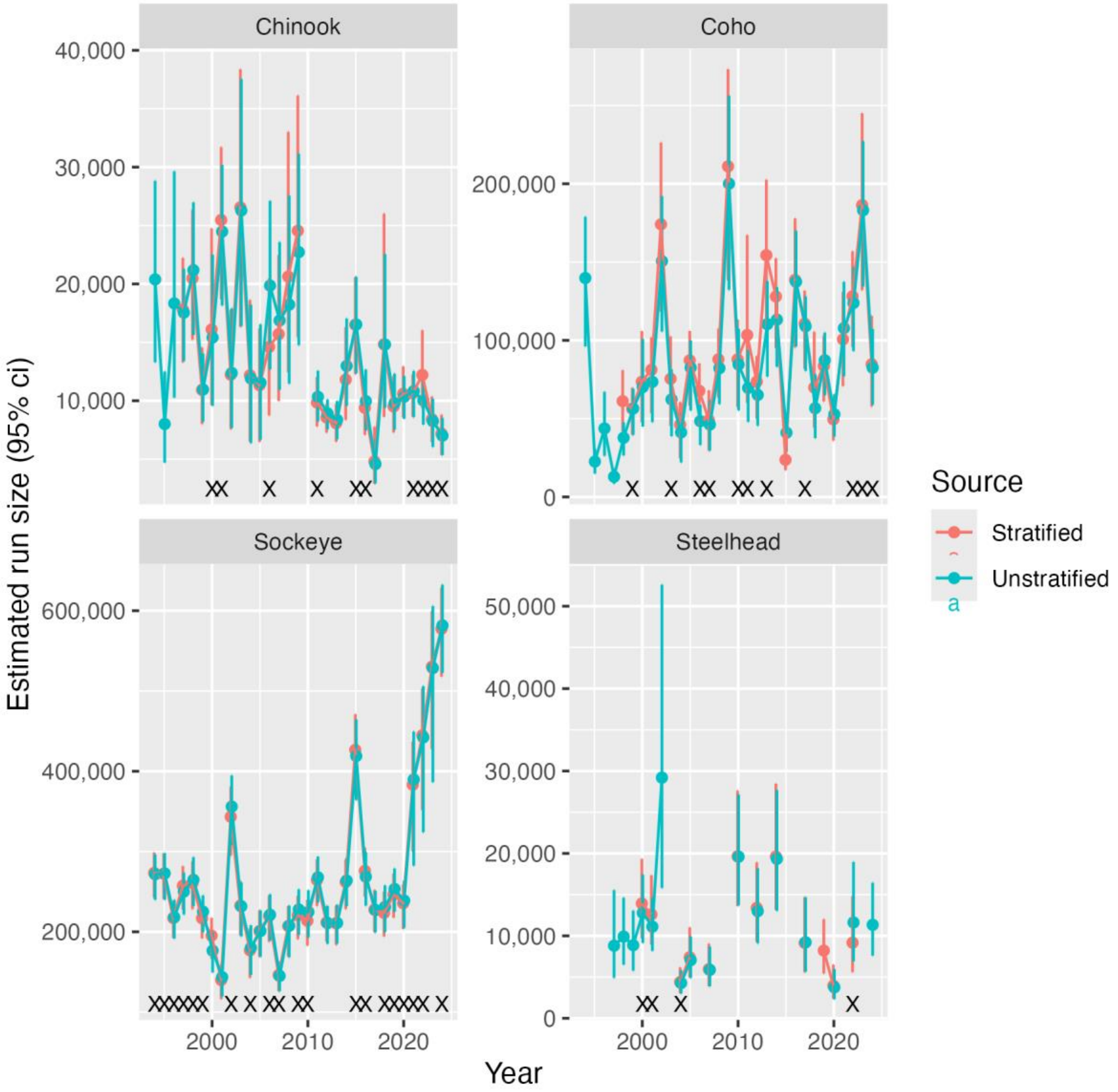


Figure 12. Comparison of the estimated run sizes at Grease Harbour between the Bayesian stratified model and the Bayesian unstratified model. Cases where the total number of tags returned is 10 or less or the number of tags in each stratum is not at least 4 are not shown (mainly affects results for steelhead). An “X” along the X-axis indicates years where there was evidence that the marked fraction differs between the two strata.

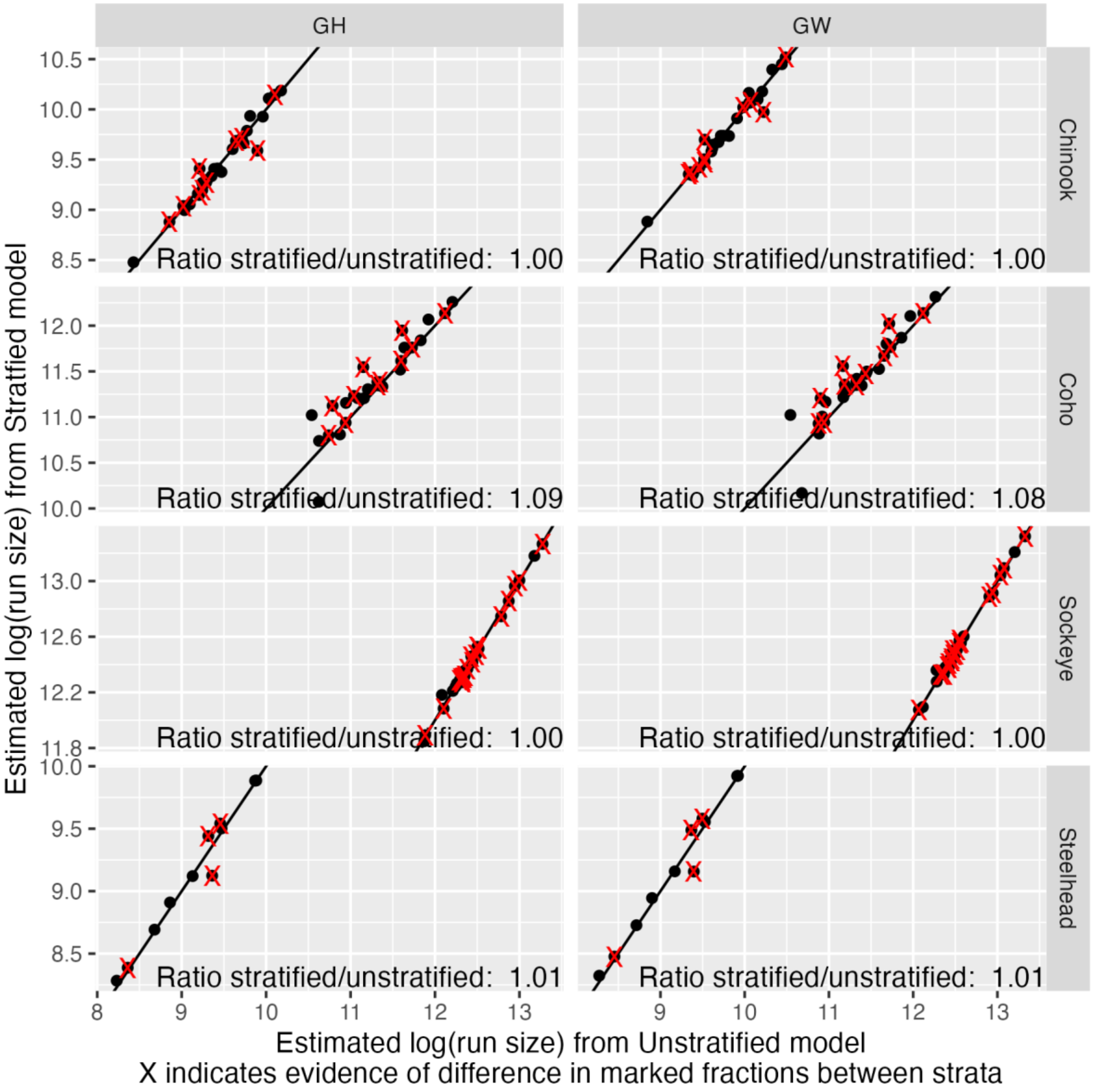


Figure 13. Comparison of the estimated run sizes at Grease Harbour between the Bayesian stratified model and the Bayesian unstratified model. Cases where the total number of tags returned is 10 or less or the number of tags in each stratum is not at least 4 are not shown (mainly affects results for steelhead). Reference line is the 1:1 line.

