ICES WGISDAA REPORT 2012

SCICOM STEERING GROUP ON ECOSYSTEM SURVEYS SCIENCE AND TECHNOLOGY

ICES CM 2012/SSGESST:18

REF. SCICOM

Report of the Working Group on Improving use of Survey Data for Assessment and Advice (WGISDAA)

10–13 January 2012

ICES Headquarters, Copenhagen

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H. C. Andersens Boulevard 44–46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2012. Report of the Working Group on Improving use of Survey Data for Assessment and Advice (WGISDAA), 10-13 January 2012, ICES Headquarters, Copenhagen. ICES CM 2012/SSGESST:18. 39 pp.

Contents

Executive summary

The Working Group on Improving use of Survey Data for Assessment and Advice (WGISDAA) met for the first time at ICES Headquarters, Copenhagen, 10–13 January 2012 under the Co-Chairmanship of Colm Lordan (Ireland) and Stephen Smith (Canada). There were eight participants from the following ICES countries; US, Canada, Scotland, Ireland, Sweden and Norway. This was the first meeting of WGISDAA and the initial work focused on interpreting the ToRs, defining the scope of the group and a strategy to address the ToRs over the next three years. Fisheries surveys usually account for the highest cost in terms of data collection for assessment and advice. They also provide potentially the most useful and unbiased information on stock development, recruitment, exploitation, population characteristics and distribution. It is important therefore to maximize the utility of this information in the assessment and advisory process. Survey data are rather imprecise so it is essential to consider creative ways to improve survey precision whenever possible. There is undoubtedly a role for WGISDAA as an interface between survey practitioners, statisticians and stock assessment scientists. The group will strategically address the ToR on a multi-annual basis while liaising with the survey planning groups, assessment and benchmark WGs on priority case studies. There is already examples where two IBTS surveys (IRGFS and EVHOE) for haddock in VIIb-k were integrated and used as an improved combined index at WKROUND 2012. In 2013 WGISDAA will look at the potential issue of changing species distribution and impact on survey indices. In WGISDAA 2014 will address the theme of comparing survey estimates of (pseudo) exploitation (BREM, AIM, etc.) with assessment model estimates.

1 Introduction

The Working Group on Improving use of Survey Data for Assessment and Advice (WGISDAA) meet for the first time at ICES Headquarters, Copenhagen, 10–13 January 2012 under the Co-Chairmanship of Colm Lordan (Ireland) and Stephen Smith (Canada). This was the first meeting of WGISDAA and the initial work focused on interpreting the ToRs, defining the scope of the group and a strategy to address the ToRs over the next 3 years. The initial term of references set out for the group are set out below.

a) Develop a framework and methodology for the analysis of fishery-independent survey information for stock assessment and advisory purposes;

b) Explore and suggest refinements to current survey designs that will improve the quality of data used to support assessment and advisory processes;

c) Investigate methods of combining and or improving indices across multiple surveys and other ways of consolidating survey-derived data;

d) Develop methods for use of survey derived indices and other survey data products as a basis for scientific advice (this should include evaluation and, if appropriate, development and implementation of the method proposed in the EC's TAC setting policy statement (COM(2010)241 Annex 4);

e) Request priority case studies from assessment working groups to support the initial activities of the WG.

WGISDAA was asked to consider the "Multi-annual Management of SCICOM Expert Groups: Implementation" draft document. There was consensus that a move to multi-annual ToRs was both a good idea and would be appropriate for WGISDAA. Given that this was the first meeting of the group and very much a scoping meeting the group discussed in detail the individual ToRs and a category 2 resolution in proposed in Annex 3. In the interim sections 3–7 of this report the details of the discussions and strategies for each individual ToR are summarized. Section 8 gives abstracts of presentations given at WGISDAA 2012. Section 9 is a proposed roadmap outlining planned activities of the group for the next few years. Annex 5 details the recommendations and conclusions from WGISDAA.

2 Background

Sampling is the process of collecting subsets of a population in order to make inferences about the population as a whole (Thompson, 2002). In 2004 an ICES workshop on survey analysis and design (ICES, 2004) identified three tiers in the process of inferring population status from fisheries survey sampling. These are summarized as:

a) Estimation of fish density at a point. This results in a series of fish density or catch per unit of effort (cpue) data, and is undertaken at sea.

b) CPUE point estimates need to be correlated to the underlying (global) population. This requires information on how samples have been allocated/stratified (survey design) and also how they should be interpolated over the survey area (survey analysis). The result can be as simple as an arithmetic mean cpue index, or as complex as conditional geostatistical simulations of the population to derive biomass and uncertainty estimates. This usually happens back at the national laboratory.

c) Finally, incorporation of these estimates into stock assessment invariably involves analysis of a time-series of survey data along with additional information on mortality, recruitment, maturity etc. For commercially exploited stocks in NE Atlantic this generally occurs at ICES assessment working groups.

In the European context, ICES expert groups are tasked directly with coordinating a) and c) above. In contrast, how accurately, precisely and consistently survey samples are representing the underlying population requires multidisciplinary survey and statistical expertise. Consequently point b) can “fall between two stools” and is often left to model and survey design assumptions. A key *raison d'être* of this group is to evaluate and support in an active role, better information exchange at this intermediate and critical point in the process of fisheries assessment and resource management.

2.1 Survey data

Fishery-independent survey data has come to the fore in recent years as many commercially exploited stocks have declined both in terms of biomass and data quality. Where surveys have historically been designed to “tune” assessment models, survey data are now often drawn into the vacuum left when commercial data becomes either unreliable or unavailable (i.e. if fisheries are closed).

Survey datasets are generally small relative to commercial data, and therefore tend to have lower precision (higher variance) than commercial landings data. However, surveys can afford a level of control over the sampling unit (effort) and sampling efficiency (catchability) to the point where they are assumed to be standardized or equal to one (eqn. 1).

$$n = qfN$$

Eqn. 1. where the catch (n) is related to population (N) after correction for fishing effort (f) and the catchability of the trawl (q). Where f can be calculated precisely, and catchability fixed, catch should then reflect changes in the population in a relative sense. Knowing q precisely then affords converting catch to population biomass.

The ability to standardize survey effort and catchability is a fundamental assumption associated with survey datasets. Changes therefore in mean survey abundance should more accurately reflect changes in mean population abundance than do commercial datasets, where gear efficiency and fishing effort are continually adapted and commercially driven.

While survey standardization can greatly constrain catchability around 1, q will never be 100% efficient or fixed due to the range of factors that affect catchability (Blanchard and Boucher, 2001; Fryer, 1991; Fryer *et al.*, 2003; Godø, 1994; Jacobson, Brodziak and Rogers, 2001). For this reason, alternate data sources are often required to quantify the unsampled individuals before converting relative indices to absolute biomass. End of survey outputs therefore largely conclude with the reporting of a simple relative index of abundance. Thereafter, scaling to absolute biomass, required by fisheries managers, comes following combination with more extensive commercial landings statistics at the stock assessment stage.

The current trend towards survey-driven or survey-only assessments is therefore a significant shift in paradigm from many traditional stock assessments. Tipping the balance towards survey datasets should afford greater certainty in terms of sampling bias (constant catchability). Survey designs should also help avoid the bias of hyperstability in commercial landings data where fish aggregations are selectively targeted leading to artificially constant and high catch rates (Harley *et al.*, 2001, Hilborn and Walters, 1992). A trade-off with survey data of course can be the potential for higher variance associated with smaller survey datasets, potentially reducing the precision of the resulting assessment.

3 Analysis of fishery-independent survey information for stock assessment and advisory purposes

Term of Reference a) essentially tasks WGISDAA with outlining best practice in the analysis and presentation of survey results. The objectives therefore of this TOR are to look at salient aspects of:

i) Survey Accuracy and Precision – promote a series of simple pre-assessment exploratory checks of benefit to data users at both survey and assessment level in quality checking survey outputs.

ii) Covariate data – to consider auxiliary information beneficial in the interpretation of how survey estimates relate to the underlying population.

iii) Presentation of results – effective and efficient methods to distil and present survey data. That is, not simply to promote its use, but to communicate a contextual understanding the strengths and limitations of survey data.

A comprehensive review of survey design and data analysis was undertaken in a series of ICES workshops (ICES, 2004, ICES, 2005). Similarly, the more technical aspects and implications of trawl design and sampling were reviewed by the ICES Study Group on Survey Trawl Standardization (ICES, 2009). Here we will focus on recommendations arising from this, and other relevant work, and by use of case studies how they might best be implemented into general survey planning and reporting.

3.1 Survey accuracy and precision

3.1.1 Survey Accuracy

A basic assumption of single or multi-vessel research surveys is that they are conducted following standard procedures and consequently maintain standard catchability (q). Catchability itself will be determined by selection characteristics of the gear, the efficiency of capture and availability to capture. From an integrated analytical assessment perspective year affects or trends in survey catchability residuals often receive much discussion at assessment working groups. Often such discussions could be better informed with appropriate input from survey practitioners.

A major area of uncertainty in trawl surveys is the impact of changes in catchability due to non-random bias in trawl geometry and performance, on the estimate of abundance (Carrothers, 1981). Monitoring of trawl performance in the field is critical in order to uphold assumptions around standard catching efficiency (ICES, 2009). Survey scientists are continually investigating and trying to mitigate sources of trawl performance variability to improve the accuracy and precision in abundance estimates. Weinberg and Kotwicki (2008), divide the variables that might explain changes in trawl performance into three categories: a) vessel operation (trawl speed, scope ratio, towing direction and vessel); b) catch weight; c) environmental conditions (depth, wave height, wind direction and windspeed).

Several recent studies have focused on the effect of environmental condition on the bottom-trawl performance and efficiency (Weinberg and Kotwicki, 2008). Stewart *et al.* (2010) discussed the possibility of suspending sampling in case of adverse sea conditions because of strong impacts on the accuracy of indices and the risk of introducing bias; these aspects cannot be overlooked when interpreting survey catches. Therefore variance in trawl performance needs to be incorporated into estimates of abundance used as tuning in assessments. As highlighted by the IBTSWG (2010 and 2011), variation in trawl geometry has, as a minimum, the direct consequence of an unavoidable alteration of the sampling unit of effort - Swept-area.

Swept-area is an important parameter for abundance estimation from bottom-trawl survey (Godø and Engås. 1989) and quite precisely measurable. It is defined by the geometry of the gear (e.g. net spread) and the distance towed by the gear in contact with the bottom (Bertrand *et al.*, 2002). Therefore a simple correction can be applied to standardize the unit of effort which would smooth intra-vessel differences and increase the precision of the estimated species-specific combined index. This may well be appropriate for combined indices such as North Sea IBTS.

Survey estimates could be improved if the effect on trawl efficiency of variability of operational factors and environmental conditions were to be incorporated into the resulting outputs.

