ICES CM 2009/I:20
Not to be cited without prior reference to the author

Minimalistic surveys clear the way for priorities in the ecosystem approach

Olav Rune Godø, Vidar Hjellvik, Dag Tjøstheim, and Knut Korsbrekke

Standardized surveys are an integral part of most fish stock assessments. Ship time is expensive and new demands set by the ecosystem approach challenge present survey and sampling strategies with respect to coverage of all components of an ecosystem. Surveys are normally designed to cover the stock distribution and to estimate relative abundance with acceptable precision. Taking ecosystem considerations means more attention to non-commercial species and ecosystem processes including predator – prey interaction and biophysical interactions.

In this paper we explore alternative use of survey effort based on data from the Barents Sea winter survey. Will a drastic reduction of survey effort reduce its consistency and will consistency change with sampling strategies? We explore the efficiency of combining minimalistic surveys with standard hydrographical sections. We analyse data of cod which is the main target species of our survey. Precision of abundance indices is reduced roughly according to reduced number of hauls. Observed trends in population abundance and overall mortality appear to be similar and so are also the variability bands around the average cohort trend. But the minimalistic surveys are more susceptible to year to year variability. There are systematic bias effects related to both age and mortality which are independent of effort and design and this represents a larger problem then sampling variability. A distribution of trawls along hydrographical sections reflects the standard survey well. We suggest that a reduced survey, e.g. carried out along hydrographical transects every second year, might clear capacity for essential ecosystem investigations including studies of unresolved year effects in surveys otherwise not funded.

Keywords: Survey strategy, bottom trawl survey, reduced effort, cod

Olav Rune Godø, Vidar Hjellvik, and Knut Korsbrekke: Institute of Marine Research, P.O. Box 1870 Nordnes, 5817 Bergen, Norway. Tel: +47 55238500, fax: +47 55238531. Dag Tjøstheim: University of Bergen, P.O. Box 7803, 5020 Bergen, Norway. Tel: +47 55582806, fax: 55589440.

Correspondence: Olav Rune Godø, Institute of Marine Research, P.O. Box 1870 Nordnes, 5817 Bergen, Norway. Office address: C Sundts gate 64, 5004 Bergen, Norway. Tel: +47 55238675; fax: +47 55238531; e-mail: olavrune@imr.no

Introduction

Scientific surveys are an integral part of assessment and management of marine resources (Gunderson, 1993). Survey effort is expensive, but cost reduction through development of advanced models to compensate for the lack of such survey data is not an alternative for precise stock assessment (Nrc, 1998). Acoustic surveys are used for pelagic resources (Simmonds and MacLennan, 2005) while bottom trawl surveys, sometimes combined with acoustic surveys, are most common for demersal or semi-demersal species (Karp and Walters, 1994, Godø and Wespestad, 1993). Due to the high cost, cost – benefit evaluation is needed for making rational decisions about survey strategies including the level of survey effort. Improved estimates of uncertainties in surveys have been requested as a high priority item, but still assessment and prediction for use in management decisions are often only based on trends in stock properties. In later years new demands set by the ecosystem approach to fisheries (EAF) requires new research and monitoring (Browman and Stergiou, 2004, Garcia, 2003). This has renewed the discussion on use of available field effort. The main motivation for this work arises from demands set by EAF where information about system health is as important as the state of the exploited stocks. Collecting ecosystem information demand new effort, which may not be available unless existing effort is reallocated. Such an option will be considered in this paper.

The Norwegian combined bottom trawl and acoustic surveys were initiated with the aim of producing fisheries independent information about gadoids with particular emphasis on estimating abundance before recruitment to the exploited stock (Hylen et al., 1985). It was early recognized that age 1-3 fish were heavily underrepresented in the catches (Engås and Godø, 1989) and a new rockhopper ground gear improved catchability substantially (Godø and Sunnanå, 1992). It is well known that standard error of survey estimates varies with the level of effort (number of stations, degree of coverage) (see e.g. (Aglen, 1994). This gives an indication of the repeatability of estimates under identical model assumptions, but is it necessarily a good indicator for the quality of the average survey estimates of stock properties? Systematic annual changes in natural fish behaviour, cannibalism and the physical environment may create unexpected variation in bottom trawl and acoustic survey results. Our findings from the present analysis will therefore also relate to the sampling needs in the parallel acoustic survey. Correct trends of stock properties rather than precise yearly estimates may be more essential to assessment of present status of stocks.

How will the level of survey effort and its distribution in time and space affect the consistency and trends of time series of survey indices? In this paper we explore these issues by using a long time series of data from a bottom trawl. The survey is a combined bottom trawl acoustic survey and the same approach might be applicable for the acoustic data. The objectives of the paper are twofold: First we will assess the impact on key survey estimates and associated time trends caused by drastically reducing survey effort. Further, we will evaluate if drastically reduced effort can be concentrated along sections generally used for monitoring the physical and biological environment of the ecosystem. With a positive outcome the approach can be further developed to enhance the EAF through alternative survey strategies.

Material and methods

Since 1982, a combined bottom trawl and acoustic survey for cod and haddock has been undertaken every winter in the Barents Sea. Independent of this, environmental data have been collected annually along some fixed transects in the same region (Fig. 1). If the trawl survey and the environmental survey could be combined by taking trawl stations along the environmental transects while collecting the environmental data, this would be a cost-effective allocation of effort. To study effects on the precision and accuracy of such an approach, we calculated abundance indices and corresponding mortality estimates using sub-sets of the historical trawl survey data from 1989 to 2009.

The data sets used were Ω_{full}, Ω_{min}, and $\Omega_{\mathrm{min+}}$, where Ω_{full} contains all of the trawl stations, Ω_{min} is the subset containing all trawl stations no more than 10 n.mi. away from the nearest environmental transect (13.5 % of all stations), and $\Omega_{\mathrm{min+}}$ contains all stations in Ω_{min} , and in addition stations closer than 10 n.mi. to one of the additional diagonal connecting transects (22.4 % of all stations) (Fig. 1). The first connecting transect joins the two westernmost environmental transects, and the other starts at the southern end of the easternmost environmental transect and goes north- and eastwards, where much of the young fish is located.

Simplified abundance at age indices $I_{\mathrm{Y,A,full}}$, $I_{\mathrm{Y,A,min}}$, and $I_{\mathrm{Y,A,min+}}$ for age A=1, …, 9 and 10 in year Y=1989,…,2009 were calculated as mean catch in numbers per towed nautical mile, and the corresponding survey mortality estimates as $Z_{\mathrm{Y,A,type}} = \log(I_{\mathrm{Y,A,type}} / I_{\mathrm{Y+1,A+1,type}})$, for type = full, min, or min+. The age-length keys used for $I_{\mathrm{Y,A,full}}$, $I_{\mathrm{Y,A,min}}$, and $I_{\mathrm{Y,A,min+}}$ were calculated from the trawl catches in $\Omega_{\mathrm{Y,full}}$, $\Omega_{\mathrm{Y,min}}$, and $\Omega_{\mathrm{Y,min+}}$, respectively.