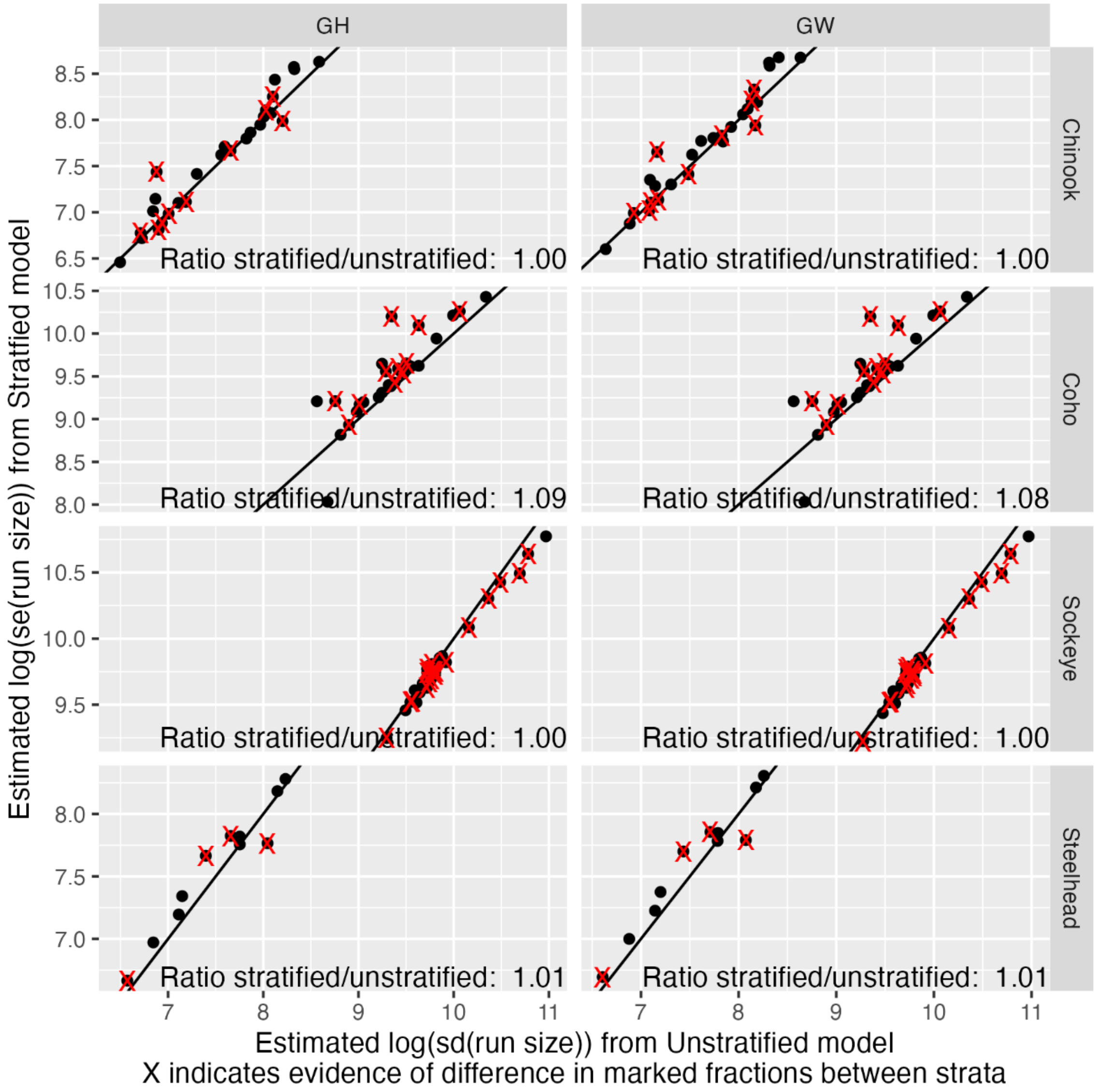


Figure 14. Comparison of the estimated uncertainty (sd from the Bayesian models) between the stratified and unstratified Bayesian models. Note that estimates are on the log-scale because of the wide range of values. The geometric mean ratio of the standard deviations from the two models is shown. Reference line is the 1:1 line.

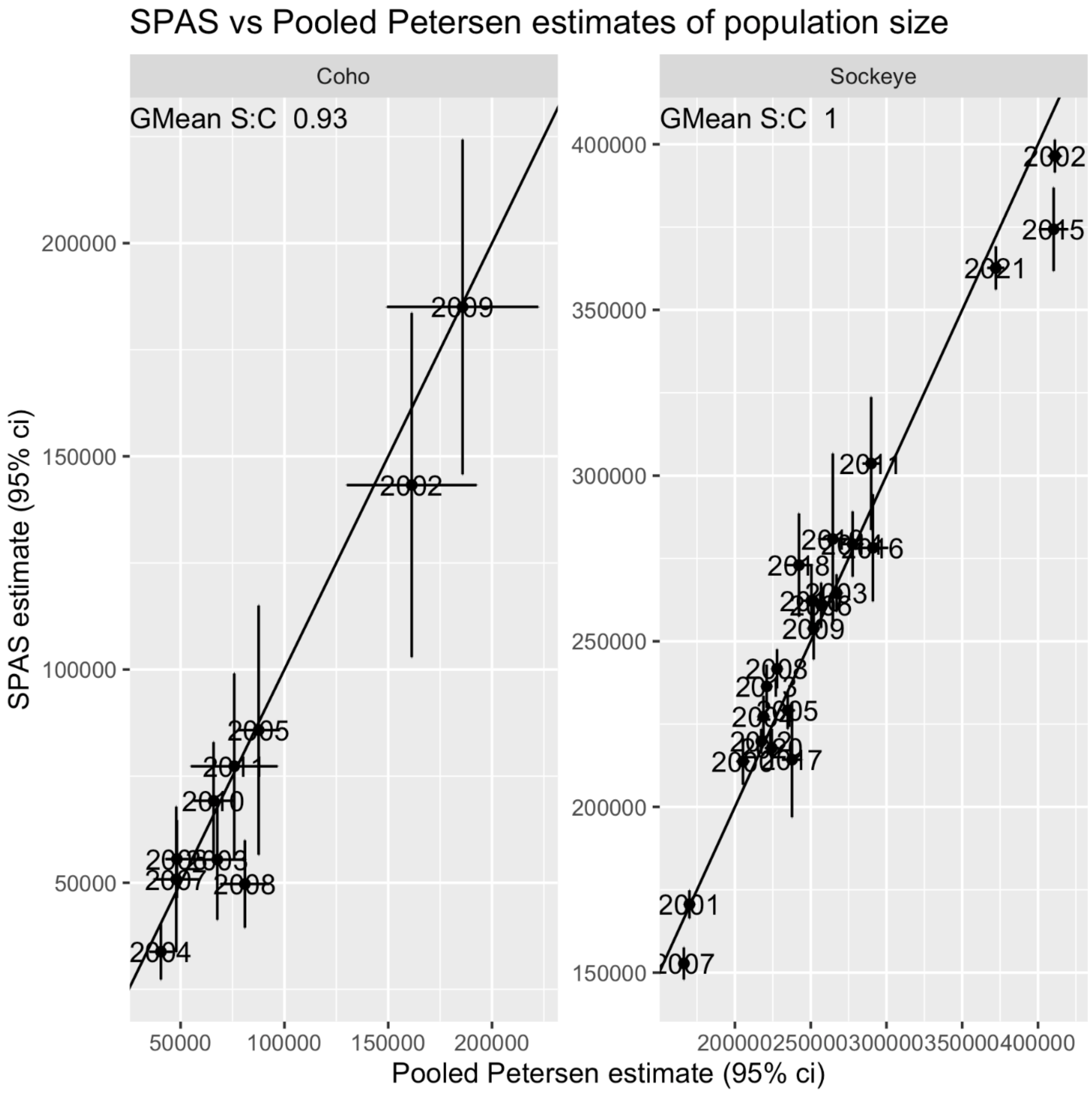


Figure 15. A comparison of temporally stratified estimator computed using SPAS vs the pooled-Petersen (Chapman) estimator for Coho and Sockeye salmon. Reference line is the 1:1 line.