3.1.2 Survey Precision

Variability (precision) is random sampling error and due to only a fraction of the population being sampled (Thompson, 2002). The mean of the samples will however, vary around the mean of the population being sampled and is therefore unbiased. While minimizing bias is a core objective of survey design in order to approach the true population mean, precision is the confidence with which we can report survey estimates and therefore equally important.

With fixed resources, precision is largely a function of how patchy the species distribution is, although survey design can help to reduce the calculated variance by grouping areas of similar abundance together (Gunderson, 1993). Susceptibility to capture may change with size or spatial and temporal distribution, changing the precision in catch numbers at each age for example over the survey area, and/or time-series. Quantifying how and where precision is distributed throughout, and between surveys, is a key component to deciding what survey outputs are best suited to addressing specific questions. Indeed, precision itself can be used actively in post analysis to weight indices in the assessment process (Simmonds, 2003) and was also explored at the WG using DATRAS output and SURBA.

Given the importance of precision in survey data, and the variety of survey sampling designs in place, a number of case studies from differing survey designs will be used to contrast appropriate methods of calculation for each. The potential application of this information when either fed forward into the assessment or indeed as feedback into the ongoing survey design and management role should be investigated.

3.2 Covariate data

The ability to predict the spatial and/or temporal distribution of the underlying distribution of the population is a significant asset to survey design and analysis. Whether due to persistent, but heterogeneous habitat (JagieloHoffmannTagart et al., 2003, Zimmermann, 2003) or more cyclic environmental changes such as oceanography (EhrichAdlersteinGoetz et al., 1998, NeudeckerFischer and Damm, 1998, Poulard and Mahé, 2004, Smith and Page, 1996) or prey, covariate data can be incorporated in a number of ways for design based surveys.

First, and most commonly, delineating the survey area into zones (strata) of similar abundance (Smith and Gavaris, 1993, Gunderson, 1993, Ault *et al.*, 1999). Pre- and post-stratification of surveys to improve precision are particularly effective where target species and objectives are not too extensive (Cochran, 1977).

Further, where auxiliary data are known for the entire survey area, it can be incorporated into a predictive approach where unsampled sites are assigned a predicted catch based on the relationship between variable and catch (Valliant, 2000; Smith, 1990). A development of this is the model-assisted approach where catches are based on empirical likelihood (see Chen *et al.*, 2004).

Where covariates are well described and persistent over time there is significant scope to incorporate them into survey design.

3.3 Presentation of results

Key requirements of assessment/advisory scientists is some measure of uncertainty for survey inputs in assessments (whether they are biomass indices, length or age structured). Bootstrapped estimates of precision for several indices are currently available from the ICES survey database (DATRAS) and were explored and used during the meeting. It became clear that in many cases these are not currently being availed of by assessment groups, possibly because they are unaware of their existence or because the product provided is not as required. This again highlighted the need for an active and succinct development of a two-way communication exchange between survey practitioners and the target audience.

Many surveys, including those in the ICES area, provide a range of survey output to varying degrees of coordinated or standardized format. Invariably results are, at a minimum, presented as a cpue with some indication of trend over the time-series. As alluded to above, a degree of separation between the folks providing the field data and those tasked with converting relative cpue indices into actual biomass for management purposes can occur conceivably due to the range of skills required in the complete process. Where this does happen, pertinent information surrounding assumptions of the pivotal relationship between cpue and biomass can get lost in a report narrative or not be reported in such a way as to appear immediately applicable to the question or analysis in hand.

WGISDAA does not see its role to recommend onerous additional analysis and outputs for survey reports. By way of case studies, it is hoped to distil down the overarching concepts mentioned above and see what and how additional information can be most effectively exchanged to the benefit of survey and assessment experts, as well as ideally non-expert stake holders alike.

Without being prescriptive initial discussions around the type of items felt important for inclusion in a survey report were:

i) Clear survey objectives

ii) Survey design

iii) Sample site allocation and any short or long-term trends

iv) Index, survey units of effort and variance calculation methods

v) Spatial distribution of the stations and catches (spatial trends over time also important)

vi) Time-series trends in the catches with confidence intervals

vii) Short interpretation of results including anything of relevance to precision or bias in the results

Some of the above may be may not change regularly and of course be simply reference to a readily accessible background document/manual providing more comprehensive information.

4 Explore and suggest refinements to current survey designs that will improve the quality of data used to support and advisory processes

In this context, survey design refers to both survey procedures and how the locations of sample units are chosen. The survey procedures covers: how catch is processed, subsampling procedures and levels, how age–length key (ALK) samples are taken and combined, and sample unit design with respect to the definition of the sample unit (gear used and tow duration/length). In addition, the choice of estimators should be a function of how the sample units are chosen.

Changes to survey designs are inevitable. They are usually initiated by those responsible for the survey to deal with such things as the replacement of survey vessels, change in survey gear and changes in sampling coverage due to changes in funding/requirements, area closures or gear conflicts. Any proposed design changes should be fully discussed with survey users. In TOR a) we discuss the kinds of information that needs to be supplied to the assessment groups on these and other changes to help evaluate potential impacts on the interpretation of the survey indices. We also envisage changes to the design may result from assessment needs identified in TOR a), such as the provision of uncertainty estimates (relative or standard errors), change in the distribution of the fishery to areas outside that covered by the survey, changes in tow length or the collection of length frequency and age data because of effective sample size considerations.

Routine scrutiny of survey precision levels such as those in DATRAS should be undertaken by assessment WGs probably as part of the Benchmarking process. Once use of this information in assessments becomes more widespread, changes or improvements may be recommend to how the estimates of precision are derived. Additionally, recommendations to find ways of increasing the precision of survey estimates may be made by the assessment working groups. Simply increasing sample size may not be an option for economic reasons but there are other ways of increasing precision of survey estimates. These other ways can include adaptive sampling designs (Thompson and Seber, 1996; Smith and Lundy, 2006), sampling with partial replacement designs (Gregoire and Valentine 2007) and using environmental and other covariates (measured over whole survey area) that are related to distribution of one or more species in a predictive survey estimate (Smith 1990).

There are concerns about effective sample size result from evaluating the actual amount of independent information available from sampling of individual survey tows compared to sampling many tows (Chih 2010; Maunder, 2011 Pennington and Helle, 2011). Fish caught in the same tow are more similar in their growth and size characteristics than fish caught in different tows and sampling large numbers of animals from the same tow may not provide more information than sampling fewer fish over many tows. Similar arguments underlay the trade-off between having fewer long tows or many shorter duration tows.

There needs to be a system in place to facilitate feedback between the assessment and survey working groups to evaluate the impact of changes made to the surveys on the assessment models and of changes requested by the assessment groups on the surveys.

There are two aspects to this TOR that need to be addressed by this group. The first is to determine the kind of changes that could be made to surveys to improve support for stock assessment. Second, the group could recommend methods for evaluating the impact of these changes on surveys and evaluate changes to surveys on the stock assessments. Initially, we propose conducting a literature review to garner examples of past experiences and determine from these a set of best practices. Readily available examples include the extensive literature on comparative survey/calibration methods, changes to the Scottish survey design (ICES, 2012 – WGCSE report), WKSAD report on evaluating the impact of design change (ICES, 2005) and the EU study on the impact of survey changes on VPA models. Next, we will request examples of case studies from the assessment and survey working groups which can be used to apply ideas from the literature review or other ideas developed by or presented to our group to provide guidance and proof of concept.

5 Methods of combining and or improving indices across multiple surveys and other ways of consolidating survey-derived data

Abundance estimates of fish stocks often require multiple surveys or multiple survey methods to cover the full extent of the species distribution. This situation often arises in case of fish stocks that are distributed across multiple countries surveyed by multiple country specific surveys. Fish stocks that are found across multiple habitats may require different habitat specific survey methods. Different life stages of the particular fish species may also require different survey methods due to survey specific selectivity issues.

Abundance indices from multiple surveys are often treated as independent in stock assessment because of significant differences in methodology of surveys and lack of methods for combining survey abundance indices and lack of methods that assess the quality of the index (e.g. Western IBTS surveys for multiple species). However combined surveys indices may be desirable because they could provide a more reliable index of abundance for stock assessment purposes, spatial dynamics studies, ecological modelling and other studies that use abundance estimates or spatial density survey data. It is likely that indices combined form multiple surveys could perform better than each separate index. Good examples of very different approaches to combine multiple survey indices are given by Conn (2009) and Gerritsen (Annex 3).

Combining estimates from the different survey types can be difficult because of the specific differences of each survey. These include unknown availability of the stock to the survey as well as survey catchability and selectivity with respect to the available stock.

Regardless of the limitations of specific surveys, it should be possible to relate data from different surveys in one modelling framework to derive combined indices of stock abundance. However at this point it is an emerging field in the stock assessment and considerable amount of work is needed to develop methods for combining indices from independent surveys.

Some ideas on how to advance studies on this topic were discussed during WGISDAA 2012 meeting. These included:

1) Combining multiple surveys of the same type and for the same stock across different spatial areas. It has been noted that it is desirable to have a spatial overlap between surveys. This overlap could be designed as a part of the standard survey or it could be created for a period needed to assess relationships between surveys.

2) Combining different types of the surveys. The following examples were actively discussed during the meeting:

    a) Bottom trawl and acoustic surveys for semi-pelagic species could be combined by modelling the relationship between acoustic data and bottom-trawl data with respect to the acoustic dead zone and bottom-trawl performance parameters. This work is currently ongoing at the Alaska Fishery Science Center for walleye pollock in the eastern Bering Sea

    b) Bottom-trawl surveys that use different methods and are suspected to have different catchability and/or selectivity should be combined using models that account for known survey specific catchability and selectivity issues. If these issues are unknown it would be desirable to estimate at least length specific catchability ratios that would allow combining surveys into a relative index of abundance

c) For the case when an existing survey is modified by either changing survey methods or design, it is desirable to perform a survey inter-calibration. Survey inter-calibration methods could be similar to the methods used in combining data from different surveys.

d) For cases where different types of gear are used (e.g. pots, bottom trawls, video equipment, etc.) it should be possible to derive length specific catchability ratios between specific surveys by either performing experiments or using data from areas of overlap.

3) Fishery-independent survey data could be combined with the commercial catches by taking advantage of the VMS data and accounting for spatial and temporal differences between survey and commercial catches.

4) In the presence of persistent spatial differences in productivity, it may be important to consider if differential exploitation patterns for the fishery. In these cases it is desirable to account for differences in predictive regimes across surveys in the stock assessment models and associated reference points. An example of this problem was presented for sea scallops in Nova Scotia (see section 8 below).