To compare the development of the various year classes over time, including their predictability, we used the Spearman rank correlation (Mann, 2001) for the year classes 1988 to 1999. If a year class has a high rank; i.e. is larger than most other year classes when it is young, the rank correlation shows whether or not it keeps its high rank as it grows older. For each age we order the year classes from largest to smallest and study the change in ranking with increasing age. If there are no changes, the rank correlation is one, if ranking is reversed it is -1, and if the rank correlation is close to zero, this indicates lack of consistency in the ranking and therefore lack of predictability. The rank correlation is better suited for this than ordinary correlation because it is more robust and does not depend on a normal distribution. This approach also permits a comparison of rankings for the minimalistic surveys against the full surveys.

The uncertainty of the indices was assessed using the bootstrap. Because of strong geographical trends in fish density, a stratified bootstrap approach was chosen. Let $N_{\mathrm{Y,A,}i}$ denote the number of fish of age A caught per nautical mile at trawl station i in year Y. Note that for a given trawl station this number will typically depend on which index we calculate ($I_{\mathrm{Y,A,full}}$, $I_{\mathrm{Y,A,min}}$, or $I_{\mathrm{Y,A,min+}}$) since the age-length key depend on the set of trawl stations used. The stratification was done by fitting a gam (generalized additive model) with latitude and longitude as covariates to $\left\{x_i \square \log(N_{Y,A,i}+1);\quad i \in \Omega_{Y,\mathrm{full}}\right\}$, i.e. using all trawl stations, to create strata each with approximately the same fish density. Based on the 0, 20, 40, 60, 80 and 100 % quantiles of the fitted catches $\left\{\hat{x}_i\right\}$, the trawl stations were allocated to five strata. For the full survey, stations i with $q_0 \le \hat{x}_i < q_{20}$ were allocated to stratum one, stations with $q_{20} \le \hat{x}_i < q_{40}$ to stratum two, etc. For $I_{\mathrm{Y,A,minl}}$ the 0, 20, …., 100 % quantiles of $\left\{\hat{x}_i; i \in \Omega_{Y,\mathrm{min}}\right\}$ were used for defining the strata limits, and for $I_{\mathrm{Y,A,min+}}$ the quantiles of $\left\{\hat{x}_i; i \in \Omega_{Y,\mathrm{min+}}\right\}$ were used. The stratification method is illustrated in Figure 2.

Bootstrapping was then done by resampling the catches $N_{Y,A,i}$ within each of the five strata separately. That is, approximately $n/5$ stations were drawn with replacement from each stratum, where n is the total number of stations in Ω_{full}, Ω_{min} , and $\Omega_{\mathrm{min+}}$, for the three index types, respectively. For bootstrap replicate j, $j = 1,\dots,m$, the bootstrapped index $I^{*,j}_{Y,A,\mathrm{min}}$ was calculated as $n_{\mathrm{min}}^{-1}\sum_i N^{*,j}_{A,Y,i}$, where n_{min} is the number of stations in $\Omega_{Y,\mathrm{min}}$, and $i \in \Omega_{Y,\mathrm{min}}$. The variance and confidence intervals for $I_{Y,A,\min}$ were then calculated from the m $I^{*,j}_{Y,A,\min}$'s. The same procedure was used to calculate the variance and confidence intervals for $I_{Y,A,full}$, and $I_{Y,A,min+}$.

To check for systematic trends over years and/or age groups in the indices based on reduced survey as compared to the full survey, we calculated the ratios $I_{Y,A,\text{full}} / I_{Y,A,\min}$ for each year Y and age A.

Results

We organize our results by first studying the variability in the abundance indices by estimating means and bootstrapping samples from the full survey. This leads to studies of effects associated with sampling strategies (systematic, random and along transects selection from the full survey data set). Assessment and management depend on trends in the development of year classes including the consistency in measures of mortalities, which is given particular attention. Finally, by means of rank correlations we study consistency in the abundance indices by year and age over the whole time series.

By estimating means and bootstrapping to obtain measures of uncertainties we found that the indices of abundance by length from full and reduced surveys followed similar trends as exemplified in Figure 3. When using 22.4 % of the effort (Ω_{min+} survey), 10-43 % of the indices fall outside the 95 % confidence limits of the full survey with the highest percentages for large and small size classes. There seem to be no systematic trend in the relationship between full and reduced survey indices. As expected, reduced surveys are more susceptible to effects caused by the skewed distribution of trawl catches (Smith, 1988). This is illustrated in Figure 4 where the index and its std is totally driven by inclusion or exclusion of the extreme catches. In general, when reducing survey effort, precision (confidence limits (cl) in Figure 3) is reduced slightly less than expected according to the square root of n law (expected cl=$\sqrt{100/22.4}\approx 2$). This will be further dealt with for age based estimates. The 13.5 % survey results underlined the same tendency with showing higher variability and being more sensitive to extreme catches.

The information used in stock evaluation and management advice are age based and thus also include errors originated from imprecise aging in addition to sampling. We found that trends in indices from reduced surveys by age/year class over the time period are similar to those from full surveys (Fig. 5 and 6). The relationships between full and reduced survey (Ω_{min+}) indices by year and year classes vary around 1 with no particular trend (Figure 5). The reduced 13.5 % survey (Ω_{min}) seems to have overall lower indices for young and old fish compared to the full survey. The development of the average indices by age (Figure 6, lower right panel) not only demonstrates similarity in trend, but also a remarkable similarity in dynamics of standard errors of the average indices.

When estimating mortalities (Figure 7) from the survey data, three features are apparent:

1. Young fish (ages 1-3) have a high occurrence of negative survey mortalities, mostly so for age 2 fish. Also variability is high and reduced survey mortalities (particularly the Ω_{min} survey) show a more erratic variation from year to year.

2. Mortalities for average sized fish (ages 4-6) are more stable and show overall positive mortalities both for reduced and full surveys.

3. Mortalities of old fish (ages 7-9) vary around 1, but variability is even higher than for age 4-6 fish.

Although uncertainties of estimates of abundance and mortality from reduced surveys increased, we found that bootstrapped standard deviations are somewhat lower than expected from the square root n rule for all ages except age 8 and 9 fish (Figure 8). This is consistent with the confidence limits of Figure 3. Mortality estimates from the three surveys, particularly for ages 4-6 vary uniformly over time. Nevertheless, the mortalities for these ages varied between 0 and 1.5, which is substantially more than expected according to the analytical assessment (Ices, 2007).

We further studied the consistency in the survey indices through a rank correlation analysis in Figure 9. This analysis gives us an insight in the consistency or lack thereof of development of the indices by year class from one year to the next, and thus also indicates the predicting power of survey indices. First, we see that the full survey generally has the highest correlations with the 22.4 % survey not far behind. Age 1 and 2 fish indices have minor predicting power of commercially sized fish (age 5+). Apparently no young fish index has any predicting power of the oldest fish (ages 9+). The negative mortalities of Figure 7 for young fish demonstrate that in absolute terms these are meaningless. Conceivably their ranking could carry some predictive power, but Figure 9 largely dispels this hope.

Discussion

Surveys provide essential information for traditional stock assessment and management (Nrc, 1998). Their importance for implementing EAF will be even greater as commercial data only include data on exploited stocks and EAF will need a variety of other information to manage human impact on the ecosystem as a whole. Our objectives are not only focused on cost efficiency of existing surveys but rather how we can improve survey time series. Further, we want to discuss survey effort in a wider perspective, e.g. to evaluate how we can free survey effort to meet research requirements set by EAF. We exemplify a strategy combining data collection along environmental transect, but other strategies might be as good or better. Finally, we will discuss the impact of our findings on the simultaneous acoustic survey.