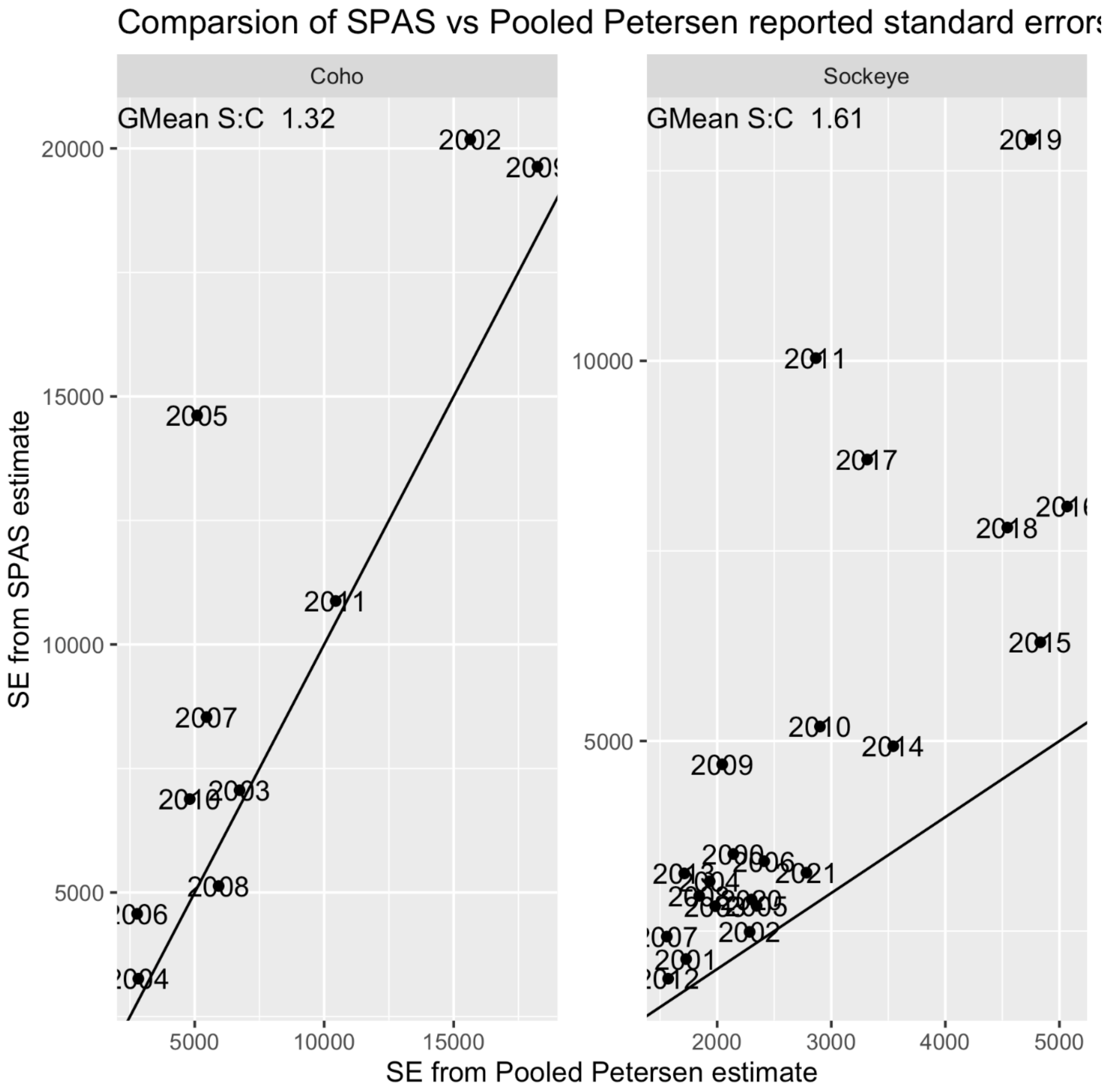


Figure 16. A comparison of standard errors from temporally stratified estimators computed using SPAS and the pooled-Petersen (Chapman) estimator for Coho and Sockeye salmon. Reference line is the 1:1 line.

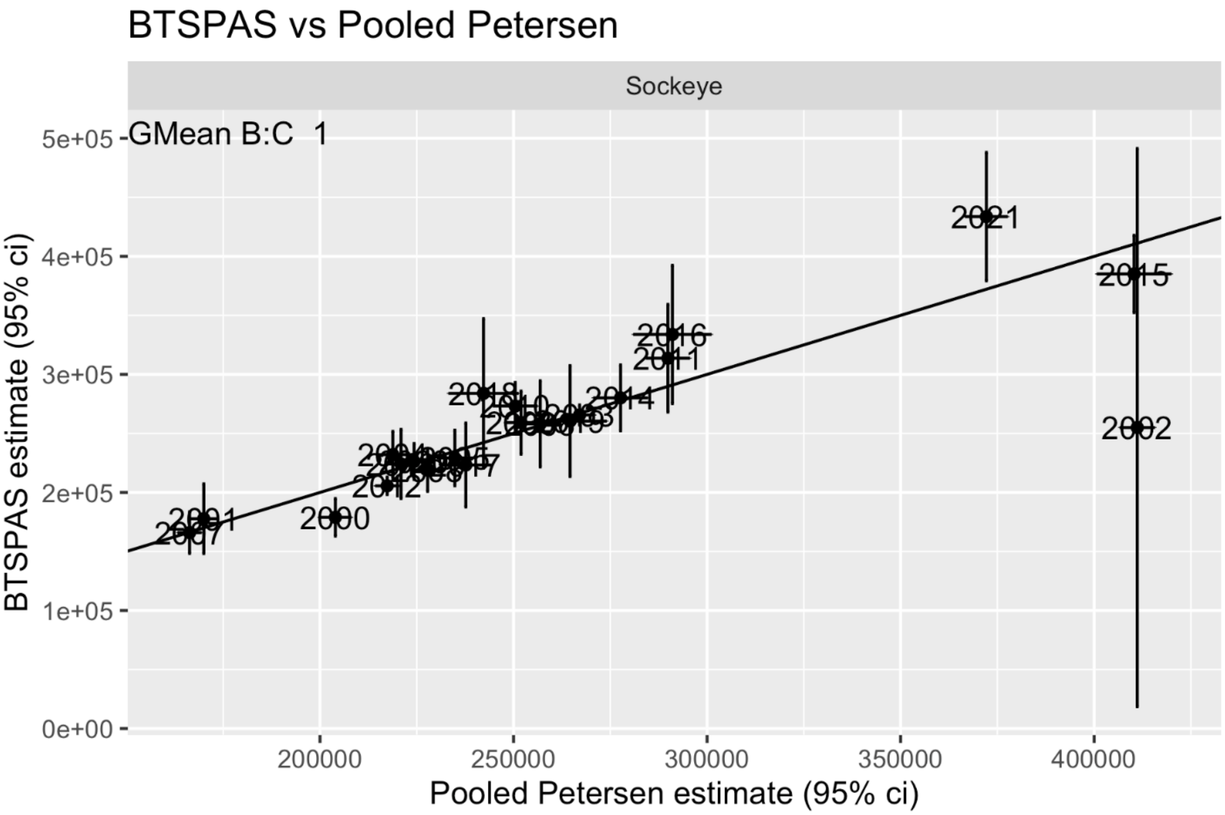


Figure 17. A comparison of temporally stratified estimators computed using BTPAS vs the pooled-Petersen (Chapman) estimator for Sockeye Salmon. Reference line is the 1:1 line.