6 Methods for using surveys directly for advice

The current form of ICES advice, for stocks for which there are accepted age-structured catch-at-age based assessments, generally follows agreed management plans, where these exist, or an MSY approach, where this seems to be appropriate. Many ICES stocks do not have such an assessment, and the current ICES template for the provision of management advice in these cases stipulates how such advice should be provided (and of what it should consist) for a number of different combinations of data availabilities and assessment methodologies (see Table 6.1). However, there are a number of stocks covered by ICES which (for example) may have excellent survey data providing absolute or relative indices of abundance or biomass, but which cannot be assessed in the traditional way because of (for example) a lack of aging data or commercial catch data. In these cases, the application of this Table could lead to inappropriate advice to reduce catch even if the survey data are indicating a rapidly increasing stock.

Table 6.1. Decision table used by ACOM (during 2011) to determine what management advice to provide in "data poor" situations.

		No overfishing	Overfishing	Unknown exploitation status
Decreasing stock trend	or	Reduce catch at rate of stock decrease. 1	Reduce catch at rate greater than the rate of stock decrease 2	Reduce catch at rate greater than the rate of stock decrease 3
Stable stock trend		Do not allow catches to increase 4	Reduce catch. 5	Reduce catch. 6
Increasing stock trend	or	Increase catch at rate of stock increase. 7	Do not allow catches to increase. 8	Do not allow catches to increase. 9
No trend information		Do not allow catches to increase 10	Reduce catch. 11	Reduce catch 12

More generally, the application of the Table leads to the classification of many stocks as being "data poor", with all the concomitant implications for management decisions, even when much good data exists that could be used as the basis for advice. There are stocks which do not fit into the catch-at-age assessment mould (e.g. where age estimation is problematic, discards are substantial, or time-series are short/partial). It is important to make the best use of all scientific knowledge and appropriate stock indicators (especially surveys) in an adaptive framework to produce pragmatic quantitative advice.

Bearing these points in mind, WGISDAA considered the issue of whether new paradigms and methodologies of assessments and advice could be developed that would be applicable to stocks with a variety of different data availabilities. Summaries of the data and methods categories considered, along with examples of stocks in ICES and elsewhere, are given in Table 6.2.

WGISDAA also considered the results of De Oliveira *et al.* (2010). That study examined in detail the so called category 6–9 rules proposed by the European Commission to make use of survey data as a basis for catch advice. WGISDAA agreed with the conclusion that the harvest control rule (HCR) proposed could be a way to stabilize the stock at its current level but did not inform managers on maximum sustainable yield. WGISDAA did have serious reservations about the applicability of the approach in the context of surveys designed as recruit indices. It critical to evaluate the appropriateness of the survey data one is using before testing these types of HCRs. Given that the De Oliveira *et al.* (2010) evaluation was carried out it is not necessary to carry forward the specific reference in the ToR of WGISDAA in future.

Table 6.2. Some management metrics and approaches that could be considered appropriate for different categories of data availability.

Available data	Management metrics	Management approach	ICES example(s)	Worldwide example(s)
Catch-at-age data and age-structured survey data	B and F from catch-at-age assessment	Agreed management plans based on Bref and Fmsy.		
Age-structured survey with swept-area abundance estimates	B and Z from survey-based assessment	HCR based on Bref and Zmsy.		Cod 2J3KL (eastern Canada)
Relative age-structured survey index	Relative B and Z from survey-based assessment	Either scale relative B using historical catch data to estimate Bref , or generate relative Bref. HCR based on Bref and Zmsy.	Irish Sea haddock	
Biomass survey	Absolute B estimate. Could also use BREM (or similar) to estimate exploitation rate.	Manage by trends in B relative to Bref (and Zmsy if exploitation rate estimated).	Northern Shelf anglerfish	Sea scallops in Canada
Abundance survey	Abundance	Proceed as for Nephrops (harvest rate appropriate for Fmsy).	Nephrops	
cpue index	Catch/effort	cpue based HCR.		Australian shark example (Little *et al.*)

6.1 Survey-based assessments and advice

If catch data for a stock are absent or considered unreliable, but a survey is available and is considered representative, then a wide range of approaches are available to estimate trends in stock abundance (and potentially exploitation) and hence management advice. It is important to define what we mean by representative. In this context it is that the surveys reflect the underlying population abundance in an unbiased way with reasonable precision.

If the survey concerned provides relative abundance estimates at age (an ICES example is haddock in Division VIIa), survey-based assessment methods (Mensil *et al.*, 2009) such as SURBA (in its various forms), CSA, TSA, delay difference and AIM (from the NOAA Toolbox) can be applied to generate time-series estimates of relative biomass and total mortality. These could then be compared with relative biomass reference points and an estimate of Z_{msy} to provide management advice in the usual manner. Alternatively, the survey may use swept-area or other density considerations to raise the relative abundance index to an absolute estimate of abundance. In this case, standard biomass reference points can be used in a familiar advice framework: such a system is currently used in the management of the 2J3KL cod off eastern Canada (using a modified version of SURBA known as SURBA+; see DFO 2011).

Some surveys provide swept-area estimates of total biomass (Northern Shelf monkfish) or total abundance (North Sea *Nephrops* using video transects, Barents Sea capelin using acoustics). In these cases, the state of the stock in relation to a biomass or abundance reference point can be readily determined, but the relevant exploitation rate is more difficult to ascertain. Here the approach taken for *Nephrops* provides a useful precedent, in which a harvest removal rate is estimated that should lead to fishing at F_{msy}. Alternatively, survey-based production models such as BREM could be used. If biomass growth is known then total mortality can be teased out otherwise BREM gives the net rate of change excluding recruitment.

Abundance indices (whether relative or absolute) from surveys rely strongly on the ability to estimate the catchability characteristics of that survey for the stock concerned. We can define catchability as the product of availability (whether a fish enters the survey gear) and selectivity (whether a fish is retained in the survey gear). Estimates of selectivity can be achieved through experiment, but availability is more difficult to determine and needs to consider factors such as the time of day, and what the survey target species is feeding on (for example, cod feeding on benthic species will be more available to a demersal survey trawl than cod feeding on midwater sprat). The spatial characteristics of the survey could also be used in the provision of advice, and it may yet prove possible to combine survey and commercial catch data on a common footing to generate spatio-temporal models of abundance. These approaches are likely to make use of raw survey data to a much greater extent than currently. Standard VPA assessments ignore much of the information available from surveys, and aside from the spatial aspect, should be replaced with statistical approaches making use of survey index variance estimates (see Annex 2 for an example of this with the SURBAR assessment method). Current assessments often assume all surveys are equally valid, whereas in reality this will not always be the case (e.g. surveys giving conflicting signals).

Fishing activity data from VMS and CCTV are becoming widely available, and could be utilized in spatial approaches in much the same way as survey data (with appropriate modifications for different catchability characteristics).

6.2 Conclusions

ICES has the ability to provide appropriate management advice on the basis of a wide range of metrics of stock abundance and exploitation, even in the absence of traditionally accepted catch-at-age based assessments. The reliability and utility of management advice based on alternative data sources will need to be analysed through such approaches as management strategy evaluation (MSE) before their use can be justified. One useful test could be to run an MSE for a stock with full data availability, and then repeat the MSE with data categories sequentially removed to understand at what point sustainable management actually becomes impossible. In any case, there is a clear need for the development of appropriate assessment and analysis methods to enable surveys to be used more appropriately in the provision of advice, alongside the need to consider improvements to the design of the surveys themselves.

7 Case studies from assessment working groups.

The ICES assessment, review and advisory framework has been evolving in the last number of years. The current approach is to have benchmark assessments every few years. In the preparation for these benchmarks WGISDAA could have a role in assisting preparations. In particular, for assessments where issues with surveys have been identified by the experts in advance of benchmarking or where the assessments are wholly dependent on the survey(s).

In the discussion of the individual ToRs elsewhere in this report WGISDAA sets out the main areas of interest to the group and a plan to address these over the next three years. Assessment Working Groups should review these and consider having case studies requiring some input from WGISDAA particularly in the preparation cycle for benchmark assessment.

At this meeting a case study was presented involving integrating two overlapping western IBTS surveys (Tor C) in preparation for the Haddock VIIb-k (Annex 3). In this case a pragmatic approach to combining the available indices was suggested weighted by the area covered in each. This approach seems valid given similarities between the surveys in the overlap areas. To estimate uncertainties in the combined index may take further consideration. Although the stratification used is common in the individual surveys the coverage and effort within strata is different. There is a potential for bias in the index by including the overlap area in each index as suggested. A spatial integration approach may be required to address these issues.

WKCOD (ICES, 2011) recommended that WGISDAA investigate some issues that had been identified with surveys for North Sea Cod. In particular there are suspected changes in catchability/availability of cod in the quarter 3 IBTS survey. This is an interesting case study that WGISDAA could follow up on but there were no participants at the 2012 meeting involved in investigating this problem to present this issue for discussion.

8 Abstracts of presentations given at WGISDAA 2012

Combining bottom trawl and acoustic data to model acoustic dead zone correction and bottom trawl efficiency parameters for semi-pelagic species.

Stan Kotwicki, Alex De Robertis, Jim Ianelli, André E. Punt, and John Horne

Abstract

Abundances of semi-pelagic fish are often estimated with acoustic-trawl or bottom-trawl surveys, both of which sample a limited area of vertical water column. Acoustic instruments are effective at sampling the water column, but have a near-bottom acoustic dead zone (ADZ), in which fish near the seabed cannot be detected. Bottom-trawl surveys cannot account for fish that are located above effective fishing height (EFH) of the trawl. We present a modelling method that combines acoustic and bottom-trawl abundance measurements and habitat data (e.g. grain size, temperature, depth, light levels) to derive ADZ correction and bottom-trawl efficiency parameters. Bottom trawl and acoustic measurements of walleye pollock (*Theragra chalcograma*) abundance and available habitat data from the eastern Bering Sea (EBS) were used to illustrate this method. Our results show that predictions of fish abundance in the ADZ can be improved by incorporating bottom habitat features such as depth and sediment particle size, as well as pelagic habitat features such as water temperature, light level, and current velocity. We also show that by modelling bottom-trawl catches as a function of acoustic measurements and environmentally dependent ADZ correction, we can obtain predictions for trawl efficiency parameters such as EFH, density-dependent trawl efficiency, and catchability ratio for trawl and acoustic data. The detectability of walleye pollock acoustic trawl surveys and catchability of bottom-trawl surveys are spatially and temporarily variable. Our results can be applied to the stock assessment for EBS walleye pollock in three ways: ADZ correction derived from the model can be used to assess detectability of acoustic trawl survey data in relation to habitat and environmental factors; environmental effects on bottom-trawl survey catchability can be assessed using estimated trawl efficiency parameters; abundance estimates can be corrected for habitat-specific or environmentally dependent catchability of bottom trawl and/or acoustic surveys. The modelling method used in this study can be easily extended to other semi-pelagic species where ADZ is of a concern or bottom-trawl survey efficiency parameters are unknown.

