Using simulation to optimize the sampling design of fisheries-independent trawl surveys

Paul M. Regular, Fran Mowbray

Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, 80 East White Hills, St. John’s, Newfoundland and Labrador, A1C 5X1, Canada

# Abstract

A central question to fish stock assessments is: how many samples are needed to adequately represent the population being monitored? For assessments based on trawl surveys, our ability to estimate population characteristics, such as abundance at age, is affected by multiple levels of sampling. We therefore need to assess the impact of haul design and length and age sampling on survey accuracy to determine optimal sampling effort. In this study we a simulated repeated stratified-random samples of fish from a spatially correlated field. These samples were then analyzed and estimates of abundance at age were compared against a known truth. By varying the sampling protocol, we demonstrate that increasing the number of sets conducted per stratum provides the greatest gains in precision when sampling a clustered population. Increasing length and age sampling efforts also improved abundance estimates, but to a lesser extent. These results indicate that length and age sampling efforts for some species (e.g. cod *Gadus morhua*) in the trawl survey conducted by Fisheries and Oceans Canada, Newfoundland region, can be scaled back without significant losses to precision. Instead, time saved by reducing sampling may afford more time to conduct additional sets and/or increase sampling efforts of other species. While this simulation is tailored to a specific case, the method may be adapted to different systems and survey procedures.

**Keywords** - length and age distribution; simulation; spatial correlation; stratified analysis; statistical assessment models; survey design

# Introduction

Fisheries-independent trawl surveys have become an increasingly important research tool for the management of dynamic fish stocks. These surveys provide reliable indices of population abundance as well as estimates of various population characteristics such as length and age frequencies (Pennington and Strømme, 1998). While costly to obtain, this information forms the basis of of many stock assessments models used throughout the world. Effective and informed management decisions therefore require surveys that maximize information while minimizing the expense of data collection. Determining the number of samples required to adequately characterize a fish population is particularly challenging since data tend to be collected across one or more stages (Aanes and Vølstad, 2015). Optimal survey design is further complicated by the fact that trawls sample clusters of fish that tend to have similar characteristics, such as length, age and stomach contents. This phenomenon results in samples with positive intra-haul correlation, which can markedly reduce the effective sample size for estimating length and age frequencies of the target population (Pennington and Vølstad, 1994; Pennington *et al.*, 2002; Stewart *et al.*, 2014). The complexities of multi-stage sampling and intra-haul correlation make it difficult to answer the question *“how much sampling is enough?”*

Two main approaches have been applied to try and answer this question: one involves statistical estimators of effective sample size (Pennington *et al.*, 2002; Stewart *et al.*, 2014), and the other employs resampling methods to test alternate sampling protocol (Cerviño and Saborido-Rey, 2006; Zhang and Cadrin, 2013). Both approaches have shown that lengths are usually oversampled and a reduction in sampling can often be done without significant loss of accuracy (Pennington *et al.*, 2002). The same result holds for other positively correlated characteristics, such as age (Coggins *et al.*, 2013) and stomach contents (Bogstad *et al.*, 1995). Pennington and Vølstad (1994) concluded that the best way to improve the accuracy of surveys with strong intra-haul correlation is to take fewer samples per sampling unit and increase the total number of sampling locations. Few studies, however, have tested these conclusions using simulations. A key advantage of a pure simulation approach is that “true” population size and all stages of sampling can be controlled. The relative effects of each stage of sampling on our perception of the “truth” can then be tested. This level of testing is of particular importance for age composition estimates, and the contemporary age-based assessment models that rely on these estimates, as they are affected by all stages of sampling. A key challenge, of course, is building a simulation that emulates reality.

The objective of this paper is to use a realistic simulation to evaluate the efficacy of various sampling protocol. First we simulated a spatially autocorrelated target population, and then we simulated surveys over this field. Set density was varied across surveys, as was length and age sampling efforts. Estimates of abundance at age were then obtained using a standard stratified approach (Smith and Somerton, 1981) and deviation from the truth was assessed. Given previous findings, we predict that the greatest gains in precision will come from increasing the number of hauls; increased sampling of correlated variables (length and age) on a set by set basis will be relatively ineffective. The bottom trawl survey of cod (*Gadus morhua*) in NAFO division 3Ps, Newfoundland, was used as a case study for this analysis.

# References

Aanes, S., and Vølstad, J. H. 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. Canadian Journal of Fisheries and Aquatic Science, 72: 938–953.

Bogstad, B., Pennington, M., and Vølstad, J. H. 1995. Cost-efficient survey designs for estimating food consumption by fish. Fisheries Research, 23: 37–46.

Cerviño, S., and Saborido-Rey, F. 2006. Using the bootstrap to investigate the effects of varying tow lengths and catch sampling schemes in fish survey. Fisheries Research, 79: 294–302.

Coggins, L. G., Gwinn, D. C., and Allen, M. S. 2013. Evaluation of Age-Length Key Sample Sizes Required to Estimate Fish Total Mortality and Growth. Transactions of the American Fisheries Society, 142: 832–840.

Pennington, M., and Vølstad, J. H. 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics, 50: 725–732.

Pennington, M., and Strømme, T. 1998. Surveys as a research tool for managing dynamic stocks. Fisheries Research, 37: 97–106.

Pennington, M., Burmeister, L.-M., and Hjellvik, V. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. Fisheries Bulletin, 100: 74–80.

Smith, S., and Somerton, G. 1981. STRAP: A User-Oriented Computer Analysis System for Groundfish Research Trawl Survey Data. Canadian Technical Report of Fisheries; Aquatic Sciences No. 1030.

Stewart, I. J., Hamel, O. S., and Rose, K. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. Canadian Journal of Fisheries & Aquatic Sciences, 71: 581–588.

Zhang, Y., and Cadrin, S. X. 2013. Estimating Effective Sample Size for Monitoring Length Distributions: A Comparative Study of Georges Bank Groundfish. Transactions of the American Fisheries Society, 142: 59–67.