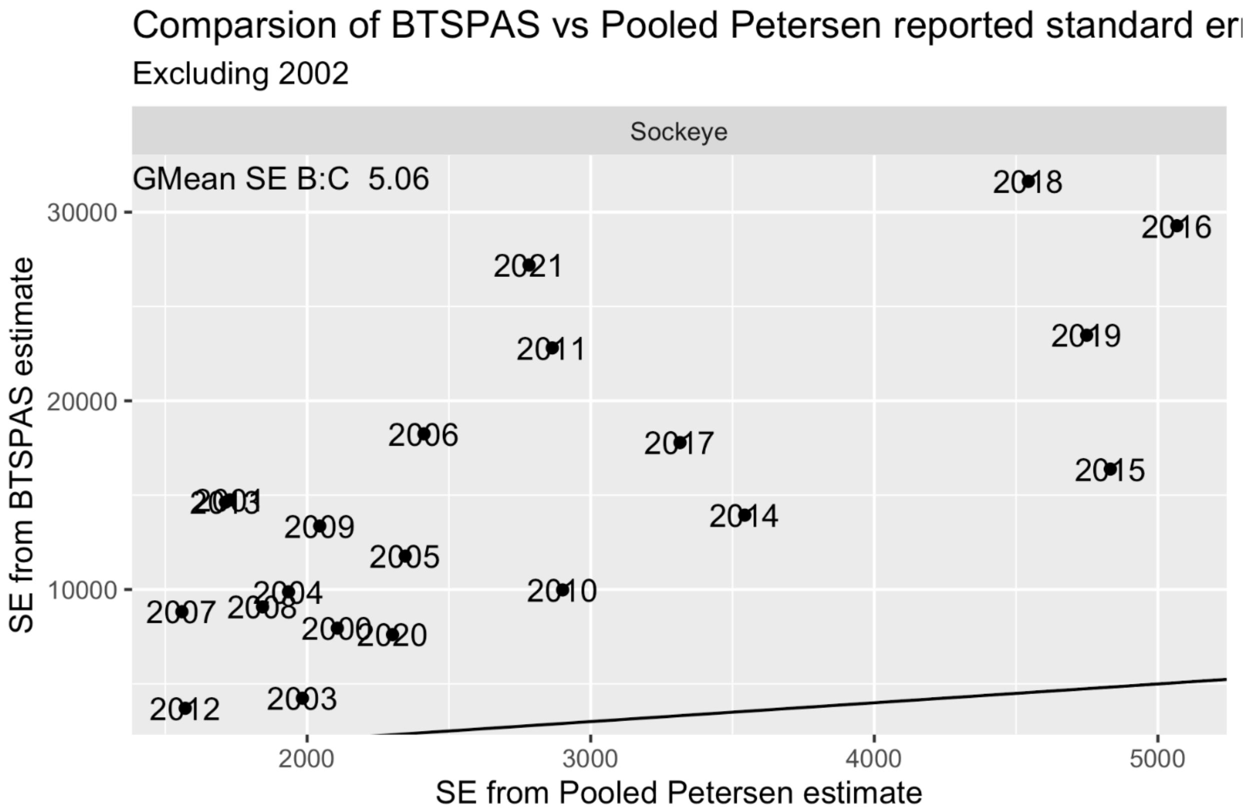


Figure 18. A comparison of measures of uncertainty from the temporally stratified estimator computed using BTPAS vs the pooled-Petersen (Chapman) estimator for Sockeye Salmon. The 2002 estimator is excluded because of peculiarities in the data (see supplemental material). Reference line is the 1:1 line.

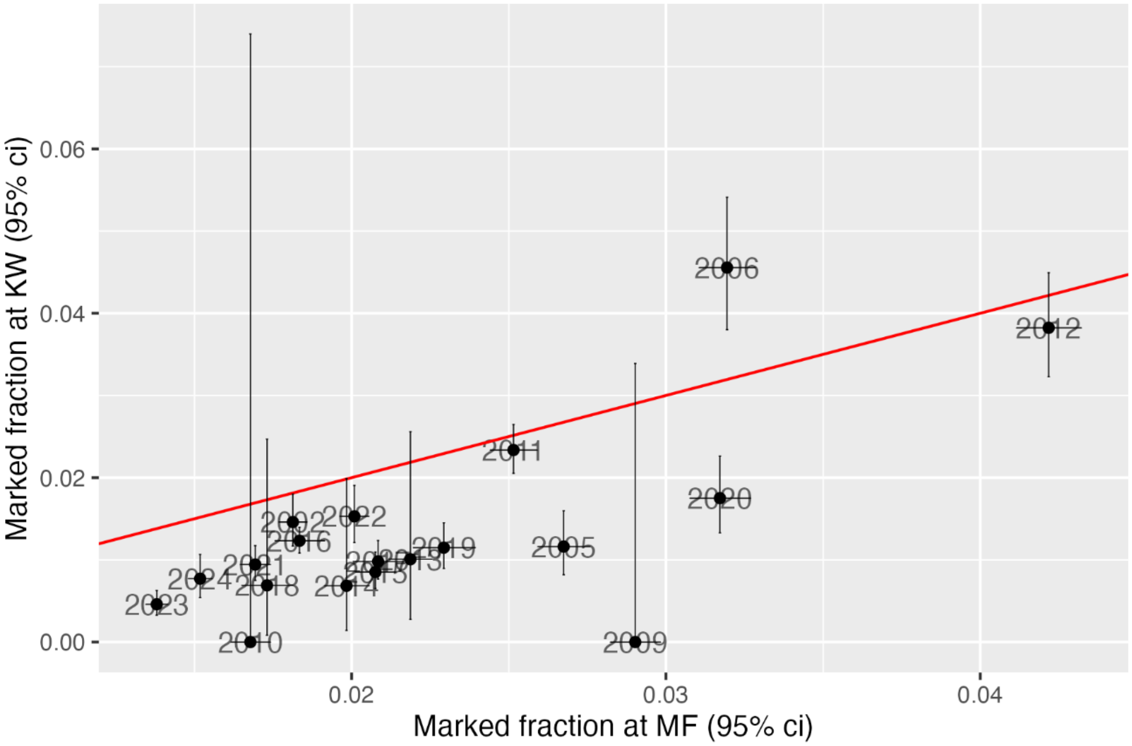


Figure 19. A comparison of the marked fractions for Sockeye Salmon at Meziadin Fishway (MF) and Kwinageese Weir (KW). Reference line is the 1:1 line.

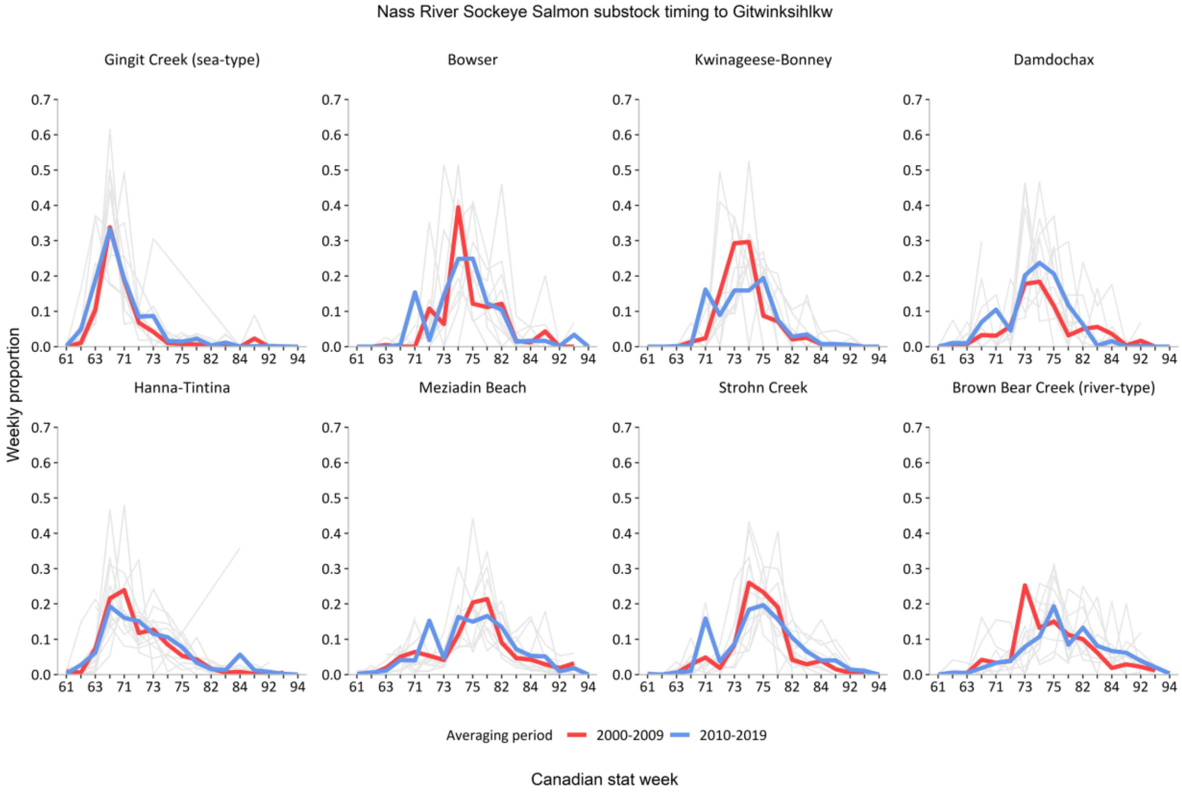


Figure 20. Nass River Sockeye Salmon sub-stock run timing to Gitwinksihlkw.

