

## **Skipped spawning in Northeast Arctic haddock *Melanogrammus aeglefinus***

**J. E. Skjæraasen\*, K. Korsbrekke, T. Nilsen, M. Fonn, O.S. Kjesbu, G.E. Dingsør,  
R. D. M. Nash**

\*Corresponding author: [jones@imr.no](mailto:jones@imr.no)

*Marine Ecology Progress Series 526: 143–155 (2015)*

---

### **Supplementary methods**

#### **Probability based survey estimation**

##### **Probability of a sample**

Swept area estimation based on trawl surveys can produce results for average catch rates or for total abundance. In the following, all equations and estimators shown represent single strata which will be combined to form estimates at the population level (summation, weighted average).

Total abundance estimates are probability based, but rely on assumptions regarding catchability of the trawl. The name “swept area” implies some assumptions about the efficiency of a trawl targeting demersal species, but can also be expressed as a probability:

$$\frac{n \cdot d_i \cdot E}{A} \tag{1}$$

is the probability assigned to each observation (trawl station) with

$n$  being the number of trawl stations (usually within a stratum)

$d_i$  being the towed distance at station  $i$

$E(l)$  is the effective fishing width (in this case assumed to dependent on length of the fish  $l$ )

and

$A$  is the area in question (usually a single stratum)

Implied in this is the assumption that a trawl station is a random sample from a population of “all” possible trawl samples and may be thought of as a population of rectangles adding up to the total survey area. The area swept associated with a trawl station is very small relative to the area sampled and there is no need for any finite population correction (FPC) (Cochran 1977).

Catches from the Barents Sea winter survey are usually subsampled and the fraction subsampled,  $q_i$ , will represent the probability of the subsample as a random sample within the trawl station and will usually be calculated as the weight of the subsample divided by the weight of the catch (by species). The probability assigned to subsample will be:

$$p_i = \frac{q_i \cdot n \cdot d_i \cdot E}{A} \quad (2)$$

Please note that this is a single subsample within our first stage sample (the trawl station).

The Barents Sea winter survey operates with two stage subsampling of catches with the first stage random sample used to describe the length distribution. The first stage subsampling is not viewed as a subsample as such, but as an adjustment of the sampling probability. The inverse of  $q_i$  is often referred to as the “raising factor” and the reduction in precision due to this subsampling depends on how it affects the number of fish measured in different length groups,  $f_{i,l}$  and the relative proportion  $f_{i,l}$  represents of the total length sample  $\sum_l f_{i,l}$ . We

assume that this error is negligible compared to between station variation.

## The use of biological samples of individual fish

The second stage subsampling for additional biological parameters such as age, sex, maturity is a fixed number of fish sampled from each 5cm length group (current practice is one haddock per 5cm length group). The probability assigned to a single fish within a length group is:

$$r_{i,l,j} = \frac{m_{i,l}}{f_{i,l}} \quad (3)$$

$m_{i,l}$  is the number of stage two biological subsamples at station  $i$  in length-group  $l$

$f_{i,l}$  is the length frequency observed in stage one subsample at station  $i$  and length  $l$ .

$j$  is a single fish index and for the ease of notation and convenience uniquely identifying every single fish sampled.

The probability associated with a single biological sample can then be written as:

$$\pi_j = p_i \cdot r_{i,l,j} \quad (4)$$

Let our indicator variable  $y_j(a) = 1$  if fish  $j$  is of age =  $a$  and 0 otherwise.

Using Horvitz-Thompson estimator (Horvitz & Thompson 1952) our index of abundance at age is:

$$\hat{I}_a = \sum_j \frac{y_j(a)}{\pi_j} \quad (5)$$

with corresponding variance:

$$\hat{V}(\hat{I}_a) = \sum_j y_j \frac{1 - \pi_j}{\pi_j^2} + \sum_{j \neq k} y_j y_k \frac{\pi_{jk} - \pi_j \pi_k}{\pi_{jk} \pi_j \pi_k} \quad (6)$$

$\pi_{jk}$  is the joint probability of any two single fish samples. Observations within single trawl stations are known to be highly correlated and the joint probability is generally not known.

Assuming that most of the variation observed stems from station to station variation, a

reasonable estimate of variance can be made using a stratified bootstrap approach (resampling trawl stations stratum by stratum). Please note that the inverse of  $\pi_j$  associated with each single can be viewed as a weighting factor and as shown above the weighting factors will sum up to population estimates (swept area). Such weighting factors are used to produce weighted estimates of population parameters. The following is an example on how an estimator of mean length at age will look:

$$\hat{L}_a = \frac{\sum_j (y_j(a) \cdot w_j \cdot L_j)}{\sum_j (y_j(a) \cdot w_j)} \quad (7)$$

with

$w_j$  being the inverse of  $\pi_j$  and

$L_j$  being the length of fish  $j$ .

And due to the intra haul correlation, variance will be estimated using a stratified bootstrap.

## LITERATURE CITED

Cochran WG (1977) Sampling Techniques. Third Edition, New York: John Wiley & Sons, Inc.

Horvitz DG, Thompson DJ (1952) A generalization of sampling without replacement from a finite universe. J Am Stat Assoc 47:663–685