We considered various bootstrap based approaches for assessing the uncertainty in the indices. The standard bootstrap assumes an identical, independent distribution of the bootstrapped entity, and was not well suited to our trawl catches due to the strong geographical trends in fish density. The second approach was to fit a gam to each of the sets $\boldsymbol{\Omega}_{\text{full}}$, $\boldsymbol{\Omega}_{\text{min}}$, and $\boldsymbol{\Omega}_{\text{min+}}$, and bootstrap the residuals, but the gam fits to the reduced surveys were generally better than the fit to the full survey, yielding smaller residuals, and hence smaller uncertainty estimates. The third approach was to fit a gam to the full survey only, assuming that this would estimate the “true” fish density best, and then bootstrap the residuals, but since the catch at age at a given station was different for the three surveys due to different age-length keys, this did not work either, yielding larger residuals (in absolute values) for the reduced surveys. In the approach we ended up with, the gam fit was used only for handling the geographic trend (Figure 2) in fish density by stratification, and should not favour any of the survey types.

As we have explained, there are strong operational reasons for choosing the hydrographical sections as a basis for the reduced surveys, since the general physical monitoring is taking place along these transects. From a purely statistical point of view there are of course "alternatives" such as systematic and random sampling. For completeness we carried out a large number of simulations, where e.g. the systematic ones were taken as the samples obtained by choosing every other observation, every third observation and every fourth observation etc.. None of the systematic or random sub-surveys gave results that were significantly better than the transect surveys when it comes to describing the trends or the size of standard errors. All of the sub-samples (including the transect-based) are sensitive to occurrence of few very large catches. These will generate extreme estimates depending on whether none, or one or more of the large catches are included in the minimalistic survey.

Predicting power of age 1-2 indices is negligible for the major age groups of the commercial catch even though the survey is designed for estimating recruitment. This might be expected as many factors like capelin abundance and cannibalism have a very strong effect on year class strength as observed at higher ages. The analytical stock assessment (ICES, 2007) uses estimates of natural mortality that implies a large variation in survival from age 1 to age 3.

In summary the analyses show that full survey estimates are more precise compared to reduced surveys along transects but difference is somewhat less than expected from pure statistical sampling arguments. All estimated trends are similar and consistency is maintained when reducing effort. Most important, changes in mortalities are similar for all surveys but these changes do not necessarily reflect mortality changes as estimated by the analytical assessment(Ices, 2007); systematic year effects are reflected similar in full and reduced surveys. These findings raise several important questions:

1. Are standard error estimates relevant measures for quality of stock property estimates?

Standard errors give an impression of the repeatability of estimates under identical model assumptions. A tool for maintaining identical model assumptions is the bootstrap. However, during a survey homogenous conditions cannot be expected to hold for the entire region surveyed. Typically, the standard error will vary from one subregion to another. We have therefore used stratification for the bootstrap. Another problem is the statistical dependence in the time series of measurements as we proceed from one station to another. This has been sought taken care of by implementing the GAM.

As a rule the standard deviation is proportional to $1/\sqrt{n}$, where n is the total number of observations. For a reduced survey where only 22.4% or 13.5% of the measurements are used, we therefore expects an increase in the standard error by a factor of $\sqrt{100/22.4}\sim2.1$ or $\sqrt{100/13.5}\sim2.7$, respectively. Our bootstrap calculations show that we are able to do somewhat better for the reduced surveys; see Figure 8 for the case of a reduction to 22.4% of the total effort, where the average increase in the standard error is below 1.9. A reason for this may be that for the reduced surveys there is a larger degree of dependence between measurements because of the larger average distance between measurement stations.

The signals from survey results, like indices of abundance and mortalities, seem to be consistent between full and minimalistic surveys. This is on one hand positive since it indicates that the surveys are robust. On the other hand the signals are much stronger than what we should expect from the perception produced in the analytical stock assessment. Random year effects in a survey (high positive correlation between age group indices) will have a tendency to produce a negative first order autocorrelation in the survey mortality series. This is not the case here and we consider this to be similar to the effects caused by variable **survey conditions** as described by (Godø and Wespestad, 1993) or as a change in bias that needs to be corrected for as exemplified for acoustic surveys of pelagic fish by(Løland et al., 2007). For example, variation in vertical distribution may affect bottom trawl estimates and inversely influence the simultaneous acoustic survey (Godø and

Wespestad, 1993, Jakobsen et al., 1997). In any survey time series there is a need to analyse what influence the stock assessment most; imprecise annual estimates or systematic year effects? In case systematic year effects dominate the uncertainty, it might be favourable to reallocate effort to collect information that enables quantification of the impact of such phenomena.

2. Can we reduce sampling without corrupting time series?

Based on the above we conclude that the signal in the survey may be severely affected by other factors than the number of stations. In fact, an important survey like the Barents Sea bottom trawl survey has serious problems related to year to year variations caused by unknown factors. In this paper we have not studied in details the uncertainties reflected in the standard errors relative to year to year changes in survey efficiency. Nevertheless we think that our analysis demonstrate the need for caution with respect to choice of survey strategy. Minimalistic survey will reduce precision in survey estimates, but if some of this effort could be used to quantify **survey condition,** and thus explain and compensate for year to year variation in survey efficiency, the overall result could be improved.

3. Are minimalistic sampling along transects an interesting avenue to pursue for efficient ecosystem monitoring?

Environmental monitoring has traditionally been carried out along transects. In the Barents Sea these are positioned across current systems so as to systematically cover areas with highestvariability. As marine life, including many of the key species of the ecosystem in this area, is totally dependent of influx of heat and recruitment from the outside through the major oceanic currents, it is not unexpected that a survey design using these transects might be favourable. It is, however, seen that by including the connecting transects improved the situation as these were positioned through hot-spots in the distribution pattern of cod.

Survey effort is expensive and there are reasons to believe that even a strong political goal like EAF may not give substantial funding effort for monitoring marine ecosystems. A scientifically based evaluation of various survey strategies is therefore essential to optimize the use of available effort to meet the new requirements. Alternatively the precautionary approach may introduce heavier regulation on commercial exploitation than needed.

Minimalistic survey strategies free effort that can be used to understand and adjust for year effects in the survey estimates. Minimalistic surveys could replace full surveys. However, it is questionable to consider this a prolongation of existing time series. Alternatively, this approach may be to check the stock situation with a minimalistic approach in years between a standard survey carried out biannually. Further work will be needed to quantify the impact on the uncertainty in the stock assessments by applying the different strategies. Freed effort from reduced surveys should be used to solve the crucial gap in understanding of pattern in year effects (or “catchability” trends in stock assessment lingo) and establish methodology for using this information in the stock assessment. These are examples. Further work is needed to find out what approach might give the best result in the long run.

Finally, minimalistic surveys may free capacity to do ecosystem monitoring that currently is limited by lack of survey effort. As ecosystem health and dynamics depends on factors like varying vertical distribution and changes in feeding, such studies could contribute to quantifying the survey condition and thereby support an approach that compensate for year to year changes not caused by variation in the stock (see above paragraph).