Improving area swept estimates from bottom trawl surveys.

Stan Kotwicki, Michael H. Martin, Edward A. Laman

Estimation of area swept is a key component for standardizing catch per unit of effort (cpue) data from fishery-independent bottom-trawl surveys and survey trawl gear experiments. Given technological advances and the proliferation of data streams from net mensuration equipment and global positioning system (GPS), techniques for estimating survey effort can be improved. Here we investigate new analytical techniques for improving the accuracy and precision of survey effort estimation (Kotwicki *et al.,* 2011). Sources of error and bias associated with two of the components used to compute area swept as a measure of fishing effort, distance fished by the trawl and net spread, are systematically examined and their influence quantified using both simulated and survey data. New analytical methods, a cubic spline smoothing algorithm to smooth GPS and net spread data, a haversine great circle algorithm to calculate distance between smoothed GPS track points, and a sequential outlier rejection algorithm to diminish the influence of noise on mean net spread estimates are shown to reduce or even eliminate the influence of biased observations on area swept estimators.

Sampling with partial replacement survey design

Stephen Smith

Survey time-series in support of stock assessment advice usually need to be running for a number of years (>15, depending on contrast in series) before they can be included in a population model. However, advice on stock status and the impact to the fishery is still required before there is enough survey data to implement a model. Survey designs are usually dictated by the objectives of a survey. According to sampling texts randomized/design based surveys are best when the objective is to estimate a mean or a total each year while fixed designs are usually optimal for estimating change from year to year. The former types of design are usually chosen when establishing a long-term time-series for models while the latter type would aid providing advice while establishing the time-series. In this presentation, I present a hybrid of the two types, sampling with partial replacement where there are two kinds of samples; samples chosen randomly each year from all possible sites the survey area and samples chosen randomly from sites sampled in the previous year. This design has many advantages when there is high correlation between observations from the current and previous year in the matched sample. The correlation can lead to increased precision for both the current mean and the difference of annual means without increasing the cost of the survey (i.e. greater precision with the same number of samples as a comparable simple random sample). Results from five years of a scallop survey off of Nova Scotia are presented demonstrating these advantages.

Spatial information from Surveys

Stephen Smith

The information content of surveys usually extends far beyond simply providing estimates of annual abundance, biomass and associated precision. The spatial distribution of species such as scallops are often determined by habitat type and knowledge of habitat type can not only lead to more accurate and precise surveys but also provide insights into the spatial aspects of the population dynamics of the species and the spatial impacts of the fishery. Scallop fishing area 29 off of the southwestern coast of Nova Scotia is an area where annual scallop surveys have been available since the fishery started in 2001, the area has been mapped using multibeam echosounders, geophysical characteristics have been determined and fishing has been monitored with logbooks and satellite monitoring systems (VMS). In addition a number of image surveys have been conducted where both the bottom type has been described and the species in the images have been identified. Species distribution modelling based on the multibeam and image data characterized the habitat suitability for scallops (Brown *et al.*, 2012). Fishing pressure metrics from the VMS data matched the habitat suitability patterns with the higher pressures occurring on the more suitable habitat areas. The scallop survey data also matched the habitat suitability patterns with the higher densities occurring on the more suitable habitat. The combination of this information showed that the higher suitability areas were more productive and tended to be fished down first until densities were similar over all suitability types. The implication of this pattern is that if management ignores the relationship between habitat and productivity, the higher productivity areas will generally be overfished. Impacts of this kind of spatial dynamics on establishing reference points were presented

Using UWTV Surveys in assessment and advice

Colm Lordan

Using underwater television surveys to monitor the abundance of *Nephrops* populations was initially pioneered in Scotland in early 1990s. Since then regular surveys have been developed for many of the main *Nephrops* fisheries around Britain and Ireland and the technique has also been used in Danish, Greek, Italian and Spanish waters. Historically either length cohort analysis (LCA) or tuned age-based assessments (XSA), where annual length distribution were sliced into pseudo-age groups, formed the basis of management advice for *Nephrops*. These methods performed relatively poorly due to the generally insensitivity nature of LCA to underlying stock and fishery dynamics and the lack of convergence in the VPAs. There were also concerns about representativeness of *Nephrops* LPUEs and tuning data as well as considerable uncertainty about accuracy of growth parameters. A more direct approach of using the UWTV surveys for applying harvest ratios (HRs) was proposed by Dobby and Bailey in 2006. Initially concerns about the accuracy of the UWTV surveys meant this approach was not widely accepted. WKNEPH 2007 discussed and documented the various uncertainties with UTV surveys and further developed the HR approach. Various work was then carried out to investigate and mitigate uncertainties in the UWTV survey methodologies (e.g. Campbell *et al*., 2009, WKNEPHBID 2008). WKNEPH 2009 debated the use of the surveys as either an absolute measure of abundance or a relative index relative. Ultimately concerns about the accuracy of landings statistics in particular lead to consensus that bias corrected survey abundance estimates could be used in the formulation of catch advice. Two modelling approaches were used to estimate sustainable stock specific Harvest Ratio reference points; SCA (a separable LCA model Bell) and Age Structured Simulation model (Dobby). Various harvest ratios are applied to bias corrected UWTV abundance, mean catch proportions retained and mean weight in the landings to give catch options in weight at different HRs. Stock specific F_{msy} proxies ($F_{0.1}$, $F_{35\%\ SPR}$, F_{max}) are chosen depending biological characteristics, level of knowledge and history of exploitation. While some concerns still remain, this approach has served to stabilize and standardize the production of catch advice for *Nephrops* stocks where UWTV surveys exist. One considerable advantage of this approach is that it can be applied to a single year's UWTV survey and doesn't necessarily require a long time-series to be useful.

A parsimonious estimator of abundance based on scientific surveys

Mike Pennington

VPA estimates of current stock size often differ significantly from the ultimate converged estimates. One reason for these revisions is that the relation between the commercial catch of cohorts still in the fishery and the actual population is usually unknown and probably changes from year to year.

An alternative to a VPA-type assessment of the current condition of a stock would be to base the assessment on, at least in theory, known relations. A survey ideally samples a stock in a consistent, standardized way, while converged VPA-type estimates based on accurate commercial catch data should provide fairly precise historical estimates of a cohort's size. Therefore it may be sensible to reverse the roles currently played by surveys and commercial catch data for some fish stocks. That is, instead of using survey data to tune a model based on catch data, use historical catch data to calibrate the survey.

As an illustration, converged ICES estimates of stock size (1981–1992) were used to calibrate the Norwegian-Russian winter trawl survey abundance indices for northeast Arctic cod. The subsequent survey-based abundance estimates, which are independent of any commercial catch-at-age data after the calibration period, were often more accurate than the annual ICES estimates (the VPA estimates are from ICES, 1995–2010). For example, the average absolute difference between the annual VPA estimates and the 2010 VPA estimates for 7+ northeast Arctic cod (which make up the bulk of the spawning stock) was 35.9 (s.d. = 27.9). While for the annual survey-based estimates, which are based on only two calibration parameters, the average absolute difference between the annual survey-based estimates and the 2010 VPA estimates was 18.5 (s.d. = 11.5). In particular, the survey-based estimates were closer to the 2010 VPA estimates in 11 out of 15 years.

Currently, based on the 2011 survey-based estimate and other information, it appears likely that the 2011 VPA significantly underestimated the abundance of 7+ cod (ICES, 2011). For more details on the survey-based estimators for northeast Arctic cod see: Pennington and Strømme, 1998; Korsbrekke *et al.*, 2001; Pennington *et al.*, 2011.

9 Time frame and roadmap for future activities of WGISDAA

The Chairs of WGISDAA will contact the assessment WGCHAIRS to solicit priority cases studies for the next meeting of the group in 2013. A dialog with WGCHAIRS will be needed to define priority requirements. In addition the 2013 meeting will also try to address the following theme;

Developing indicators for change in distribution of target species in a survey (and/or fishery) could involve looking at overlaying VMS on survey distributions as a first step. It may also involve looking at cases where changing underlying population distributions are causing problems with survey indices.

The subsequent meeting in 2014 will address the theme of comparing survey estimates of (pseudo-)exploitation (BREM, AIM, etc.) with assessment model estimates.

10 References

Ault, J. S., Diaz, G. A., Smith, S. G., Luo, J., and Serafy, J. E. 1999. An efficient sampling survey design to estimate pink shrimp population abundance in Biscayne Bay, Florida. North American Journal of Fisheries Management, 19: 696–712.

Blanchard, F., and Boucher, J. 2001. Temporal variability of total biomass in harvested communities of demersal fishes. Fisheries Research (Amsterdam), 49: 283–293.

Bertrand, J. A., Leonori, I., Dremière, P-Y., and Cosimi, G. 2002. Depth trajectory and performance of a trawl used for an international bottom trawl survey in the Mediterranean. Scientia Marina, 66(Suppl. 2): 169–182.

Brown, C. J., Sameoto, J. A., and Smith, S. J. 2012. Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. Journal of Sea Research, 72: 1–13.

Carrothers, P. J. G. 1981. Catch variability due to variations in groundfish otter trawl behavior and possibilities to reduce it through instrumented fishing gear studies and improved fishing procedures. *In* Bottom Trawl Surveys, pp. 247–257. Ed. by W. G. Doubleday, and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 58. 273 pp.

Chih, C.-P. 2010. Incorporating effective sample sizes into sampling designs for reef fish. Fisheries Research, 105: 102–110.

Chen, J., Thompson, M. E., and Wu, C. 2004. Estimation of fish abundance indices based on scientific research trawl surveys. Biometrics, 60: 116–123.

Cochran, W. G. 1977. Sampling techniques (third edition). John Wiley and Sons, New York, 428 pp.

Conn, P. B. 2010. Hierarchical analysis of multiple noisy abundance indices. Canadian Journal of Fisheries and Aquatic Sciences, 67: 108–120.

De Oliveira, J., Darby, C., Earl, T., and O'Brien, C. 2010. Technical Background Evaluation of Annex IV Rules. Cefas, October 2010. 28 pp.

Ehrich, S., Adlerstein, S., Goetz, S., Mergardt, N., and Temming, A. 1998 Variation of meso scale fish distribution in the North Sea. ICES CM, J:25: 7

Fryer, R. J. 1991. A model of between-haul variation in selectivity. ICES Journal of Marine Science, 48: 281–290.

Fryer, R. J., Zuur, A. F., and Graham, N. 2003 Using mixed models to combine smooth size-selection and catch-comparison curves over hauls. Canadian Journal of Fisheries and Aquatic Sciences, 60: 448–459.

Godø, O. R., and Engås, A. 1989. Swept area variation with depth and its influence on abundance indices of groundfish from trawl surveys. Journal of Northwest Atlantic Fishery Science, 9: 133–139.

