Species composition estimation using hydro-acoustic fish length data within a Bayesian mixture model

Catherine G.J. Michielsens, Fiona J. Martens, Jacqueline L. Nelitz and Yunbo Xie

Pacific Salmon Commission, Vancouver, B.C

Introduction

Hydro-acoustic data are routinely used to estimate daily salmon passage in the Fraser River but thus far these data have not been able to identify salmon species.

Dual frequency **ID**entification **SON**ar (DIDSON) hydro-acoustic data was examined to determine if it can be used to estimate the relative abundance of pink and sockeye salmon based on the frequency distributions of hydro-acoustic fish length as measured from the DIDSON acoustic fish images.

Objectives

1) Evaluate biological and hydro-acoustic fish length difference between pink and sockeye salmon
2) Estimate the relative abundance of both species using simulated mixtures of hydro-acoustic length data collected when only one species is dominating the migration.
3) Evaluate the impact of including prior knowledge about hydro-acoustic fish length or relative abundance of pink and sockeye salmon on relative abundance estimates
4) Estimate in-season changes in species proportions

Data collection

The post-orbital-to-fork (POF) lengths of 3465 sockeye and 1530 pink salmon caught by salmon test-fisheries, were measured (**Figure 2**).

Hydro-acoustic total length data were collected at Mission, BC, using a standard DIDSON (Xie et al. 2005) (**Figure 1a**). Fish lengths were measured using the Mark Fish feature in the DIDSON software (**Figure 1b**).

To collect hydro-acoustic length data from fish of a known species, 644 length measurements were taken in 2006 when no adult pink salmon migrate in the river. Similarly 1861 hydro-acoustic fish lengths were measured in 2007 when the salmon migration was dominated by pink salmon (**Figure 2**).

From July to September 2009, 4215 sockeye and pink salmon were measured during their migration using the DIDSON (**Figure 4**).

Methods

The proportion of pink and sockeye salmon are estimated using a mixture model which assumes that the frequency distribution of the hydro-acoustic fish length is a mixture of two normal distributions with different means.

These two distributions contribute to a different degree to the joint distribution of hydro-acoustic fish length depending on the proportion of salmon observed for each species (Fleischman and Burwen, 2003).

By applying a Bayesian version of the mixture model, additional information can be incorporated. Sockeye salmon migrate earlier than pink salmon to the river. Therefore it is possible to include mean hydro-acoustic sockeye length obtained prior to pink salmon migration and expected increases in pink proportion as additional information into the model.

Collection of high frequency hydro-acoustic fish length data using DIDSON sonar

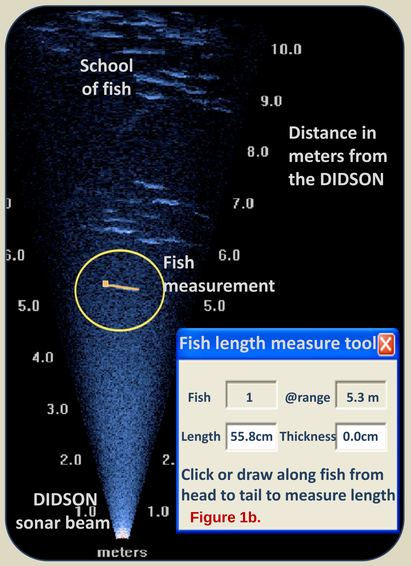

Figure 1a. Figure 1b.