4. What impact can we expect on the simultaneous acoustic survey

There is all reason to believe that acoustic surveys are influenced by similar factors as the bottom trawl survey. Nevertheless, these factors do not necessarily change the estimates in the same direction (see e.g. (Godø and Wespestad, 1993, Karp and Walters, 1994, McQuinn et al., 1999). The current results again demonstrate that survey catches not necessarily sample representatively what is available with the profound effects this may have on the overall acoustic abundance estimates. The next step should therefore be to study minimalistic approaches in the simultaneous acoustic survey and study to what extent effects on the signals are parallel or opposite. Further, in the light of the outcome of these analyses we should reconsider the survey effort trying to find a balanced design taking into account the needs for precise estimates, quantification of **survey condition,** and representative sampling of species and sizes of the acoustically recorded biomass. We are convinced that this approach also will generate key information about the state of the ecosystem and may thus support EAF.

Our analysis demonstrate that excess of effort is no solution to improve the Barents Sea bottom trawl survey, which contrast the general belief that number of stations improves quality of survey estimates. Our study shows that when surveys suffer from year to year effects, more stations improve precision of the estimates but maintain the year effect. The unresolved issue is thus how to optimise the use of available effort. We can for example suggest annual reduced surveys along transects or biannual reduced surveys as approaches to free effort to do ecosystem research. In the long run this may improve quality of survey estimates as well as support other requirements set by EAF. We think that survey strategies in a wider ecosystem perspective are an underemphasized key issue that needs attention in the future. The wide range of acoustic technologies now available (Demer et al., 2009), is an important incitement for a renewed discussion of survey strategies including development of new approaches to meet the demand from EAF. Our findings actualise the following citation from social economics “It is better to be roughly right than precisely wrong” (John Maynard Keynes, 1883-1946).

References

AGLEN, A. 1994. Sources of Error in Acoustic Estimation of Fish Abundance. pp. 107-133. Ed. by A. FERNO and S. OLSEN. Fishing News Books, Blackwell Science Ltd, Oxford.

BROWMAN, H. I. and STERGIOU, K. I. 2004. Perspectives on Ecosystem-Based Approaches to the Management of Marine Resources. Marine Ecology-Progress Series, 274: 269-270.

DEMER, D. A., KLOSER, R. J., MACLENNAN, D. N. and ONA, E. 2009. An introduction to the proceedings and a synthesis of the 2008 ICES Symposium on the Ecosystem Approach with Fisheries Acoustics and Complementary Technologies (SEAFACTS). ICES Journal of Marine Science, 66: 961-965.

ENGÅS, A. and GODØ, O. R. 1989. Escape of fish under the fishing line of a Norwegian sampling trawl and its influence on survey results. Journal du Conseil International pour l'Exploration de la Mer, 45: 269-276.

GARCIA, S. M. 2003. The ecosystem approach to fisheries : issues, terminology, principles, institutional foundations, implementation and outlook, FAO, Rome.

GODØ, O. R. and SUNNANÅ, K. 1992. Size selection during trawl sampling of cod and haddock and its effect on abundance indices at age. pp. 293-310.

GODØ, O. R. and WESPESTAD, V. G. 1993. Monitoring changes in abundance of gadoids with varying availability to trawl and acoustic surveys. pp. 39-51.

GUNDERSON, D. R. 1993. Survey of Fisheries Resources, John Wiley & Sons, New York.

HYLEN, A., NAKKEN, O. and SUNNAN†, K. 1985. The use of acoustic and bottom trawl surveys in the assessement of North-East Arctic cod and haddock stocs. pp. 473-497.

ICES 2007. Report of the Arctic Fisheries Working Group (SFWG) 18-27 April 2007, Vigo, Spain. ICES CM 2007/ACFM:16: 1-651.

JAKOBSEN, T., KORSBREKKE, K., MEHL, S. and NAKKEN, O. 1997. Norwegian combined acoustic and bottom trawl surveys for demersal fish in the Barents Sea during winter. pp. 1-26.

KARP, W. A. and WALTERS, G. E. 1994. Survey assessment of semi-pelagic gadoids: the example of walleye pollock (*Theragra chalcogramma*) in the eastern Bering Sea. Marine Fisheries Review, 56: 8-22.

LØLAND, A., ALDRIN, M., ONA, E., HJELLVIK, V. and HOLST, J. C. 2007. Estimating and decomposing total uncertainty for survey-based abundance estimates of Norwegian spring-spawning herring. ICES Journal of Marine Science, 64: 1302-1312.

MANN, P. S. 2001. Introductory Statistics 4th Edition. Chapter 15.5, Wiley, London.

MCQUINN, I. H., SIMARD, Y., STROUD, T. W. F., BEAULIEU, J. L., MCCALLUM, B. R. and WALSH, S. 1999. An Adaptive Integrated Acoustic-Trawl Survey on Atlantic Cod. pp. 1-21.

NRC 1998. Improving Fish Stock Assessments, National Academic Press, Washington D.C.

SIMMONDS, J. and MACLENNAN, D. 2005. Fisheries acoustics : theory and practice, Oxford : Blackwell Science, London.

SMITH, S. J. 1988. Evaluating the Efficiency of the Delta-Distribution Mean Estimator. Biometrics, 44: 485-493.

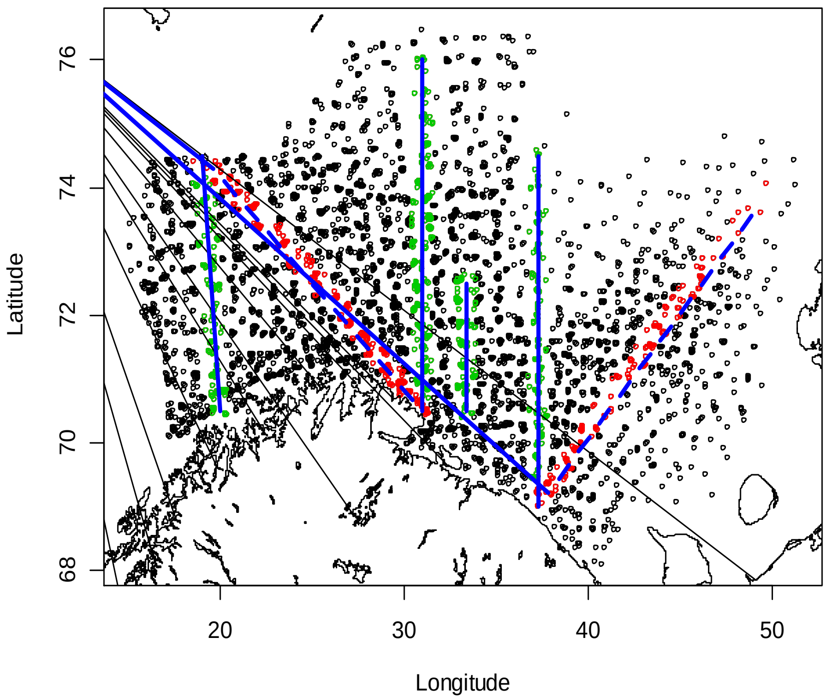

Figure 1. All trawl stations taken from 1989 through 2009 (black points), environmental transects (solid blue lines) and connecting transects (dashed blue lines). The data set Ω_{min} contains the green stations, Ω_{min+} contains the green and red stations, and Ω_{full} contains all stations.