Godø, O. R. 1994. Factors affecting the reliability of groundfish abundance estimates from bottom trawl surveys. In: Marine fish behaviour in capture and abundance estimation. (Eds A. Fernö and S. Olsen), pp. 166–199. Fishing News Books.

Gregoire, T. G., and Valentine, H. T. 2007. Sampling strategies for natural resources and the environment. Chapman & Hall/CRC Applied Environmental Statistics, 3: pp. 496.

Gunderson, D. R. 1993. Surveys of fisheries resources, John Wiley & Sons, Inc., New York, NY.

Harley, S. J., Myers, R. A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences, 58: 1760–1772.

Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment; choice, dynamics and uncertainty. Chapman & Hall, New York.

ICES. 1995–2011. Reports of the Arctic Fisheries Working Group (AFWG).

ICES. 2004. Report of the Workshop on Survey Design and Data Analysis (WKSAD). In: WKSAD. (Eds. FRS Marine Laboratory Aberdeen, Aberdeen, UK.

ICES. 2005. Report of the Workshop on Survey Design and Data Analysis (WKSAD). In: WKSAD. (Eds. Ifremer, Sète, France.

ICES. 2009 Report of the Study Group on Survey Trawl Standardisation (SGSTS). In: ICES Council Meeting documents [ICES Council Meeting Documents. Copenhagen]. (Eds, p. [np], Vol. ICES CM 2009/FTC:09.

ICES. 2010. Report of the International Bottom Trawl Survey Working Group (IBTSWG). ICES CM 2010/SSGESST:06. 243 pp.

ICES. 2011. Report of the International Bottom Trawl Survey Working Group (IBTSWG). ICES CM 2011/SSGESST:06. 243 pp.

ICES. 2011. Report of the Workshop on the Analysis of the Benchmark of Cod in Subarea IV (North Sea), Division VIId (Eastern Channel) and Division IIIa (Skagerrak; WKCOD 2011), 7–9 February 2011, Copenhagen, Denmark. ICES CM 2011/ACOM:51. 94 pp.

Jacobson, L. D., Brodziak, J., and Rogers, J. 2001. Depth distributions and time-varying bottom trawl selectivities for Dover sole (*Microstomus pacificus*), sablefish (*Anoplopoma fimbria*), and thornyheads (*Sebastolobus alascanus* and *S. altivelis*) in a commercial fishery. Fishery Bulletin, 99: 309–327.

Jagielo, T., Hoffmann, A., Tagart, J., and Zimmermann, M. 2003. Demersal groundfish densities in trawlable and untrawlable habitats off Washington: implications for the estimation of habitat bias in trawl surveys. Fishery Bulletin, 101: 545–565.

Korsbrekke, K., S. Mehl, O. Nakken and M. Pennington. 2001. A survey-based assessment of the Northeast Arctic cod stock. ICES Journal of Marine Science, 58: 763–769.

Kotwicki, S., Martin, M. H., and Laman, E. A. 2011. Improving area swept estimates from bottom trawl surveys. Fisheries Research, 110: 198–206.

Maunder, M. N. 2011. Review and evaluation of likelihood functions for composition data in stock-assessment models: Estimating the effective sample size. Fisheries Research, 109: 311–319.

Maynou, F., and Sardà, F. 2001. Influence of environmental factors on commercial trawl catches of *Nephrops norvegicus* (L.). ICES Journal of Marine Science, 58: 1318–1325.

Mesnil, B. Cotter, J., Fryer, R. J., Needle, C. L., and Trenkel, V. M. 2009. A review of fishery-independent assessment models, and initial evaluation based on simulated data. Aquatic Living Resources, 22: 207–216.

Neudecker, T., Fischer, J., and Damm, U. 1998. Influence of tidal currents on fishing performance in the Wadden Sea, ICES, Copenhagen (Denmark).

Pelletier, D. 1998. Intercalibration of research survey vessels in fisheries: a review and an application. Canadian Journal of Fisheries and Aquatic Sciences, 55, 2672–2690.

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

Pennington, M., and Helle, K. 2011. Evaluation of the design and efficiency of the Norwegian self-sampling purse-seine reference fleet. ICES Journal of Marine Science, 68: 1764–1768.

Pennington, M., Shevelev, M. S., Vølstad, J. H., and Nakken, O. 2011. Bottom trawl surveys: In: T. Jakobsen and V. K. Ozhigin (eds.), The Barents Sea: Ecosystem, Resources, Management. pp 570–583. Tapir Academic Press, Trondheim.

Poulard, J.-C., and Mahé, J.-C. 2004. Structure and spatial distribution of fish assemblages in the Celtic Sea., p. 14pp. Ifremer.

Simmonds, E. J. 2003. Weighting of acoustic- and trawl-survey indices for the assessment of North Sea herring. ICES Journal of Marine Science, 60, 463–471.

Smith, S. J. 1990. Use of statistical models for the estimation of abundance from groundfish trawl survey data. Canadian Journal of Fisheries and Aquatic Sciences, 47: 894–903.

Smith, S. J., and Gavaris, S. 1993. Improving the precision of abundance estimates of Eastern Scotian Shelf Atlantic cod from bottom trawl surveys. North American Journal of Fisheries Management, 13: 35–47.

Smith, S. J., and Page, F. H. 1996. Associations between Atlantic cod (*Gadus morhua*) and hydrographic variables: implications for the management of the 4VsW cod stock. ICES Journal of Marine Science, 53: 597–614.

Smith S. J., and Lundy, M. J. 2006. Improving the precision of design-based scallop drag surveys using adaptive allocation methods. Canadian Journal of Fisheries and Aquatic Sciences, 63: 1639–1646.

Stewart, I. J., Keller, A. A., Fruh, E. L., Simon, V. H., and Horness, B. H. 2010. Throwing in the towel: when do adverse conditions dictate a weather day during a bottom trawl survey? Fisheries Research, 102: 130–140.

Thompson, S. K. 2002. Sampling. John Wiley & Sons, Inc., New York.

Thompson, S.K., and Seber, G. A. F. 1996. Adaptive Sampling. New York: Wiley.

Valliant, R., Dorfman, A. H., and Royall, R. M. 2000. Finite population sampling and inference: A prediction approach. John Wiley & Sons.

Weinberg, K. L., and Kotwicki, S. 2008. Factors influencing net spread and sea floor contact of a survey bottom trawl. Fisheries Research, 93: 265–279.

Wieland, K., Olesen, H. J., Pedersen, E. M. F., and Beyer, J. E. 2011. Potential bias in estimates of abundance and distribution of North Sea cod (*Gadus morhua*) due to strong winds prevailing prior or during a survey. Fisheries Research, 110: 325–330.

Zimmermann, M. 2003. Calculation of untrawlable areas within the boundaries of a bottom trawl survey. Canadian Journal of Fisheries and Aquatic Sciences, 60:657–669.

Annex 1: List of participants

Name	Address	Telephone/Fax	E-mail
Colm Lordan **Chair**	Marine Institute Rinville Oranmore Co. Galway Ireland	+353 91 387 387 (or *200) / mobile +35 876 995 708 +353 91 387201	colm.lordan@marine.ie
Stephen Smith **Chair**	Fisheries and Oceans Canada Bedford Institute of Oceanography 1 Challenger Drive PO Box 1006 Dartmouth, NS. B2Y 4A2 Canada	+1-902-426-3317	Stephen.smith@dfo-mpo.gc.ca
Knut Korsbrekke	Institute of Marine Research PO Box 1870 Nordnes 5817 Bergen Norway	+47 55 23 86 38 +47 55 23 85 31	knut.korsbrekke@imr.no
Stan Kotwicki	National Marine Fisheries Services Alaska Fisheries Science Center 7600 Sand Point Way N.E. Seattle WA 98115 United States	+1 206-526-6614	stan.kotwicki@noaa.gov
Coby Needle	Marine Scotland Science Marine Laboratory PO Box 101 375 Victoria Road Aberdeen AB11 9DB Scotland	+44 1224 295456 +44 1224 295511	needlec@marlab.ac.uk coby.needle@scotland.gsi.gov.uk
Michael Pennington	Institute of Marine Research Nordnes PO Box 1870 5817 Bergen Norway	+47 55238601	michael.pennington@imr.no
Francesca Vitale	Swedish University of Agricultural Sciences Institute of Marine Research Turistgatan 5 PO Box 4 SE-453 21 Lysekil Sweden	+46 523 18774 +46 523 13977	francesca.vitale@slu.se

Name	Address	Telephone/Fax	E-mail
David Stokes	Marine Institute Rinville Oranmore Co. Galway Ireland	+353 (0)91 387200 +353 (0)91 387201	david.stokes@marine.ie

Annex 2: Estimates of survey variance and effects on survey-based assessments

In our discussion we discovered that estimates of CV were available and here is a preliminary look at the consequences for survey and assessment working groups.

The IBTS Q1 survey indices for North Sea haddock (ages 1–6, years 1983–2011) were extracted from DATRAS, along with bootstrap resamples for each age and year provided by ICES in DATRAS. For most years there were 500 resamples in this dataset, although there were 510 resamples for the years 1983–1989 and 1500 for 2007. Figure A.2.1 gives the estimated CV (standard error divided by the mean) for each year and age, along with a variance metric provided by ICES on the DATRAS website (range divided by median). The Figure shows that the CVs of this survey are generally low, around 10% – 20%, although there has been an increase in CV for all ages towards the end of the time-series. The CVs on older ages towards the end of the time-series exceeds the 30% level and may not be contributing much useful information to the assessment. The increasing CVs are a source of concern and possible causes should be investigated by IBTSWG.

Figure A.2.2 shows the summaries from a standard SURBAR run for North Sea haddock, with no downweighting of index data during sum-of-squares minimization. Figure A.2.3 shows the effect of applying inverse-variance down-weighting, using the inverse of the estimated CVs as weights, while Figure A.2.4 applies additional downweighting to age 1 indices only (they are effectively weighted out of the estimation). Finally, Figure A.2.5 compares the results of these three runs.

The application of inverse-variance downweighting makes very little difference to the stock estimates from SURBAR in this case, although the variance of the recruitment estimates from the third run is inflated. Aside from some of the CVs on the older ages towards the end of the time-series, the estimated variances on the survey indices are quite stable for this particular case. The effect of downweighting in a case with more extreme and changeable CVs still needs to be explored.