Comparison of biological fish length data with hydro-acoustic fish length data

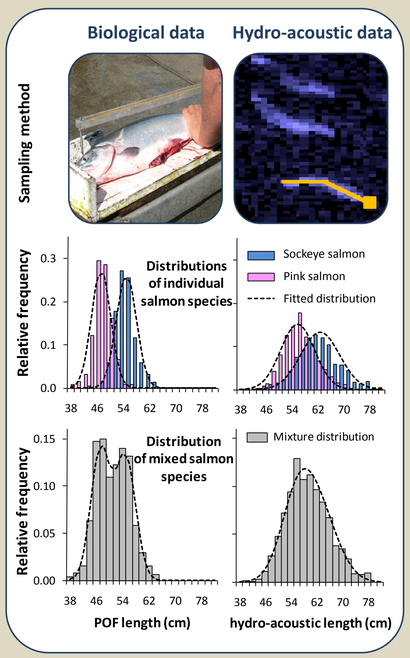

Figure 2.

Comparison of model predicted proportions of pinks against true proportions assuming different amounts of prior knowledge on mean hydro-acoustic length and pink proportions

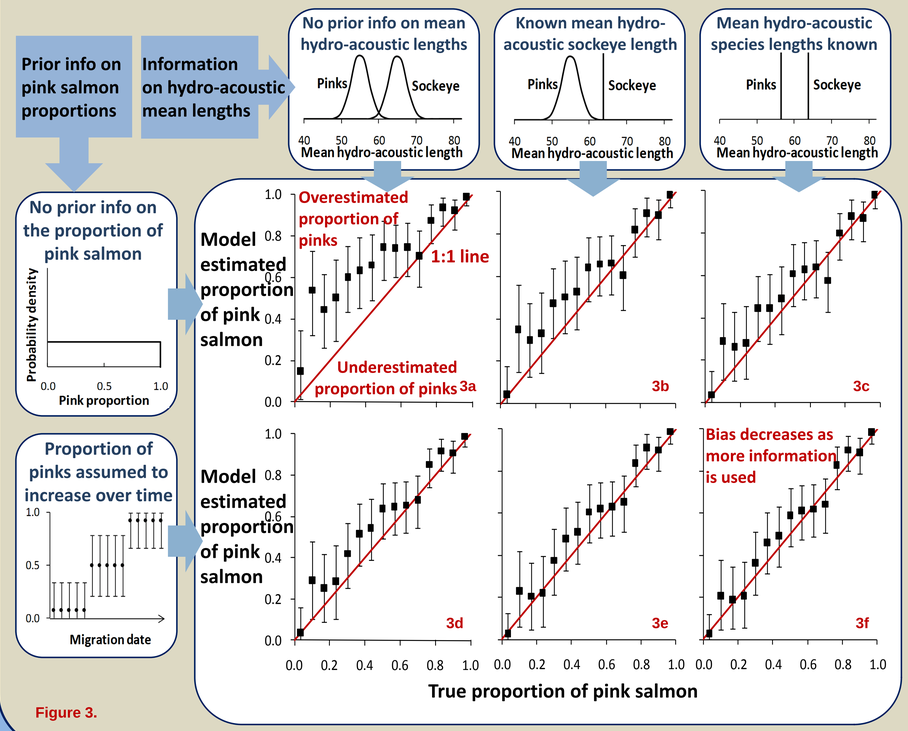

Figure 3.

Estimates of changes in proportion of pink salmon using hydro-acoustic length data collected at Mission, BC in 2009 and comparison with alternative proportion estimates