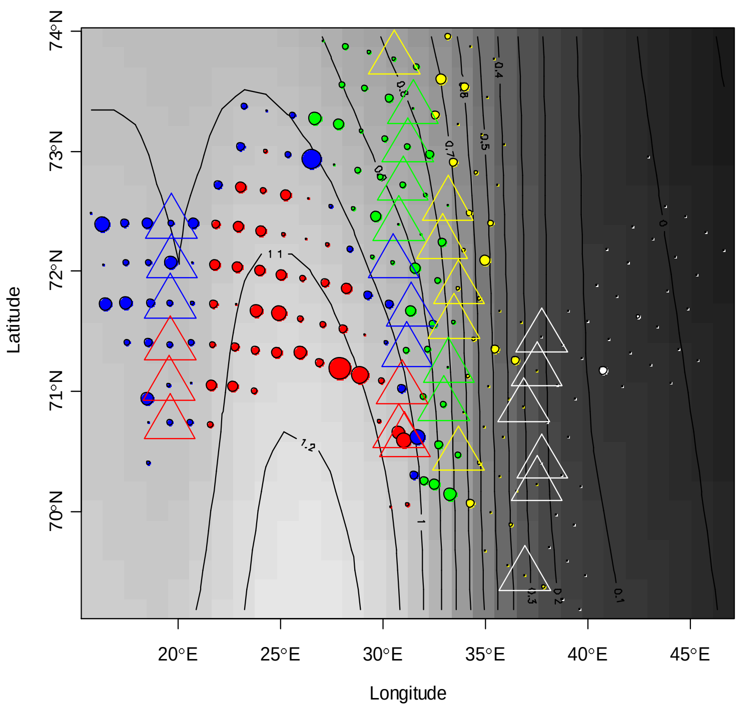

Figure 2. An example of stratification of trawl stations for bootstrapping using the 0, 20, 40, 60, 80, and 100 % quantiles of the predicted catches from a GAM fit to the catches of 7 year old cod in 1992. The grey shading indicates the level of the fitted GAM with light shading corresponding to high density. The circles represent trawl stations, and the diameter is proportional with log ($N_{1992,7,i}$+1). The colour inside the circles indicates which stratum the stations are allocated to for calculation of $I_{1992,7,full}$, and the color of the triangles indicates the stratification for calculation of $I_{1992,7,min}$. White, yellow, green, blue, and red represents increasing density.

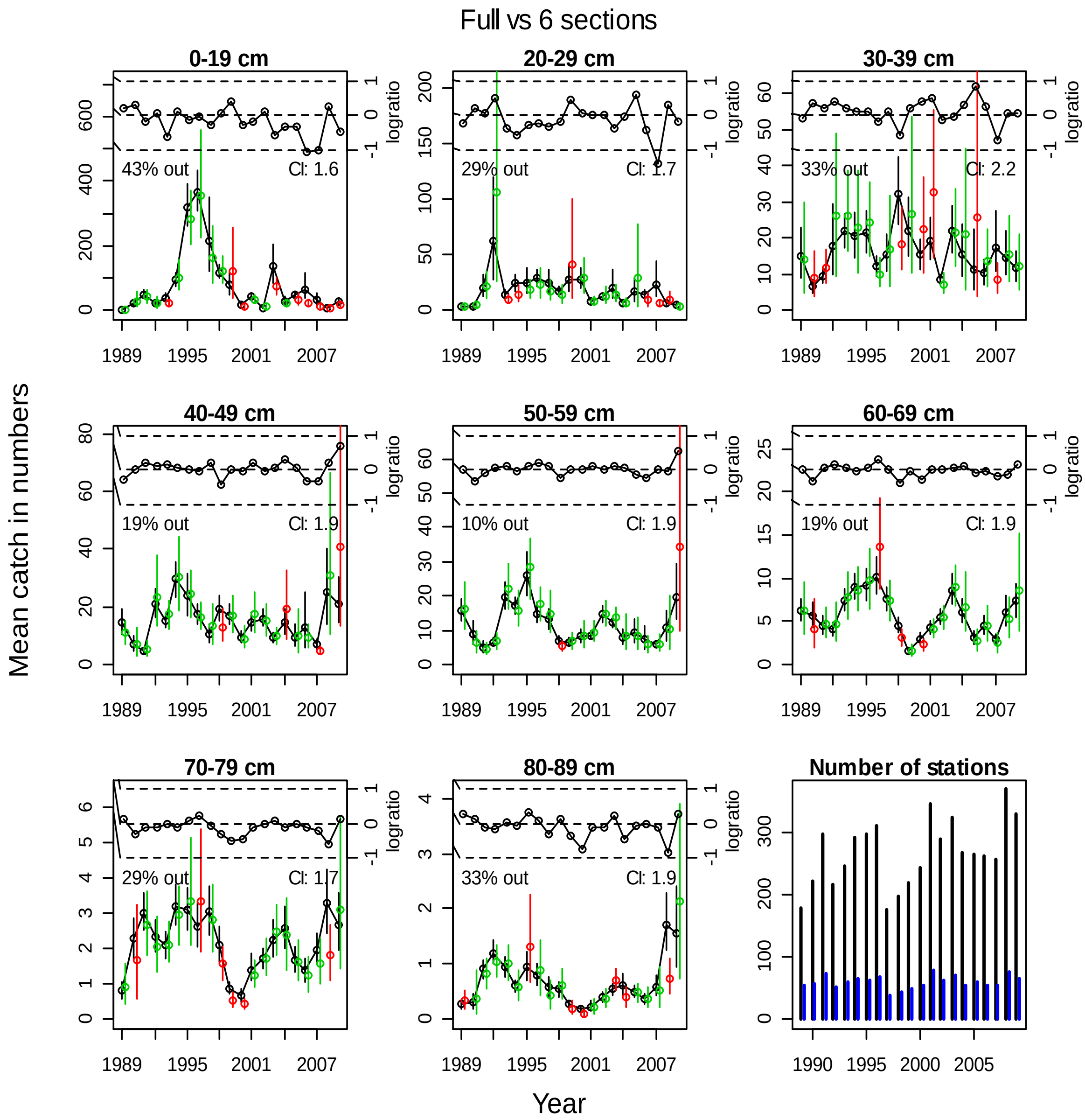

Figure 3. Comparison of length based indices and the associated 95 % confidence limits from full surveys (black) and reduced survey ($\Omega_{Y,min+,}$ 22.4 %) (green/red). Red indicates indices from reduced surveys outside the confidence limit of the full survey indices and percentage of red years are given by the numbers on the left hand side. Cl to the right gives the relationship between confidence limits of reduced and full surveys. At the top of each panel we show the relationship between indices (log ratios) of reduced and full surveys by year. Number of stations per year is shown in the lower right panel.