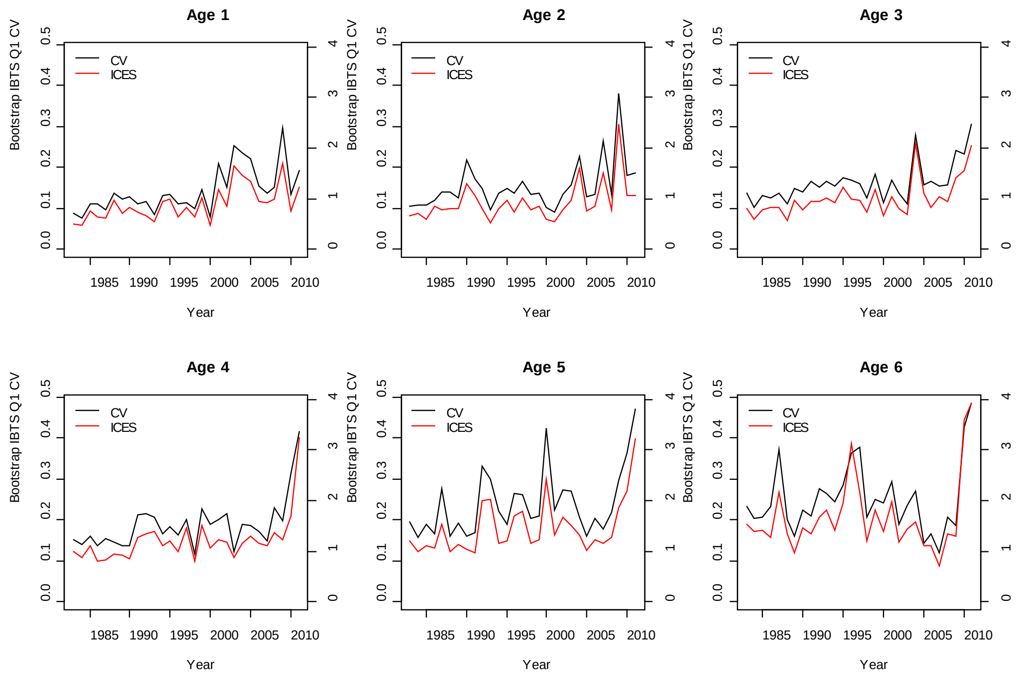

Figure A.2.1. Variance estimates for the IBTS Q1 survey indices for North Sea haddock.

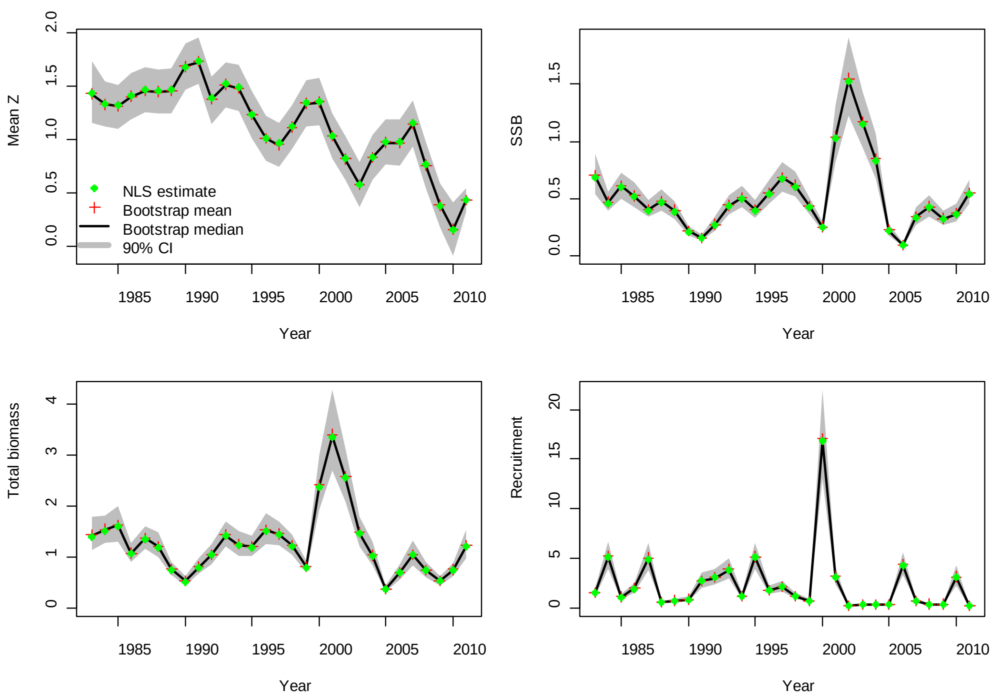

Figure A.2.2. SURBAR summaries for North Sea haddock, using IBTS Q1 with no down-weighting.

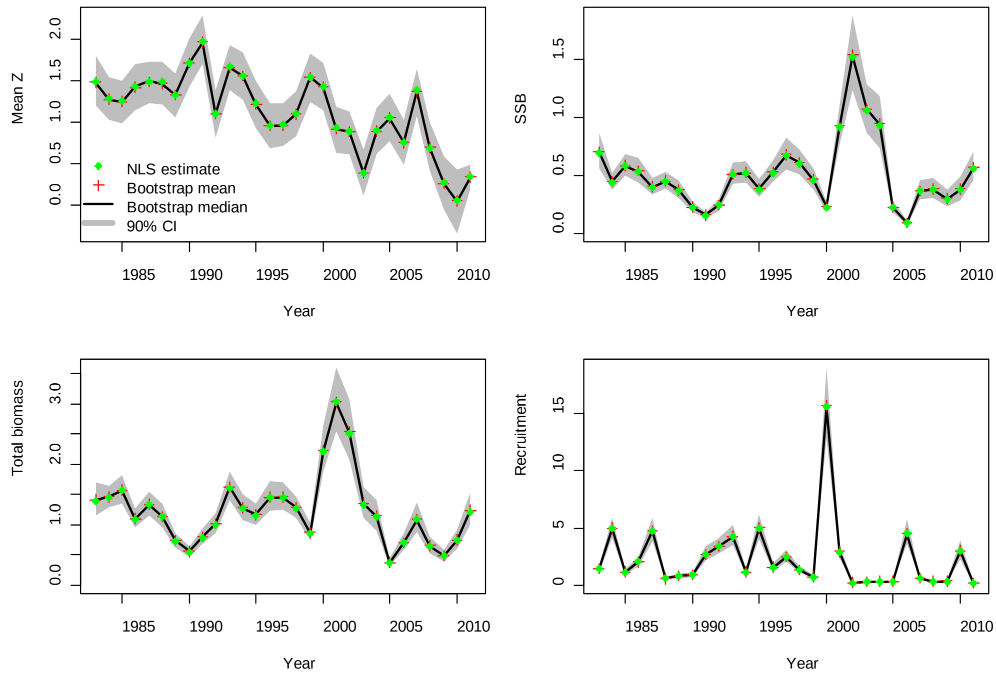

Figure A.2.3. SURBAR summaries for North Sea haddock, using IBTS Q1 with inverse variance down-weighting.

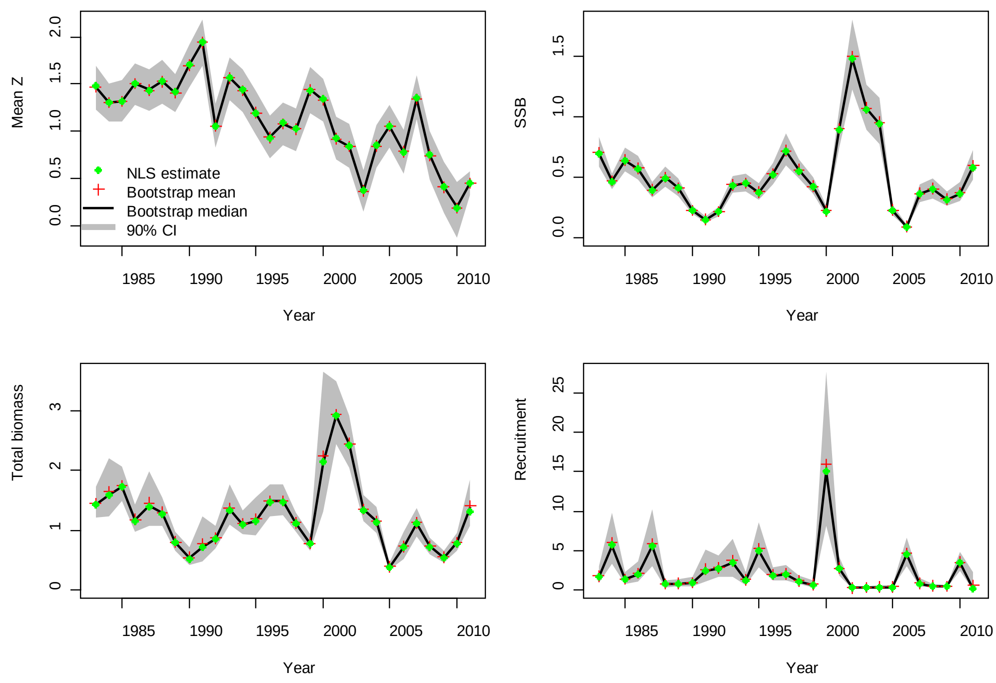

Figure A.2.4. SURBAR summaries for North Sea haddock, using IBTS Q1 with inverse variance down-weighting and all data for age 1 strongly down-weighted further.

Figure A.2.5. Comparison of SURBAR summaries for North Sea haddock.

Annex 3: Haddock 7b-k Combined IGFS-WIBTS-Q4 and EVHOE_WIBTS-Q4 surveys

Working Document to WGISDAA and WKROUND 2012

Hans Gerritsen

Marine Institute, Rinville, Oranmore, Co Galway, Ireland

The IGFS-WIBTS-Q4 and EVHOE-WIBTS-Q4 surveys have nearly full coverage of the haddock VIIb-k stock, with the exception of VIIef (Figure A.3.1). The landings from VIIef have been fairly minor at around of 15% of the total landings from VIIb-k. Both surveys follow the IBTS manual on gear specification and fishing and sampling protocols. This document is aimed at investigating whether it would be appropriate to create a combined IGFS/EVHOE tuning index for the haddock VIIb-k assessment.

The two surveys take place at the same time of year (quarter 4) and a number of inter-calibration tows have been performed whereby the two vessels fish on parallel tracks a short distance apart at the same time. In addition to this, there are a number of trawl tracks that have been sampled by both vessels during the same season (but not at the same time). Figure 1 shows the location of the shared stations and Figure A.3.2 shows the time difference between the shared stations. In most cases the EVHOE survey sampled the stations later with a maximum time difference of 42 days.

The catch rates of the two surveys on the shared stations are reasonably correlated on the log-scale (Figure A.3.3). A linear model was fitted to the log cpue (numbers per hour) of haddock on the shared stations, this model was compared to a model with a slope of one and intercept of zero and the two models were not significantly different (ANOVA: $F(2,63)=2.31$; $p=0.11$). Therefore there is no significant difference in the catch rate of the two surveys.