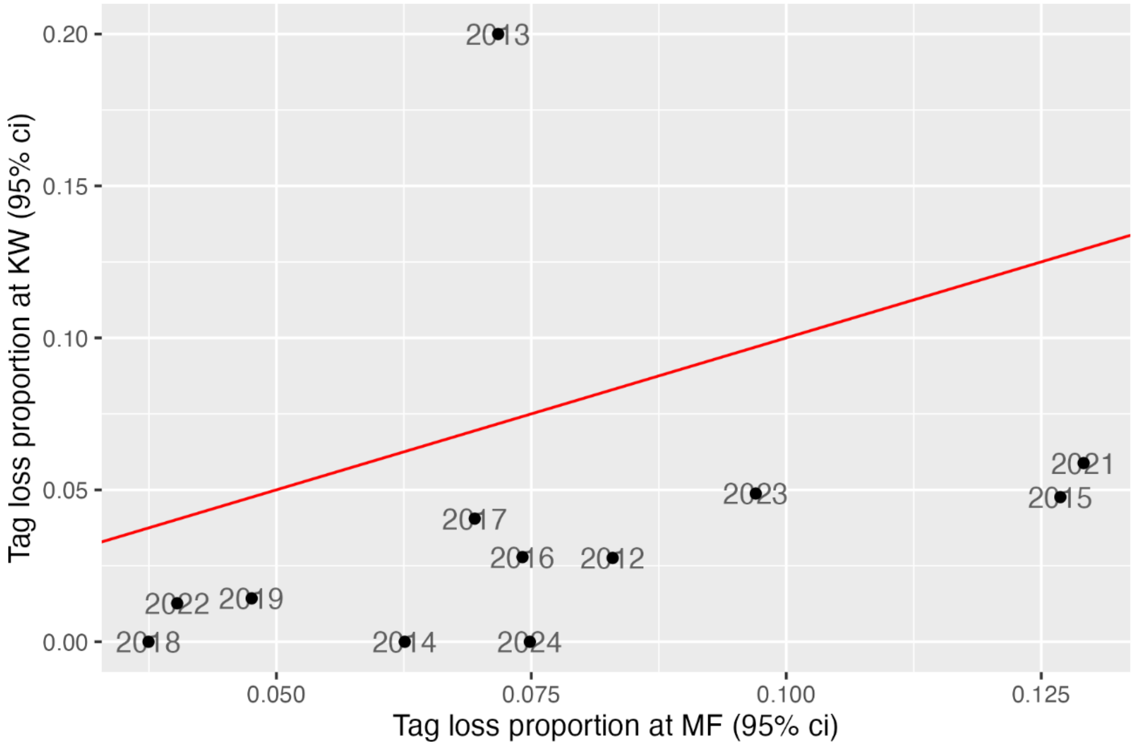


Figure 21. Comparison of tag loss estimates for Sockeye Salmon at Meziadin Fishway (MF) and Kwinageese Weir (KW). Reference line is the 1:1 line.

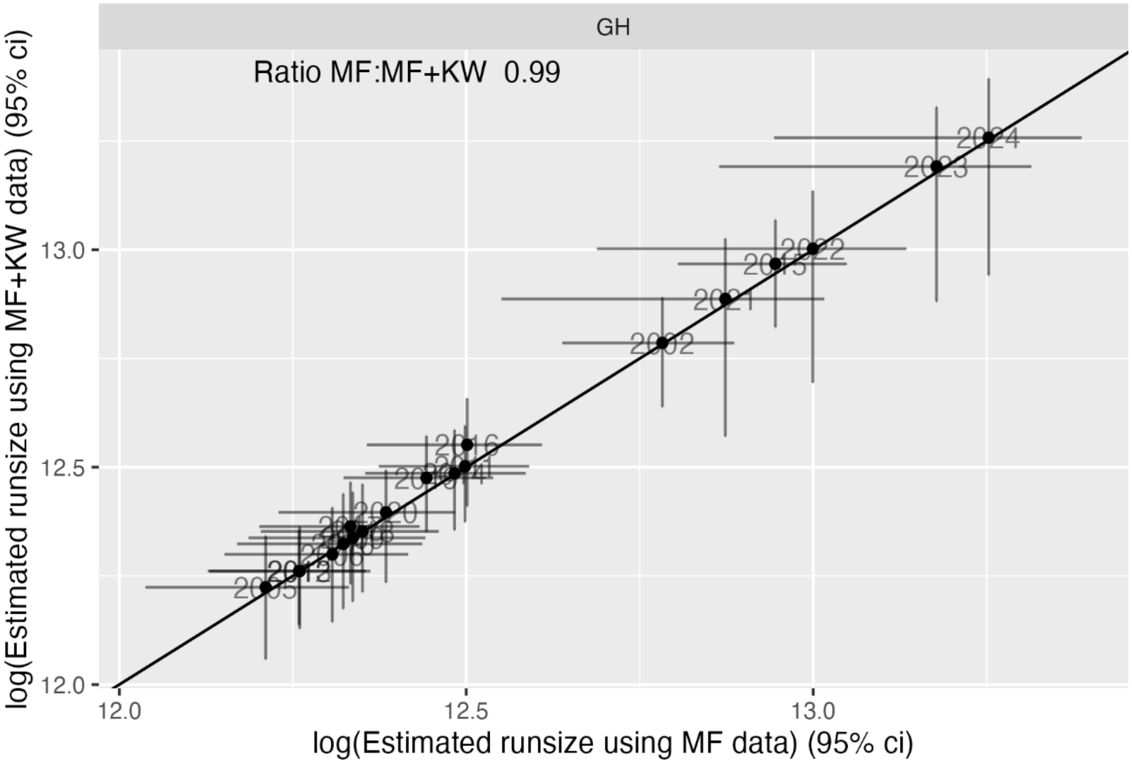


Figure 22. A comparison of the estimated run size for Sockeye Salmon just upriver of Grease Harbour when computed using Meziadin Fishway (MF) data only and when the Kwinageese Weir (KW) data were added to the MF data. Reference line is the 1:1 line.

APPENDICES

Appendix A: Mathematic description of the Bayesian model

The input data used in the Bayesian model is defined in Table A-1. The notation used in the model description below is for the stratified model. The unstratified model is similar. Data values from the unstratified model may not equal the sum of the values from the stratified model because some (a small number) of fish were not stratified and/or stratification was not possible at some recovery locations. The stratified model with a single stratum is equivalent to the unstratified model.

Underlying parameter or unobserved stochastic nodes (e.g., number of tagged fish that lose their tags)[2] are defined in Table A-2.

A.1 Part 1. Modelling the net number of tags just above GH available for recapture at upstream locations (MF, KW, DC, CR or other).

Fish are tagged at GW (and at GH for some species). The number of tagged fish is adjusted for initial tagging mortality/ fall back, tag loss (for Coho and Sockeye Salmon), spawning between GW and GH (for Chinook Salmon and steelhead), removals at fishwheels, removals in the Nisg̲a'a Domestic Harvest downstream of GW, etc.

All Chinook Salmon are fully inspected for the presence of a tag and a batch mark, so tag loss is not an issue. Steelhead are currently fin-clipped so tag loss is not an issue.

Because Chinook Salmon and steelhead are tagged at both GW and GH, it is necessary to also adjust the number of tags available just upriver of GH for fish tagged at GW that spawn between GW and GH.

Because Coho and Sockeye Salmon are only tagged at GW, no adjustment for spawning between GW and GH is needed, but now the estimated run size at GH also includes spawning escapement between GW and GH.

The subscript to indicate fish-stratum (e.g., by length) has been dropped for clarity in the equations that follow.

In Part I of the model (see page 2), the net number of tags available for recapture just upstream of GH is estimated. This modelling is done because some removals (e.g., Nisg̲a'a Domestic Harvest) are thought to differentially affect tagged and untagged fish.

---

[2] In the Bayesian paradigm, these are treated identically.

This first step is defined as:

GH.tags.net	= *GW.tags* + *GH.tags*	Tags applied
	- *tags.init.mort*	Tags removed due to initial mortality or fallback
	- *tags.spawn.GW.GH*	Tags removed from fish that spawn between GW and GH (Chinook and steelhead only)
	- *tags.tagloss*	Tags removed from fish that lose tags (Coho and Sockeye Salmon only)
	- *tags.NH.GW.GH*	Tags removed by the Nisga'a Domestic harvest between GW and GH
	- *tags.GW.fishwheel.harvest*	Tags removed in the GW fishwheel harvest
	- *tags.GH.fishwheel.harvest*	Tags removed in the GH fishwheel harvest
	- *tags.demo.harvest*	Estimated tags removed in the demonstration harvests
	- *tags.sport.GW.GH*	Estimated tags removed by the sport fishery between GW and GH.

The computation of these values is explained in the following sections.

