Quantifying uncertainty in abundance estimates of herring

Magne Aldrin (1), Anders Løland (1), Alf Harbitz (2)

(1) Norwegian Computing Center, Norway. (2) Institute of Marine Research, Observation Methodology, Norway.

Norwegian spring spawning herring

The largest fish stock in the Northeast Atlantic.

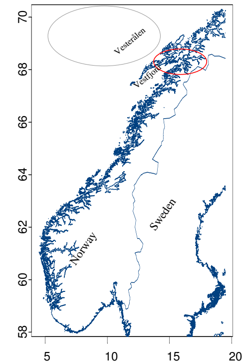
Precise abundance estimates with quantified uncertainties are important for a sustainable management of the stock.

Our goal is to estimate with a quantified uncertainty

N = the total number of fish, and

N_a = the number of fish of age a.

Fig. 1



Winter survey

In the winter, the whole stock is located either in the sea outside Vesterålen (grey) or in the Vestfjord system (red), see Fig. 1.

Abundance estimates are based on a survey over 2-3 weeks, with (Fig. 2)

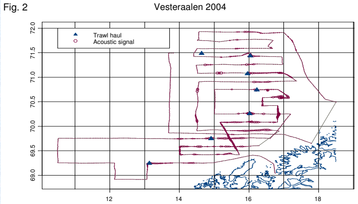
acoustic measurements of the relative density of fish in transects,

•each fish reflects a signal assumed to be proportional to its squared length,

trawl samples with age and length measurements of individual fish and there are

•10-20 trawl hauls, with age and length measurements of 100 fish in each haul.

Fig. 2



Relationship between number of fish, acoustic signal, length and age

Let S be the hypothetical undisturbed acoustic signal (per area unit) over an area of size A, if one could survey all the area in a snapshot, without measurement errors.

Let L^2 be the averaged squared length over all fish in the area.

Then the total number of fish could be calculated as

$$N = S \cdot A / (c \cdot L^2),$$

where c is a constant (assumed known exactly here).

Let further P_a be the true proportion of fish with age a.

Then the number of fish of age a is given by

$$N_a = N \cdot P_a.$$

Abundance estimate

The official abundance estimates are based on the formulas above, with simple plug-in estimates for the unknown quantities S, L^2 and P_a.

1) First, the raw signal is corrected for the shadowing effect, i.e. that fish in lower layers reflect weaker signal if they are in the shadow of fish in upper layers.
2) Then S is estimated by the sample mean of the corrected signals along the transects.
3) L^2 is estimated by the sample mean of all squared fish lengths in the trawl hauls.
4) P_a is estimated by the sample proportion of all fish of age a in the trawl hauls.

The abundance estimate is not based on a model, and it is not obvious how to quantify its uncertainty.

Three sources to uncertainty

1) The shadow correction (see box to the right of Fig. 3) is uncertain due to the uncertainty in the parameter γ.
2) Acoustic measurements are made in only a small part of the area over a period of several weeks. Note that, since the fish are moving around, measurements at the same location, but at different time, are in general unequal.
3) The lengths and ages are measured for only a small amount of the fish, and on few locations. There are large heterogeneities between trawl samples, i.e. one sample may consist of young fish, and another sample of older fish (Fig. 3).

Note that the three sources are independent, since they are based on different and independent data.

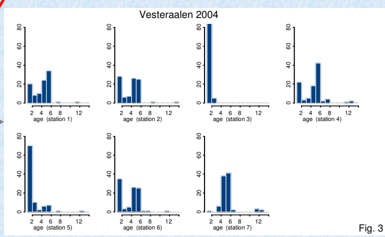


Fig. 3

Shadow correction

The raw acoustic signal is shadow corrected by the equation (Zhao and Ona, 2003)

$$S_{corr} = \frac{1}{K \cdot \gamma} \ln\left(\frac{1}{1 - K \cdot \gamma \cdot S_{raw}}\right),$$

where K is a known constant. The parameter γ has been estimated by separate experiments, with estimate 2.41 and standard deviation 0.33.

Quantification of the estimation uncertainty by bootstrapping

1) First, 1000 replicates of shadow corrected signals are generated by simulating values of γ from a normal distribution N(2.41,0,33²).
2) Each replicate consists of a space-time series of corrected signals along the transect. We ignore the spatial structure, and consider the measurements as a time series along the time axe. Note, however, that the spatial pattern still is partly covered by the time axis. Then we generate one bootstrap sample based on block bootstrap (Künsch, 1989) for a stationary time series.
3) Each replicate of the acoustic signal is linked to a bootstrap replicate of the age/length data. The bootstrap of the trawl samples has to take into account the inhomogeneity between trawl hauls. First, n trawl hauls are drawn with replacements from the n real trawl hauls. Within each bootsrap haul, 100 fish are drawn with replacements.
4) For each set of bootstrap replicates, bootstrap estimates N^* and N_a^* are calculated. The estimation uncertainty is finally found from considering all 1000 replicates.

Uncertainty in the estimate of N

The coefficient of variation for total number of fish N in Vesterålen and Vestfjord in 2004 is 0.16. Below it is decomposed into the contribution from each of the three sources.

Age/length	Shadow correction	Spatial
4%	1%	95%

The uncertainty is dominated by the spatial uncertainty due to limited coverage of the acoustic measurements. If the precision of the estimate of the total number of fish is to be increased, the acoustic measurements should be extended by denser transects.

Uncertainty in the estimate of N_a

The uncertainty of the estimates of the number of fish N_a is illustrated by Fig. 4, which shows the bootstrap distribution.

For each age we can calculate the standard deviation $\hat{\sigma}_a^2$. A suitable summary of the total uncertainty within age is given by

$$\sqrt{\sum_a \hat{\sigma}_a^2}.$$

This is decomposed into the contribution from the three sources below.

Age/length	Shadow correction	Spatial
40%	1%	59%

Thus, to reduce the uncertainties in the age specific estimates one should both extend the acoustic measurements as well as take more trawl hauls (more hauls are more important than more fish within each haul.) The uncertainty due to the shadowing effect is still very small.

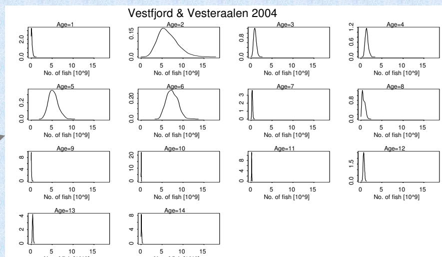


Fig. 4

Presented at the 25th European Meeting of Statisticians, 24th-28th of July 2005, Oslo, Norway.