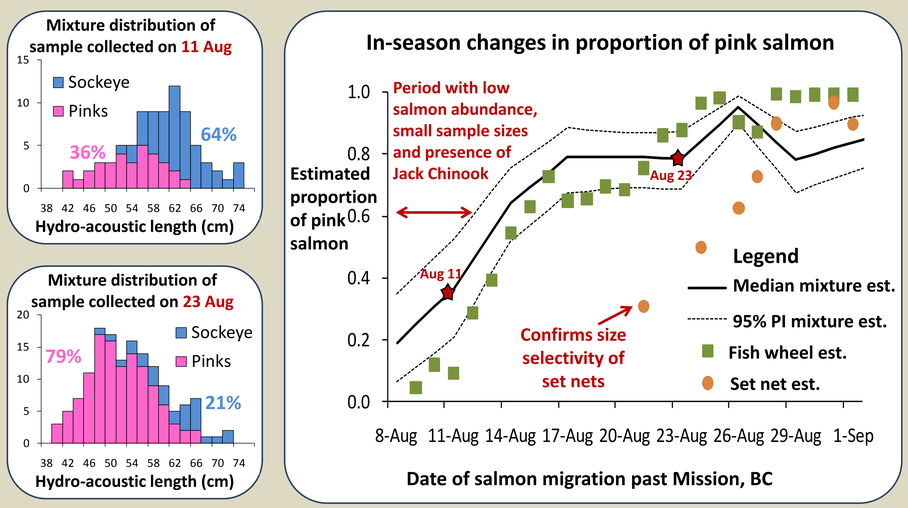

Figure 4.

Results and Discussion

Hydro-acoustic versus POF length data

- Because hydro-acoustic fish lengths are estimated from fish images obtained through sonar they are subject to larger uncertainty than POF measurements (**Figure 2**). Compared to POF length distributions, the distributions for the hydro-acoustic fish length are wider distributions, indicating larger measurement error.
- The joint distribution for hydro-acoustic length data for pink and sockeye salmon does not provide a clear indication of the peak of the length distributions of the individual species (Figure 2).

Using simulated data to evaluate model performance

- Data sets of 150 hydro-acoustic fish measurements were simulated assuming different proportions of 2007 Pink and 2006 Sockeye data. The model was run for each of these data sets to compare the estimated proportion of pink salmon against the true proportion (**Figure 3**). Prior information on mean hydro-acoustic length and expected increasing proportion of pink salmon improves the accuracy of the estimates.
- Because pink salmon migrate later than sockeye salmon, it is not possible to know the mean hydro-acoustic pink salmon length until the end of the season when all sockeye migrations have completed. When including known mean lengths, proportion estimates are less biased (**Figure 3c** vs. **3a**). It is possible to obtain similar information in-season by comparing the POF length in the current year with the POF length in previous years and to adjust the mean hydro-acoustic length accordingly.

Proportion estimates for 2009

- When using the Bayesian mixture model to estimate the proportion of pink salmon in 2009 (**Figure 4**), results can be compared against the species information obtained by fish wheel (Karl English, pers.com.) and set net sampling (Keith Forrest, pers.com.) within the same area.
- The DIDSON-derived pink proportion estimates show a similar increase in stock proportion estimates as the data collected by the fish wheel (**Figure 4**). The discrepancies near the start and the end of the time series might be due to bias caused by co-migrating species especially when abundances of pink and sockeye salmon are low. The discrepancies with the set net might be caused by the fact that the net has a greater efficiency for sockeye than for the smaller-sized pink salmon.

Precision and bias in the proportion estimates

- The higher the sockeye and pink abundance and the larger the DIDSON sample size, the better the hydro-acoustic length distributions are defined. Increasing the sample sizes would increases the precision of the estimates but would not necessarily reduce the bias.
- Biases in estimates of proportions can result when the hydro-acoustic length distribution deviates from the normal distribution. For sockeye salmon this can be due to the presence of fish of different ages. Also the occurrences of co-migrating species, especially during periods of low sockeye or pink abundance, can bias results.

References

Fleischman, S. J. and D. L. Burwen. 2003. Mixture models for the species apportionment of hydroacoustic data with echo-envelope length as the discriminatory variable. ICES Journal of Marine Science 60: 592-598.

Xie, Y., A. P. Gray, F. J. Martens, J. L. Boffey and J. D. Cave. 2005. Use of dual-frequency identification sonar to verify salmon flux and to examine fish behaviour in the Fraser River. Pacific Salmon Comm. Tech. Rep. No. 16: 58 p.

Acknowledgements

We thank Karl English (LGL limited) and Keith Forrest (PSC) for the use of stock proportion estimates obtained using fish wheel and set net sampling methods.