A.1.1 Tagged fish lost due to initial handling mortality or fallback

Radio-tagging studies have been conducted in the past. In each study, the number of radio tags that were classified as initial mortality, fall back, drop out is summarized in the Part I of the review and can be summarized as follows. It will be assumed that the initial mortality probability is the same for all strata. Let

- n_{IM} represent the number of initial mortality studies conducted for a species.
- p_i^{IM} represent the initial tagging mortality / fallback for the mortality study i, $i=1,\ldots,n_{IM}$
- R_i^{IM} represent the net number of radio-tagged fish released in mortality study i after accounting for regurgitation, etc., $i=1,\ldots,n_{IM}$
- M_i^{IM} represent the number of radio-tagged fish that suffered mortality or fall back in mortality study i, $i=1,\ldots,n_{IM}$

For each mortality study, the mortalities are modelled as:

$$M_i^{IM} \sim Binomial(R_i^{IM}, p_i^{IM}) \qquad i=1,\ldots,n_{IM}$$

Because the number of fish released with radio tags is often small, a hierarchical model is adopted (see Section 2.1, on the logit scale):

$$logit(p_i^{IM}) \sim Normal(\mu = \mu_{logit(p)}^{IM}, \sigma = \sigma_{logit(p)}^{IM}) \qquad i=1,\ldots,n_{IM}$$

Priors (on the logit scale) are adopted[3]

$$\mu_{logit(p)}^{IM} \sim Normal(\mu = -2.5, \sigma = 3)$$

$$(1/\sigma_{logit(p)}^{IM})^2 \sim Gamma(.001, .001)$$

With a small number of mortality studies, estimation of the year-to-year standard deviation in the *logit*(mortality probability) will be difficult and more suitable priors can easily be substituted.

Finally, for the year in which tagging takes place, we model the probability for the number of tags lost due to initial mortality or fallback using

$$logit(p_{tagging.year}^{IM}) \sim Normal(\mu = \mu_{logit(p)}^{IM}, \sigma = \sigma_{logit(p)}^{IM})$$

The posterior distribution for $p_{tagging.year}^{IM}$ will account for the year-to-year variation in the tagging mortality and the uncertainty in the estimate based on a limited number of studies and a limited number of radio-tagged fish released each year.

The number of tagged fish lost to mortality/ fall back after tagging is then modelled as:

$$tags.init.mort \sim Binomial(GW.tags + GH.tags, p_{tagging.year}^{IM})$$

The modelled number of tagged fish lost to initial mortality/ fallback will incorporate the year-to-year variation in the mortality probability, the uncertainty due to the small number of mortality studies and small sample sizes in each year, and binomial variation in the actual number that are lost.

A.1.2 Tagged fish lost due to spawning between Gitwinksihlkw and Grease Harbour (Chinook Salmon and steelhead only)

After fish are tagged and released at GW, some fish may spawn between GW and GH. Only those fish tagged at GW are subject to “removal” due to spawning between GW and GH before arriving at GH.

For Coho and Sockeye Salmon, fish are only tagged at GW and so adjustment is done to the number of tags available from GH, but the estimated run size then INCLUDES the number of fish that spawn between GW and GH.

For Chinook Salmon and steelhead, fish are tagged at GW and GH and the estimated run size EXCLUDES the number of fish that spawn between GW and GH.

Genetic samples taken at GW provide an estimate of the proportion of the fish captured at GW that spawn between GW and GH. It is assumed that this proportion is applicable to the total run size at GW as well.

---

[3] The mean for the prior varies by species. The standard deviation of around 3 will give a shape for the prior that has a peak at smaller values, and then tapers off towards 1. See the supplemental reports for details.

Let

- $\hat{p}_i^{SP}$ represents the estimated probability of spawning between GW and GH in genetic study *i, i=1,…,n_{SP}*
- $SE(p_i^{SP})$ represents the uncertainty in this estimate. This will be a combination of binomial uncertainty (number of fish in the genetic sample that belong to the stocks that spawn between GW and GH) and genetic assignment uncertainty.

The actual yearly proportion of spawning between GW and GH is modelled on the logit scale as

$$logit(\hat{p}_i^{SP}) \sim Normal(\mu = logit(p_i^{SP-observed}), \sigma = \frac{SE(p_i^{SP-observed})}{p_i^{SP-observed}(1-p_i^{SP-observed})})$$

The complex term for the standard deviation converts the SE in the probability of stock assignment to the standard error on the *logit scale based on a Taylor-series expansion (Delta Method)*

A hierarchical model (on the *logit* scale) is adopted for the probability of spawning between GW and GH in genetic study *i.*

$$logit(p_i^{SP}) \sim Normal(\mu = \mu_i^{SP}, \sigma = \sigma_i^{SP})$$

where

- μ_i^{SP} is the estimated mean (on the logit scale) of the probability of spawning between GW and GH over all years.
- σ_i^{SP} is the estimated standard deviation (on the logit scale) of the probability of spawning between GW and GH over all years.

The prior distributions for the mean and standard deviation are

$$\mu_i^{SP} \sim Normal(\mu = -1.5\ (Sockeye)\ or\ -4\ (steelhead), \sigma = 1)$$

$$1/(\sigma_i^{SP})^2 \sim Gamma(.001, .001)$$

Finally, there are two cases to consider when generatingg the number of fish with tags that spawn between GW and GH. First, suppose that in a particular year, the mark-recapture study includes a genetic study. Then

$$tags.spawn.GW.GH \sim Binomial(GW.tags, p_{tagging.year}^{SP})$$

i.e., the genetic data from the study year corresponding to the tagging year after being “shrunk” by the hierarchical model is used.

Second, if there is no genetic study in the mark-recapture year, then we sample the probability from the hierarchical model,

$$logit(p_{tagging.year}^{SP}) \sim Normal(\mu = \mu_i^{SP}, \sigma = \sigma_i^{SP})$$

and then

$$tags.spawn.GW.GH \sim Binomial\left(GW.tags, p_{tagging.year}^{SP}\right).$$

A.1.3 Tagged fish that lose tags (Coho and Sockeye salmon only)

Tagged fish that migrate upstream can suffer tag loss which can lead to bias in the capture-capture estimates that assume no tag loss. As shown in Part I of the review, there are two mathematically equivalent ways to deal with tag loss. We have chosen to adjust the number of fish tagged at GW and GH for tag loss.

Tag loss is not an issue for Chinook Salmon because all fish are carefully inspected for the presence of a tag and the batch secondary mark. Tag loss is not an issue for steelhead because all fish are fin-clipped which is assumed not to close.

As fish pass the MF, they are classified as having an observable tag or not. In addition, a sub-sample of fish with no observable tags is taken in several years and fully inspected for the presence of the batch secondary mark.

Let

- N_{TL} represent the number of tag loss studies conducted for a species.
- p_i^{TL}represent the probability of tag loss in study *i, i=1,…,n_{TL}*
- NOT_i^{Total}represent the total number of fish inspected at MF that have no observable tag in study *i, i=1,…,n_{TL}*
- OT_i^{Total}represent the total number of fish inspected at MF that have an observable tag in study *i, i=1,…,n_{TL}*
- I_i^{TL}represent the number of fish that undergo a secondary inspection after failing to have an observed tag in study *i, i=1,…,n_{TL}*
- L_i^{TL}represent the number of fish in the secondary inspection that have lost their tag in study *i, i=1,…,n_{TL}*
- p_i^{STL}represent the conditional probability of actual tag loss in the secondary sample rather than a fish that was never tagged given the number of fish inspected in the primary inspection in study and the number of tags observed in study *i, i=1,…,n_{TL}* Its form is complex, but an approximation is:

$$p_i^{STL} = \frac{OT_i^{Total}}{NOT_i^{Total}} \times \frac{p_i^{TL}}{(1 - p_i^{TL})}$$

The distribution of L_i^{TL} is very complex, but a binomial distribution serves as an approximate model:

$$L_i^{TL} \sim Binomial(I_i^{TL}, p_i^{STL})$$

A hierarchical model (on the logit scale) is adopted for the probability of tag loss in tag loss study year *i.*

$$logit(p_i^{TL}) \sim Normal(\mu = \mu_i^{TL}, \sigma = \sigma_i^{TL})$$

where

p_i^{TL} is the probability of tag loss in tag loss study year *i,*

$\mu_i^{TL}$is the estimated mean (on the logit scale) of the probability of tag loss over all years.

σ_i^{TL} is the estimated standard deviation (on the logit scale) of the probability of tag loss over all years.

The prior distributions for the mean and standard deviation are

$$\sim Normal(\mu = -3, \sigma = 3)$$

$$1/(\sigma_i^{TL})^2 \sim Gamma(.001, .001)$$

Finally, there are two cases to consider when generating the number of fish with tag loss from those tagged at GW and GH. First, suppose that in a particular year, the mark-recapture study includes a tag loss study this year. Then

$$tags.loss \sim Binomial(GW.tags\ +\ GH.tags, p_{tagging.year}^{TL})$$

i.e., the estimated tag loss data from the study year corresponding to the tagging year is used.