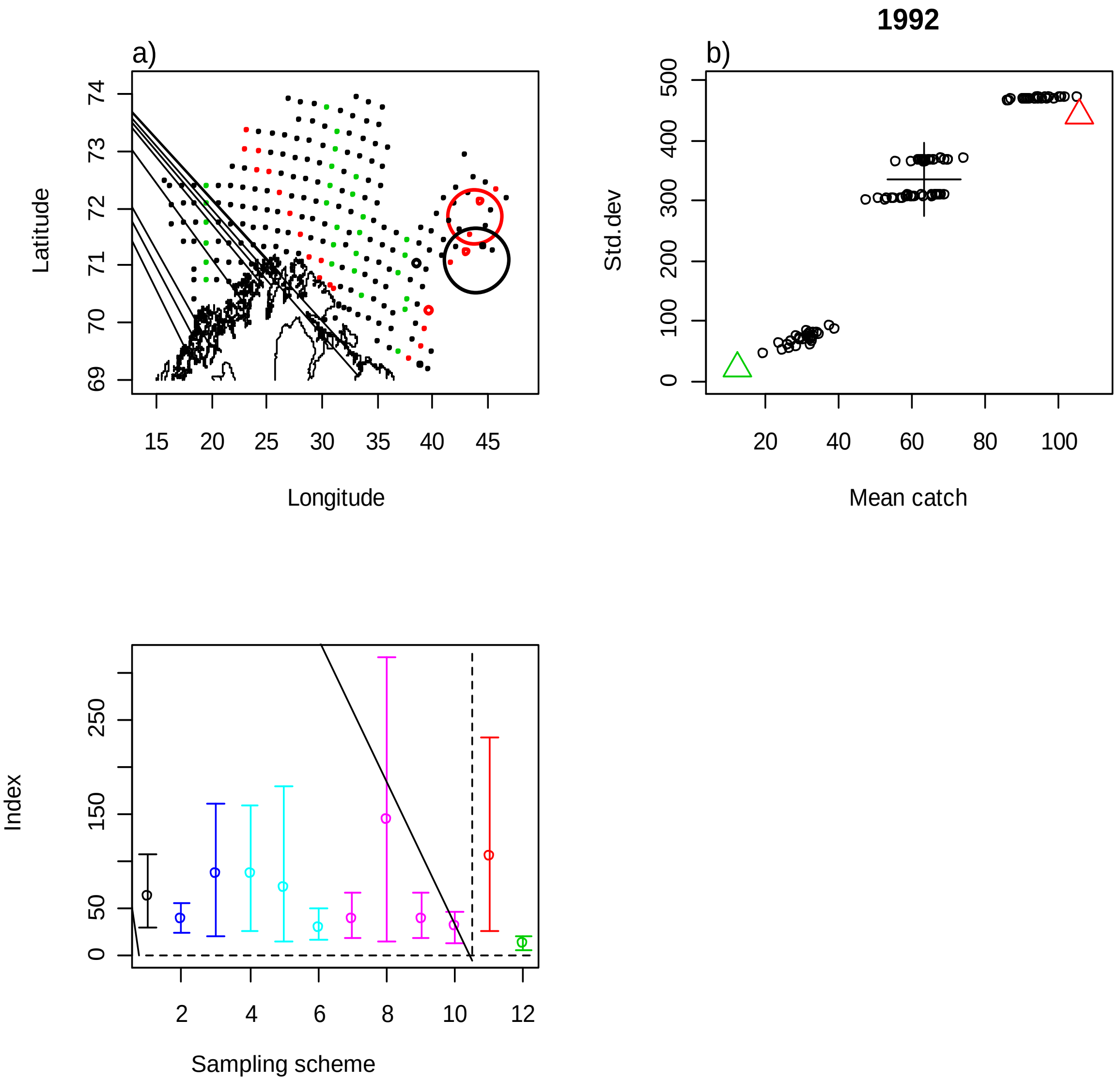

Figure 4. Example illustrating the impact of extreme catches on overall index and its uncertainty. a) Catches (proportional to bubble size) of 20-24 cm fish, black full survey, green the four sections, red the connecting sections. b) Std. deviation and mean catch for 100 selections of 50 % of catches compared to 22.4 % (red) and 13.5 % (green). C) Average index with standard deviation for various levels of sampling; black – full survey, green 13.5% and red 22.4% surveys. The other sampling schemes numbered from 2-10 show results from a gradually reduced sampling level.

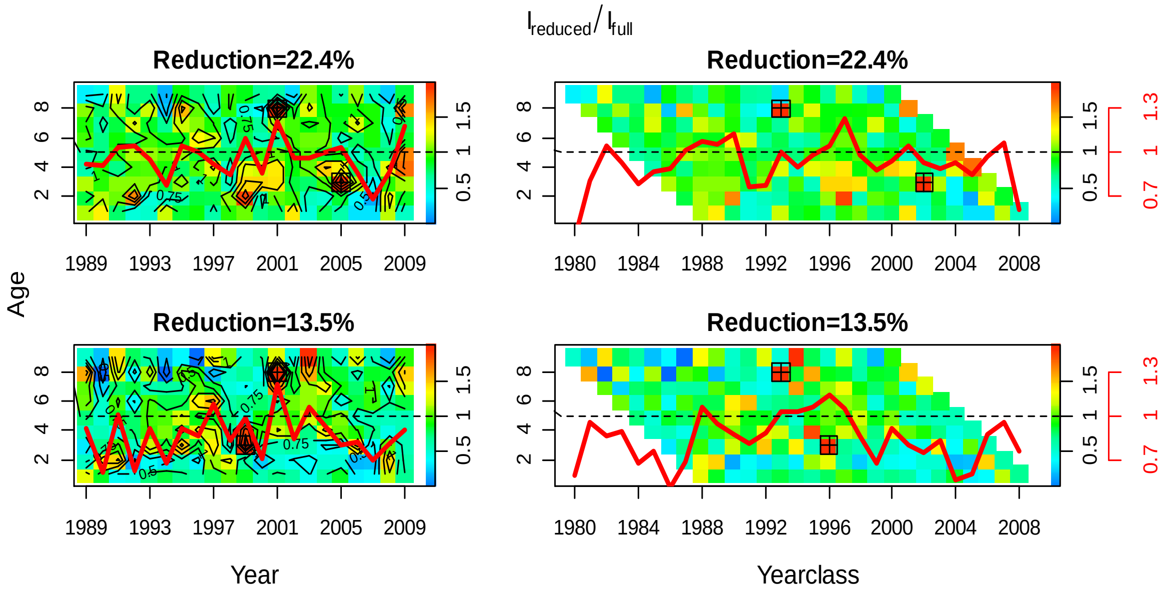

Figure 5. The relationship between index from reduced survey ($I_{reduced}$) and corresponding from full survey (I_{full}) by age and year (left panels), and age and year class (right panels). Colours indicate deviance from 1 and according to colour bar on right vertical axis. Red line shows average $I_{reduced}/I_{full}$ over all age groups by year and year class according to red axis on the right hand side.