The length frequency distribution of the catches at the shared stations was also compared. The average length distribution is very strongly dominated by a single station that had a very high catch rate. Therefore the relative length distribution was first calculated for each station by dividing the numbers-at-length by the total catch number at each station. Then the mean relative length distribution was calculated, using the log cpue at each station as a weighting factor. Figure A.3.4 shows that the length frequency distributions of the two surveys are nearly identical.

Because the catch rate and size distribution of the catches at the shared stations are very similar, the two surveys can be combined into a single index. The surface area covered by the IGFS survey in VIIb-k is approximately 30 000 nm^2 while that of the EVHOE survey in VIIb-k is approximately 37 000 nm^2. It is suggested that these values are used as weights when combining the surveys. This way, each country can continue to work up their survey indices which can be combined afterwards.

Figure A.3.5 shows the geometric mean catch numbers per hour. There seems to be a hot-spot of haddock off the west coast of Ireland and haddock are generally abundant in VIIgj but not so much in the main area covered by the EVHOE survey further south.

Figures A.3.6–A.3.8 show diagnostics plots for the combined EVHOE-IGFS index. The new index appears to be reasonably internally consistent.

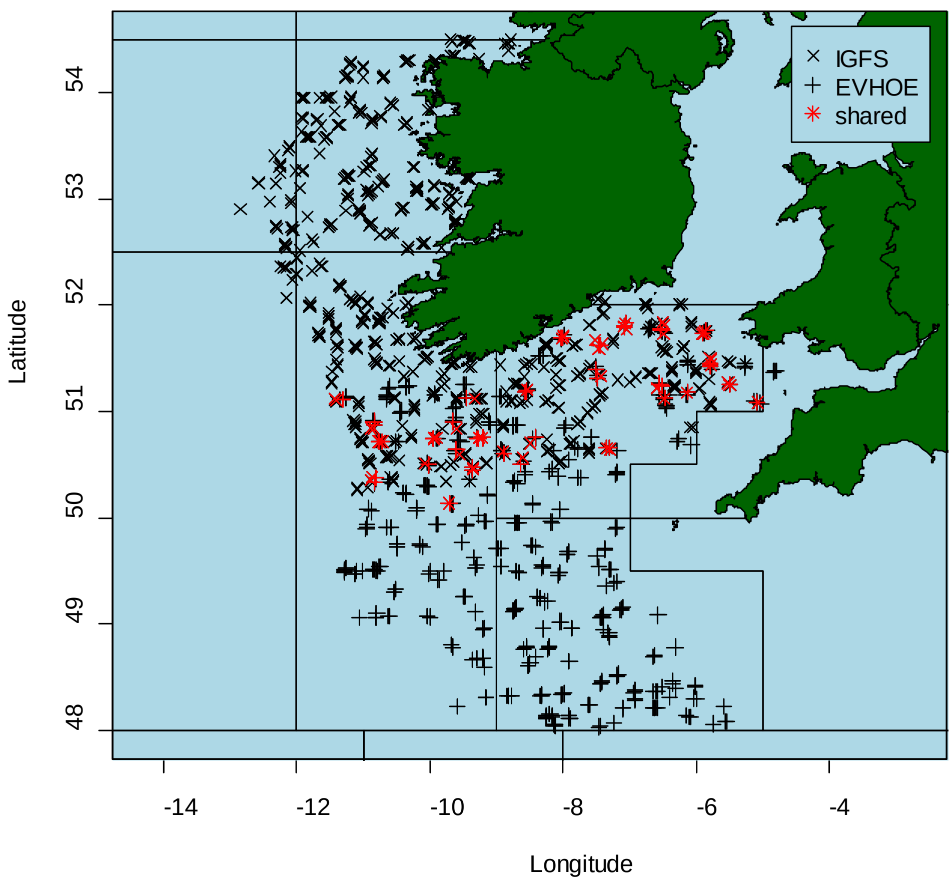

Figure A.3.1. Survey coverage (all stations, full time-series) of the IGFS and EVHOE surveys in VIIb-k. Stations that were sampled in the same year by both surveys are highlighted in red (65 stations in total).

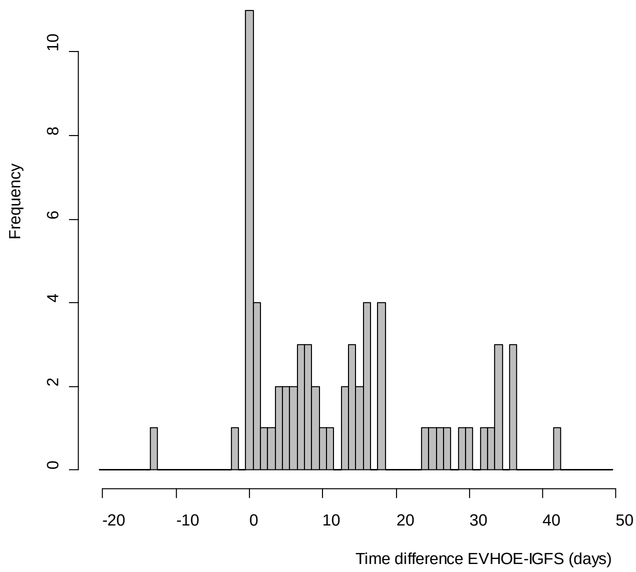

Figure A.3.2. Time difference between the shared hauls of the two surveys.

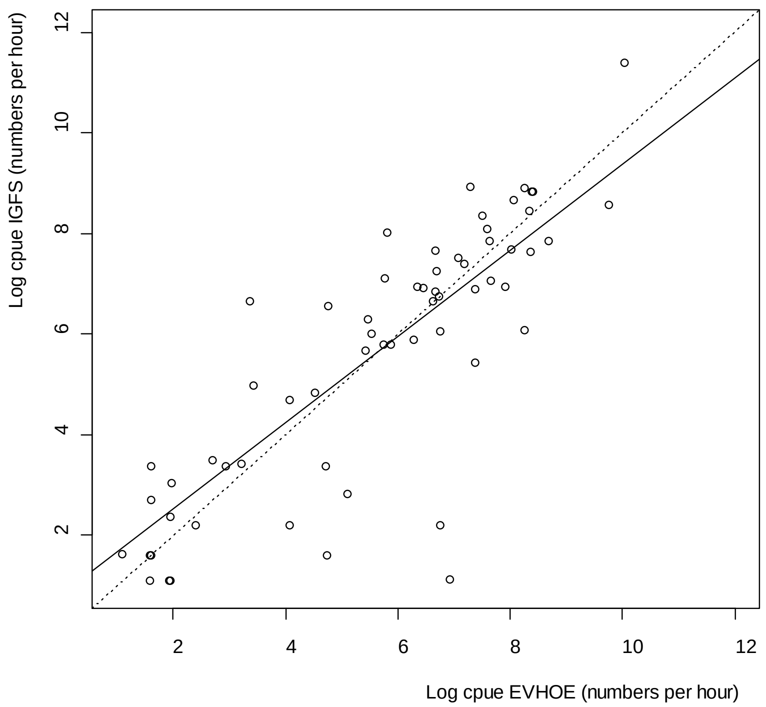

Figure A.3.3. The haddock catch numbers (standardised per hour towed) are reasonably correlated on the log-scale (R^2=0.71). The regression line (solid) was not significantly different from the 1:1 line (dotted); ANOVA: F(2,63)=2.31; p=0.11.

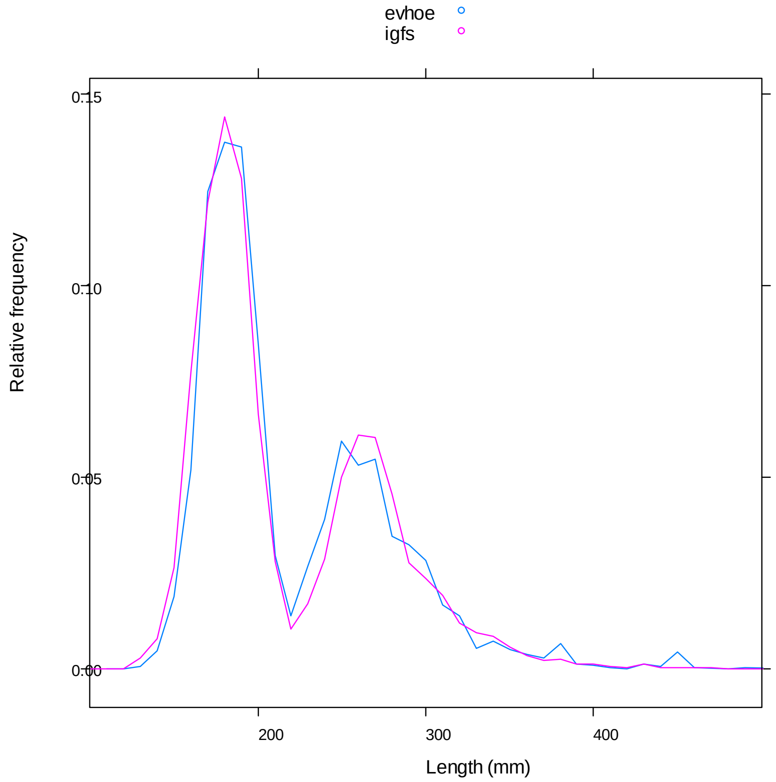

Figure A.3.4. The geometric mean length frequency of the two surveys.

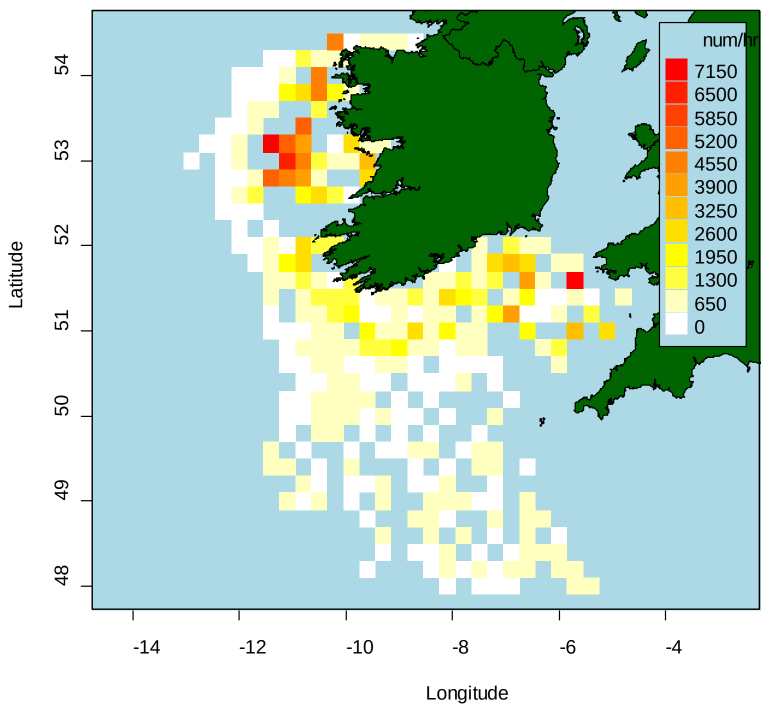

Figure A.3.5. The geometric mean catch numbers per hour of the combined EVHOE and IGFS surveys (all years included).