Second, if there is no tag loss study in the mark-recapture year, then we sample a tag loss probability from the hierarchical model,

$$logit(\ p_{tagging.year}^{TL}) \sim Normal(\mu = \mu_i^{TL}, \sigma = \sigma_i^{TL})$$

and then

$$tags.loss \sim Binomial\big(GW.tags\ +\ GH.tags, p_{tagging.year}^{TL}\big).$$

A.1.4 Tagged fish removed in the Nisg̲a’a Domestic Harvest between Gitwinksihlkw and Grease Harbour

The Nisg̲a'a Domestic Harvest consists of two components:

1. the Domestic Fishery whose catch is estimated using survey methods. Not all tags that are removed in the Domestic Fishery are reported. The estimates and precision of the Domestic Fishery Harvest catch is found using a survey methods and estimates are produced yearly; and
2. the Individual-Sale Fishery. Part of the total IS fishery is closely monitored (typically 90%+) and all tags in the monitored fishery are counted. The IS fishery catch is included in the total catch estimate from the Domestic Fishery.

The number of fish with tags removed in the Nisg̲a'a Domestic Harvest (DF and ISF combined) is estimated based on the tag ratio in the fully observed Individual-Sales Fishery.

Compliance information has been recorded since 2000 in three river-strata around and downstream of GH. Notice that in some years, no Individual Sales occurred and that in many years the number of tags returned from the *Lower* river-stratum is very small in the Individual Sales fishery. Similarly, in recent years, monitoring has been reduced in response to Covid-19.

The total catch is also not known with certainty. The standard error of the estimates of total catch was extracted from the Catch Estimation reports.

A simple model is applied for each year using the data collected each year. First, the number of tags observed in the IS fishery in each river-stratum *h* is modelled as.

$$IS_h^{TagsObserved} \sim Binomial(IS_h^{Catch}, p_h^{IS.tag.ratio})$$

The prior distribution for the proportion of the catch that is tagged in the individual sale fishery depends on the river-stratum:

$$p_h^{IS.tag.ratio} \sim Beta(\alpha_h^{IS.tag.ratio}, \beta_h^{IS.tag.ratio})$$

Even if the observed number of tags in the Individual-Sale Fishery is zero, the posterior distribution for $p_h^{IS.tag.ratio}$ can be computed. Cases where there is no Individual Sale Fishery (or none is monitored) is problematic because there is no data available. I suggest in these cases that "pseudo"-values be created to represent an "average" in past years.

The actual domestic harvest is modelled around the estimated domestic harvest to account for uncertainty in the harvest

$$TotalCatch_h \sim Normal(\mu = EstTotalCatch_h\,, \sigma = SE(EstTotalCatch_h\,))$$

The estimated number of fish with tags removed in the Domestic Harvest in river-stratum *h* is then modelled as:

$$tags.NW.GW.GH_h \sim Binomial(TotalCatch_h, p_h^{IS.tag.ratio}),$$

We will ignore any cases where the modelled number of fish with tags removed is less than the observed tags reported in the Domestic fishery for simplicity.

Note that the total catch used to expand the tag-fraction in the ISF ($TotalCatch_h$) may differ from the total catch reported in harvest monitoring due to timing issues. For example, harvest monitoring may measure harvest starting in May, but fishwheels may not start operating until June and so only the harvest from June onwards (when the fishwheels are operating and tagging fish) is used. In most cases, the difference in estimates of catch is small except for Sockeye Salmon in 2000-2006. **In the current Bayesian implementation, the value of $\boldsymbol{TotalCatch_h}$ is taken from the harvest reports and not adjusted for timing issues.**

Finally, the total number of fish with tags removed in the Domestic Harvest is the sum over the three river-strata.

$$tags.NH.GW.GH = \sum_h tags.NH.GW.GH_h$$

In the unstratified model, no further computations are needed. In the stratified model, the allocation of the total number of fish with tags removed in the Nisga'a harvest to the individual strata is based on a sub-sample of fish from the harvest that are measured and classified into the individual strata. We assume a multinomial model for this subsample and then apply the multinomial model to the total catch (assuming that a maximum of two fish strata are used):

$$\left(tag.NH.GW.GH_1^{harvest\ sub\ sample}, tags.NH.GW.GH_2^{harvest\ sub\ sample}\right) \sim$$
$$Multinomial\left(tag.NH.GW.GH;\ p_1^{harvest\ sub\ sample}, p_2^{harvest\ sub\ sample}\right)$$

A weak Dirichlet prior is placed on the multinomial proportions based on multiple years of sub-samples and differs among the species (see supplemental documents for details).

A.1 5 Tags lost to the Grease Harbour fishwheel harvest

This is an exact observed value and used directly as is without any uncertainty.

A.1.6 Tags lost in the demonstration harvests

These are estimated values, and the estimate and standard error are used to generate values using a (discretized) Normal distribution i.e., rounded to the nearest integer:

$$Tags.demo.harvest \sim \text{Discretized Normal}(\mu = tags.demo.harvest.est, \sigma = tags.demo.harvest.se)$$

If the number of tags lost to the demonstration harvest is known exactly, then the SE can be set to very small value (e.g., 0.001).

A.1.7 Tags lost in the sport harvest between Gitwinksihlkw and Grease Harbour

These are estimated values, and the estimate and standard error are used to generate values using a (discretized) Normal distribution:

$$Tags.sport.GW.GH \sim \text{DiscretizedNormal}(\mu = tags.sport.GW.GH.est, \sigma = tags.sport.GW.GH.se)$$

If the number of tags lost to the sport fishery between GW and GH is known exactly, then the SE can be set to a very small value (e.g., .001).

A.2 Part II. Modelling the marked fraction just upriver of Grease Harbour

The marked fraction (i.e., proportion of the run just upriver of GH that has a tag) is assumed to be constant over the rest of the journey upstream (i.e. all tag loss has occurred, all initial mortality has occurred etc.). Any subsequent removals upstream of GH that operates on tagged and untagged fish equally will not affect the marked fraction and so can be “ignored”.

Recoveries could occur at a number of locations (e.g., MF, KW, DC, CR and others). At each location the number of fish inspected for tags and the number of tags observed is assumed to be recorded exactly without error and modelled using Bailey (1951, 1952) Binomial model.

$$MF.tags \sim Binomial(MF.inspect, GH.marked.fraction)$$

$$KW.tags \sim Binomial(KW.inspect, GH.marked.fraction)$$

$$DC.tags \sim Binomial(DC.inspect, GH.marked.fraction)$$

$$CR.tags \sim Binomial(CR.inspect, GH.marked.fraction)$$

Not all species have data from all locations in all years. For years where data are not taken or values are not available, values of 0 can be used for the number of fish inspected and the number of tags observed and will not contribute any information.

Prior information about the marked fraction in the first fish-stratum (entire population when an unstratified analysis is done) is modelled using a beta distribution

$$GH.marked.fraction_1 \sim Beta(\alpha_{GH.marked.fraction}, \beta_{GH.marked.fraction})$$

Choices for α and βwill indicate the strength of the prior information. For example, the mean of this prior distribution is $\frac{\alpha}{\alpha+\beta}$, and a choice of (1,29) for α and β will induce a prior distribution for the marked fraction with a mean of 1/30=3.3%.

In the stratified analysis, the marked fraction for the second fish-stratum is allowed to differ from that on the first fish-stratum after transforming to the *logit* scale:

$$logit(GH.marked.fraction_2) = logit(GH.marked.fraction_1) + delta$$

The *delta* value measures the odds-ratio of the marked fraction between the two fish-strata.

The *delta* parameter has prior distribution of

$$\text{delta} \sim \text{Normal}(\mu = \mu_{delta}, \sigma = \sigma_{delta})$$

Current choices for the prior on delta are a mean of 0 and a standard deviation of 1.

This is equivalent to modelling the odds ratio for the marked fraction between the two fish-strata with a prior distribution centered at 1 (or 0 on the logit scale) (corresponding to no difference in the marked fraction). If the data in each stratum is strong, the two estimated marked fractions will be different, but in cases for which the data are very sparse, the two marked fractions will then be very similar. This is a form of “automated” switching between a fully stratified and fully pooled Petersen estimator. An example of this automated stratification is available in the supplemental information.

Prior information on the marked fraction induces a prior on the run size just upriver of GH through the relationship

$$marked.fraction = \frac{GH.tags.net}{GH.run.size}$$

A.3 Part III. Modelling the run size at Grease Harbour

The information on the marked fraction and the net tags available just upriver of GH leads to the estimate of the run size at GH as:

$$GH.runsize \ = \frac{GH.tags.net}{marked.fraction} + \text{GH. fishwheel. harvest}$$

The posterior distributions for the net number of tags available just upriver of GH and for the marked fraction just upriver of GH induce the posterior distribution for the run size just upriver of GH.

Notice that for Coho and Sockeye Salmon, this estimated run size just upriver of GH already includes spawning escapement between GW and GH because tags were only applied at GW and no adjustment for escapement between GW and GH is made. [This could be done in a similar fashion as for Chinook Salmon and steelhead.] The run estimate at GH post fishwheel harvest is available for each fish-stratum, but the run size at GH prior to the fishwheel harvest is not stratified because the fishwheel harvest values are not available for the individual fish=strata.