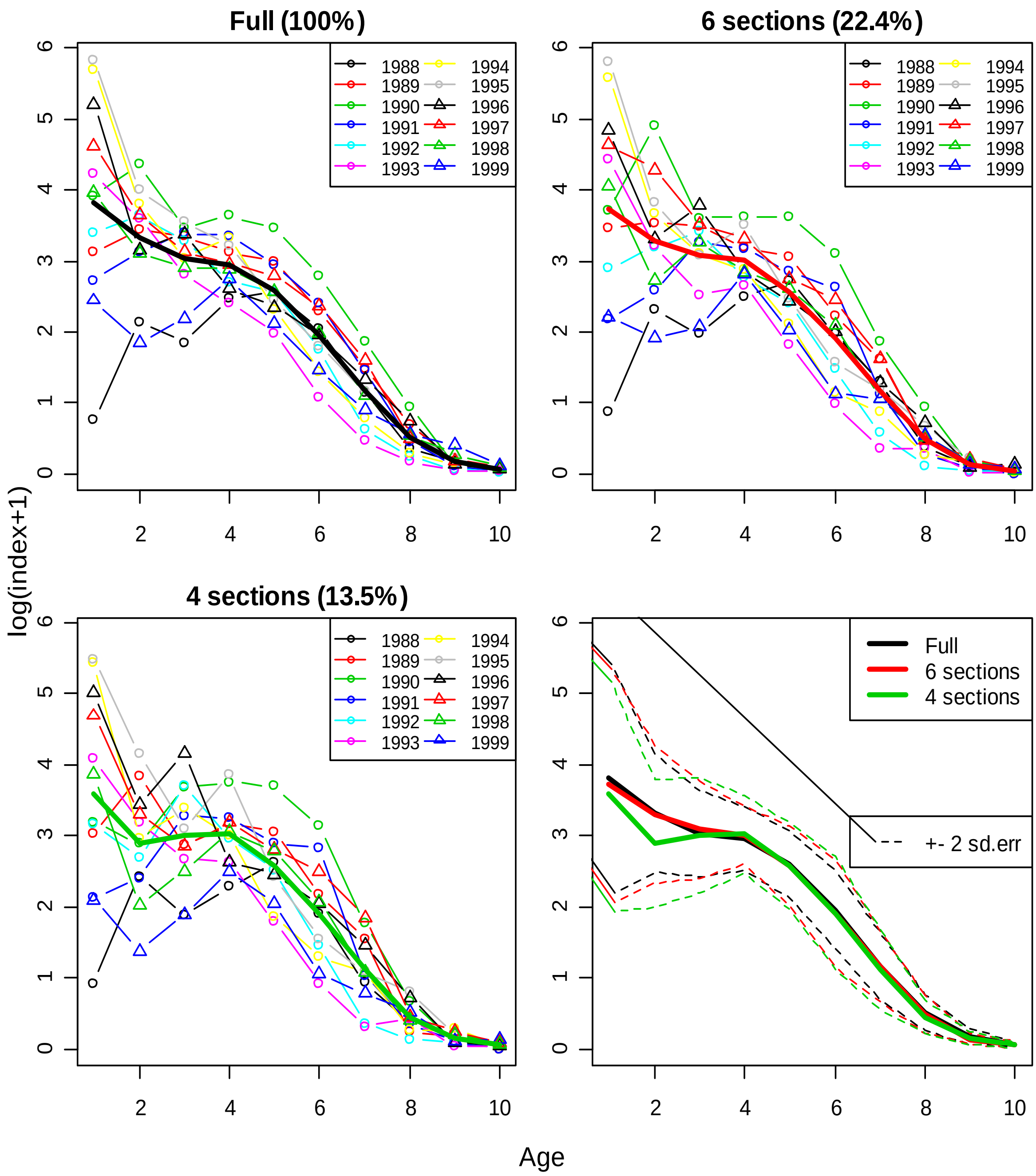

Figure 6. Index development by year class and age for full and reduced surveys. Lower right panel shows average index by age for full and reduced surveys with associated standard errors.

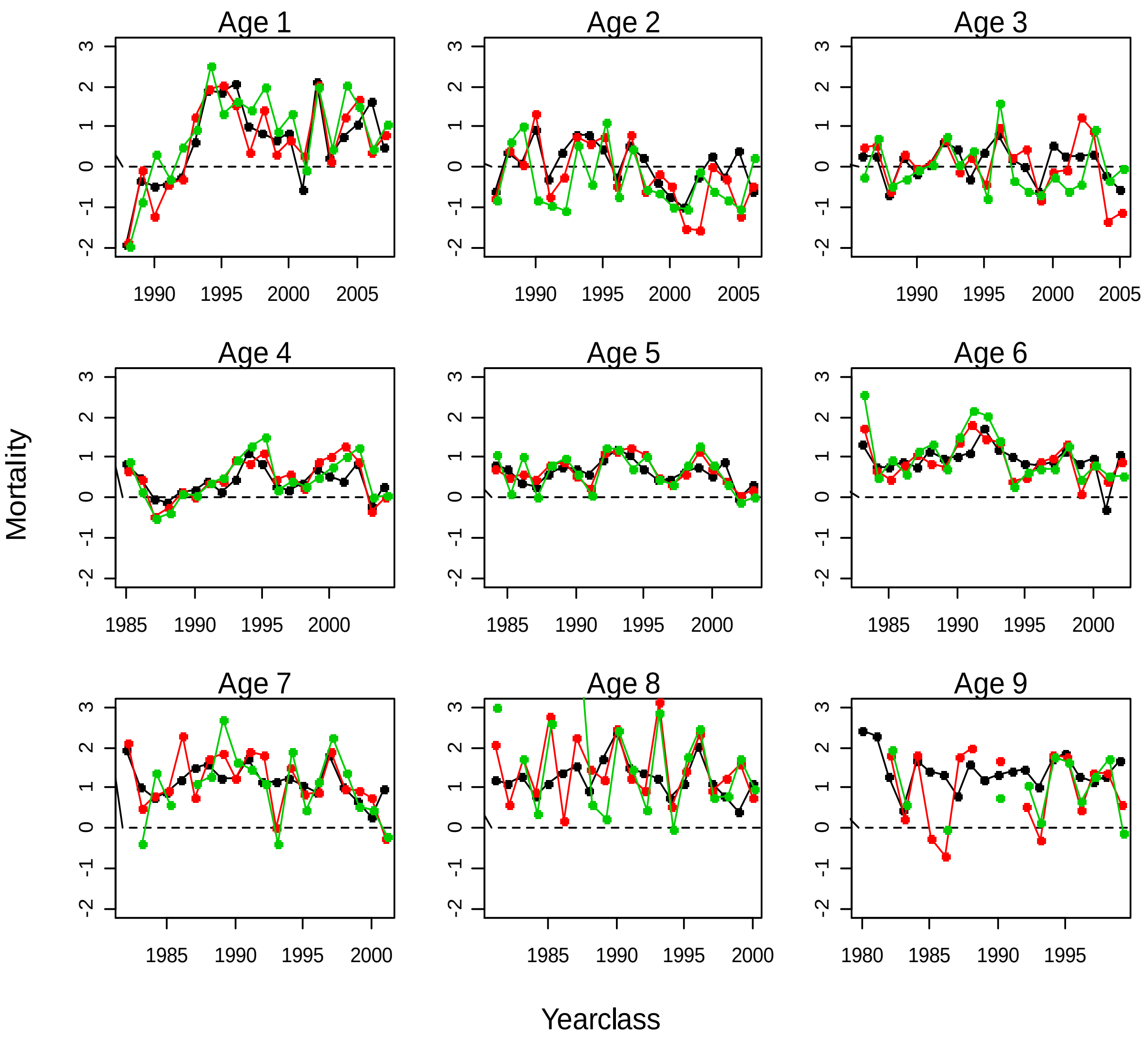

Figure 7. Estimated mortalities (Z) by age as estimated from the index of the age group one year and the index the following year at age a+1. More precise information on how indices are compared?

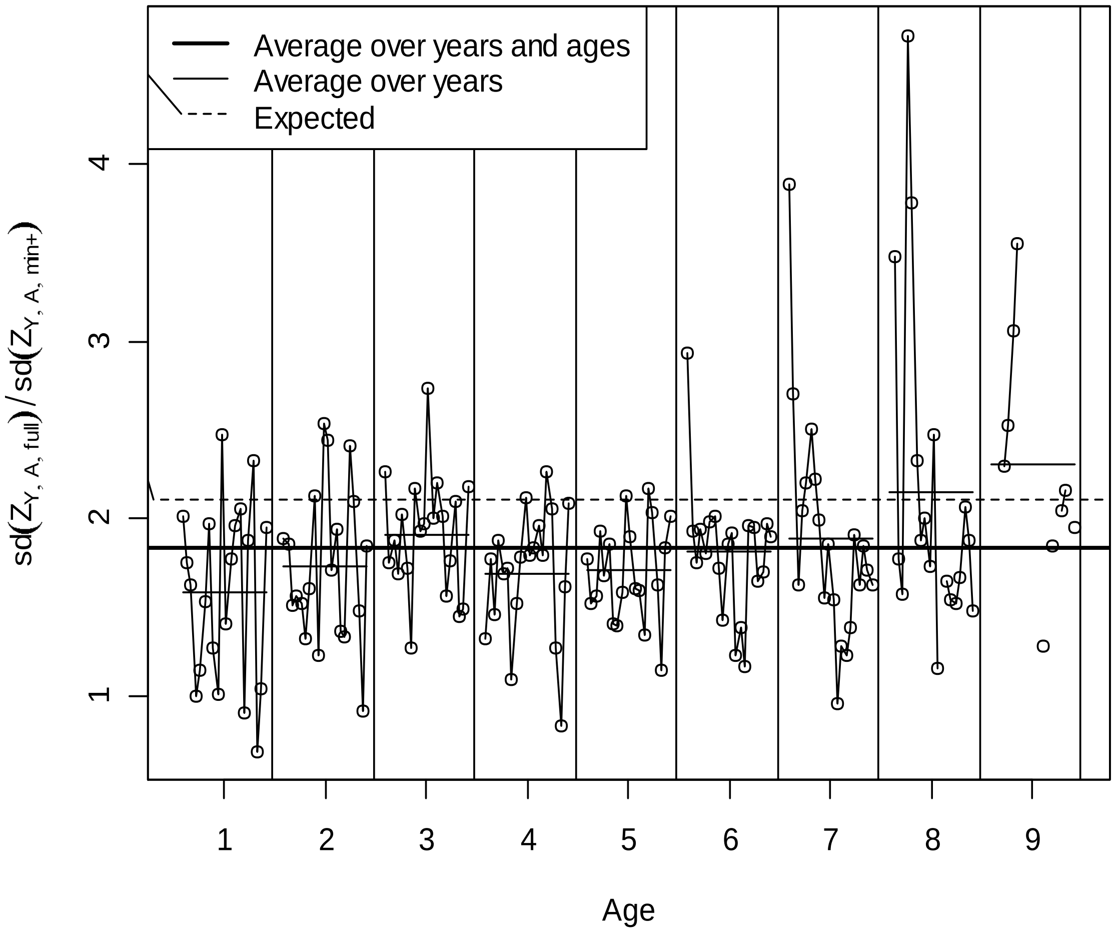

Figure 8. Bootstrapped standard deviations for one year mortality estimates for full and reduced ($\Omega_{Y,min+}$, 22.4 %)surveys. The ratio $sd(Z_{Y,A,full})/sd(Z_{Y,A,min+})$ is shown for 1 to 9 year old fish from 1989 through 2008. The standard deviations are calculated by stratified bootstrap with 1000 bootstrap replicates, using the GAM approach to do the stratification. The expected ratio, based on reduction in number of trawl stations, is $1/\sqrt{0.224}$.

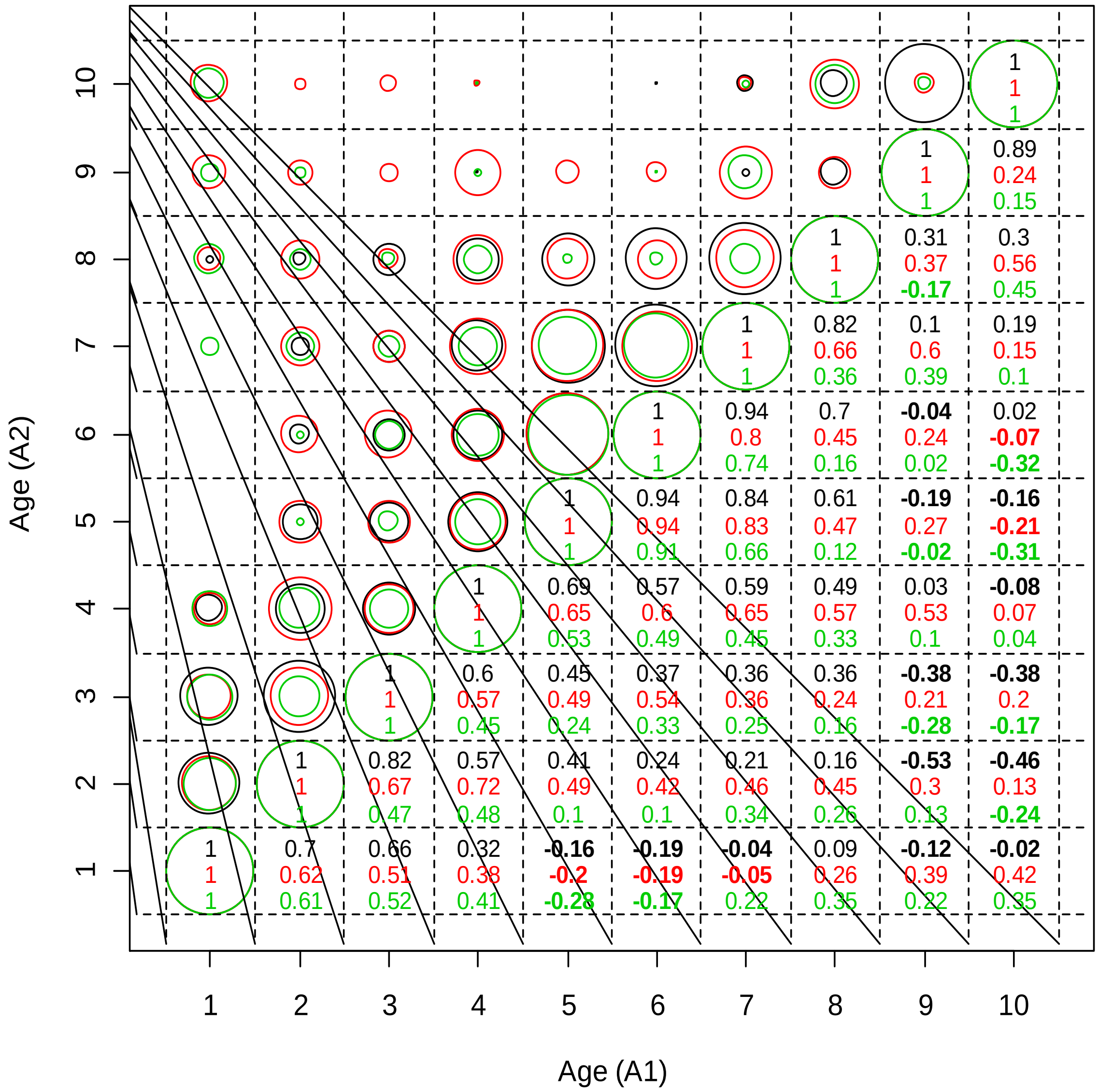

Figure 9. Rank correlation (cor(rank(I_{A1}), rank(I_{A2}) for the year classes 1988 to 1999) for studying consistency or lack thereof in measures of abundance of a year class over years. Correlations in the lower right hand side are visualised through corresponding bubble sizes in the upper left part of the figure. Full, 22.4 % and 13.5 % reduced surveys are given by black, red and green numbers and circles, respectively.