A.4 Part IV. Modelling the run size at Gitwinksihlkw

The total run size at GH (i.e., summed over all fish-strata), is then adjusted for sport harvest, Nisga'a harvest, and escapement that occurs between GW and GH. All of these adjustments are only done at the pooled level, i.e., the estimates of run size at GW are not stratified.

For Sockeye and Coho salmon, tagging only occurred at GW. Consequently, the run size estimate at GH already includes escapement between GW and GH and no further adjustments are needed. The run size at GW is found by adding the estimated sport and Nisga'a harvest between GW and GH, and the fishwheel harvest at GW. The sport and Nisga'a harvest are modelled based on the estimate and standard error for these quantities, i.e.,

$$SpH.GW.GH \sim Normal(\mu = SpH.GW.GH.mean, \sigma = SpH.GW.GH.sd)$$

$$NiH.GW.GH \sim Normal(\mu = NiH.GW.GH.mean, \sigma = NiH.GW.GH.sd)$$

where the mean and standard deviation are available from harvest surveys. Then

$$GW.runsize = GH.runsize + SpH.GW.GH + NiH.GW.GH + GW.fishwheel.harvest$$

For Chinook Salmon and steelhead, the run size at GH plus the Nisga'a and sport harvest between GW and GH must be inflated by (1-p(escapement between GW and GH). We then have

$$\begin{aligned} GW.runsize = {} & (GH.runsize + SpH.GW.GH + NiH.GW.GH)(1 - p(escapement)) \\ & + GW.fishwheel.harvest \end{aligned}$$

Table A-1. Input data (and notation) used in the Bayesian analysis. Data and analysis can proceed at the Unstratified (U) or Stratified (S) level.

Data	**Notation**	**Chinook**	**Coho**	**Sockeye**	**Steelhead**
Tagged at GW	*GW.tagged*	U,S	U,S	U,S	U,S
Tagged at GH	*GH.tagged*	U,S	No tagging done	No tagging done	U,S
Mortality / fallback studies These are not stratum specific					
Net number of radio tagged fish released	R_i^{IM}, i=1,…, n_{IM}	U	U	U	U
Observed number of mortality or fall back	M_i^{IM}, i=1,…, n_{IM}	U	U	U	U
Genetic studies for the proportion of fish that spawn between GW and GH. These are not stratum specific					
Estimated proportion of fish that spawn between GW and GH	$\hat{p}_i^{SP}$ with uncertainty $SE(p_i^{SP})$	U	U	U	U
Tag loss studies These are not stratum specific					
Number of fish inspected at KW without an observable tag	NOT_i^{Total}	U	U	U	U
Number of fish inspected at KW with an observable tag	OT_i^{Total}	U	U	U	U
Number of fish subsampled from those	I_i^{TL}	U	U	U	U

with no observable tag					
Number of fish in subsample that have lost a tag	I_i^{TL}	U	U	U	U
Tags removed in the Nisga'a Domestic Harvest					
ISF monitored catch and tags seen	$IS_h^{TagsObserved}$ IS_h^{Catch}	U, S by fish variables and river stratum	U, S by fish variables and river stratum	U, S by fish variables and river stratum	U, S by fish variables and river stratum
Estimated Total Catch in domestic harvest from harvest reports	$EstTotalCatch_h$ $SE(EstTotalCatch_h)$	U, S by fish variables and river stratum	U, S by fish variables and river stratum	U, S by fish variables and river stratum	U, S by fish variables and river stratum
Tags removed in the GW fishwheel harvest	*tags.GW.fishwheel.harvest* Known exactly.	U,S	U,S	U,S	U,S
Tags removed in the GH fishwheel harvest	*tags.GH.fishwheel.harvest* Known exactly.	U,S	U,S	U,S	U,S
Tags removed in the demonstration harvests	*tags.demo.harvest.est* *tags.demo.harvest.se* estimated by surveys.	U,S	U,S	U,S	U,S
Tags removed in fishwheel harvests	*tags.GH.fishwheel.harvest.est* *tags.GH.fishwheel.harvest.se* estimated by creel (and other) surveys.	U,S	U,S	U,S	U,S
Tags removed in the	*tags.demo.harvest.est* *tags.demo.harvest.se* estimated by surveys.	U,S	U,S	U,S	U,S

demonstration harvests					
Number of fish inspected at MF, KW, DC, CR etc	*MF.inspect* *KW.inspect* *DC.inspect* *CR.inspect*	U,S U,S U,S U,S	U,S	U,S	U,S
Number of tagged fish observed in those inspected	*MF.tagged* *KW.tagged* *DC.tagged* *CR.tagged*	U,S U,S U,S U,S	U,S	U,S	U,S
Prior on the marked fraction in stratum 1	*alpha.GH.marked.fraction* *beta.GH.marked.fraction* specify parameters of the prior distribution Beta(alpha,beta).	U,S	U,S	U,S	U,S
	Default values are	(1,5)	(1,5)	(1,5)	(1,5)
Prior on the difference in the marked fraction between the two strata	*mu.delta* *sigma.delta* specify the parameters for the prior distribution for the log(odds ratio) of the marked fraction between the two strata. N(mu, sigma)	S	S	S	S
	Default values are	(0,1)	(0,1)	(0,1)	(0,1)

U – values unstratified by size or sex are available
S – values stratified by size are available

Table A-2. Parameters and unobserved stochastic nodes used in the Bayesian model.

Parameter or unobserved stochastic node	Notation	Chinook	Coho	Sockeye	Steelhead
Tagged fish removed between GW and GH					
Tagged fish lost to initial mortality / fall back	*tags.init.mort*	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Tagged fish that spawn between GW and GH	*tags.spawn.GW.GH*	S but same probability used for all strata	Not applicable	Not applicable	S but same probability used for all strata
Tagged fish that lose tags between tagging and recovery at MF	*tags.tagloss*	Not applicable	S but same probability used for all strata	S but same probability used for all strata	Not applicable
Tags removed in the Nisga'a Domestic Harvest	$tags.NH.GW.GH$	S by fish variables and river stratum	S by fish variables and river stratum	S by fish variables and river stratum	S by fish variables and river stratum
Tagged fish lost to the demonstration harvest	*tags.demo.harvest*	S	S	S	S
Tagged fish lost to the sport harvest	*tags.sport.GW.GH*	S	S	S	S
Initial Mortality /fallback					
Probability of mortality in mortality study year *i*	p_i^{IM}	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Mean and standard deviation (on	$\mu_{logit(p)}^{IM}, \sigma = \sigma_{logit(p)}^{IM}$	S but same probability	S but same probability	S but same probability	S but same probability

logit scale) or initial mortality probability across mortality studies		used for all strata	used for all strata	used for all strata	used for all strata
Probability of initial mortality or fall back for the tagged fish from GW and GH	$p^{IM}_{tagging.year}$	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Genetic studies to estimate probability of spawning between GW and GH					
Probability spawning between GW and GH in genetic study year *i*	p_i^{SP}	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Mean and standard deviation (on logit scale) of probability of spawning between GW and GH	$\mu^{SP}_{logit(p)}, \sigma = \sigma^{SP}_{logit(p)}$	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Probability of spawning between GW and GH for genetic study year ip_i^{SP}	$p^{SP}_{tagging.year}$	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Tag loss studies					
Tag loss probability	p_i^{TL}	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata
Mean and standard deviation of logit(tag loss probabilities)	$\mu_i^{TL}, \sigma_i^{TL}$	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata	S but same probability used for all strata

Nisga’a Domestic Harvest					
Tagged proportion in the ISF	$p_h^{IS.tag.ratio}$	U, River stratum	U, river stratum	U. river stratum	U, river stratum
Actual catch in river stratum h	$TotalCatch_h$	U, S, River stratum	U, S, river stratum	U, S, river stratum	U, S, river stratum
Marked fraction upriver of GH					
Marked fraction just upriver of GH	$GH.marked.fraction$	S but prior placed that odds ratio is 1	S but prior placed that odds ratio is 1	S but prior placed that odds ratio is 1	S but prior placed that odds ratio is 1
log-odds ratio of marked fraction between the two strata	*Delta*	Controls difference in marked fraction between strata	Controls difference in marked fraction between strata	Controls difference in marked fraction between strata	Controls difference in marked fraction between strata
Run size					
Run size just above GH	$GH.runsize$		Includes spawning between GW and GH	Includes spawning between GW and GH	
Run size at GW	$GW.runsize$	Includes fishwheel harvests	Includes fishwheel harvests	Includes fishwheel harvests	Includes fishwheel harvests