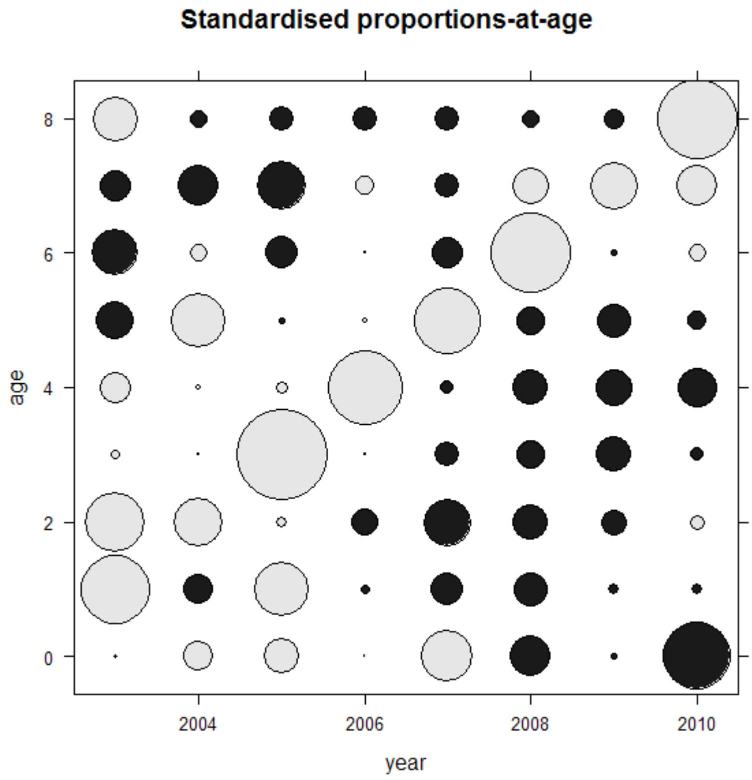

Figure A.3.6. Bubble plot of the standardized proportions-at-age of the combined EVHOE-IGFS survey. Grey bubbles represent larger than average while black represents smaller than average catch-numbers-at-age.

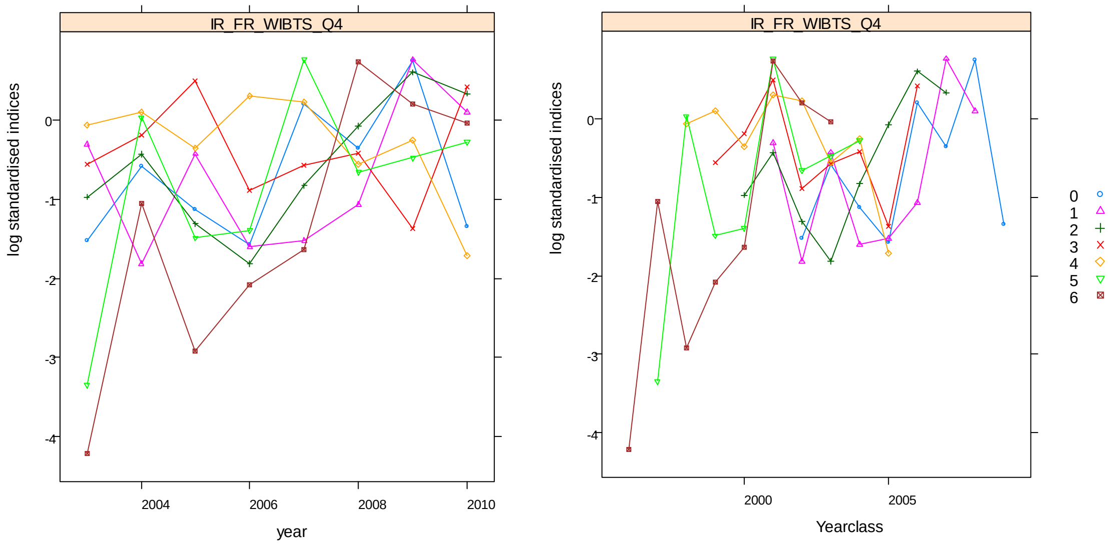

Figure A.3.7. Log-standardized indices by year (left) and year class (right) for the combined EVHOE-IGFS survey. No year-effects are obvious while the strong 2002 and 2009 year classes are clearly identified.

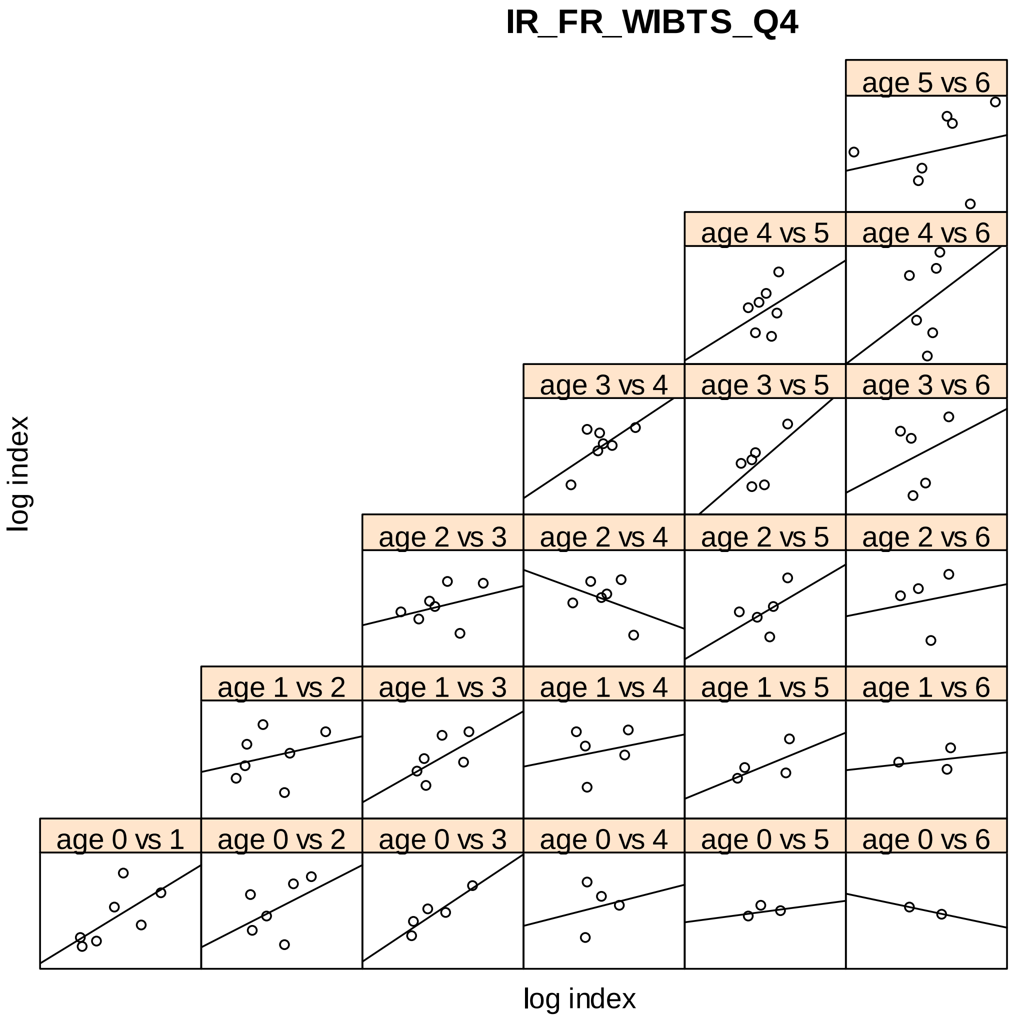

Figure A.3.8. Pairwise plots of the log cpue-at-age for the combined EVHOE-IGFS survey.

Annex 4: WGISDAA terms of reference for the next meeting

The **Working Group on Improving use of Survey Data for Assessment and Advice** (WGISDAA) chaired by Colm Lordan, Ireland and Stephen Smith, Canada, will meet in Dublin, Ireland, 19–21 March 2013 to:

a) Develop a framework and methodology for the analysis of fishery-independent survey information for stock assessment and advisory purposes.

b) Explore and suggest refinements to current survey designs that will improve the quality of data used to support assessment and advisory processes.

c) Investigate methods of combining and or improving indices across multiple surveys and other ways of consolidating survey-derived data.

d) Develop methods for use of survey derived indices and other survey data products as a basis for scientific advice.

e) Request priority case studies from assessment working groups to support the initial activities of the WG.

WGISDAA will report by 20 April 2013 (via SSGESST) for the attention of SCICOM.

Supporting Information

Priority	The formation of such a group is considered a high priority as the outputs will help improve the basis for scientific advice on fisheries from ICES and support the desire for survey reviews from some client commissions.
Scientific justification	Annual surveys provide critical fishery-independent indices for the majority of stocks assessed by ICES. In many instances, survey catch data represent the primary or sole source of information to estimate stock biomass and exploitation rate. A number of ICES assessment working groups including WG Celtic Seas Ecoregion (WGCSE), WG on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK), WG Hake, Monk and Megrim (WGHMM) and WG Widely Distributed Stocks (WGWIDE) have identified particular problems in some survey indices. Lack of stability, conflicting signals between surveys, possible changes in survey catchability and hypersensitivity may occur. All of these issues can severely affect an assessment and subsequent basis of advice. Given the analysis time required, it is not possibly to investigate these in sufficient detail during assessment working groups. The assessment working groups above have independently identified the need for a 'go to' expert group with the appropriate expertise to resolve such data problems (particularly in advance of benchmark assessments). The group would also explore appropriateness of survey design and to develop tools and methods to combine or add in new surveys that are not currently used. This proposal is responsive to three of the thematic areas identified in the current ICES Science Plan: Understanding Ecosystem Functioning. (primary topic – integration of surveys in support of the EAM) Understanding Interactions of Human Activities with Ecosystems (primary topic – impacts of fishing on marine ecosystems), and Development of options for sustainable use of ecosystems (primary topic – marine living resource management tools).

Resource requirements	No specific resource requirements beyond the need for members to prepare for and participate in the meeting.
Participants	These would include stock assessment scientists, survey statisticians and survey technologists.
Secretariat facilities	Sharepoint plus normal secretariat support
Financial	None specific.
Linkages to advisory committees	Close link with all assessment working groups and survey planning groups.
Linkages to other committees or groups	The work of this expert group has strong linkage with assessment expert groups under ICES Advisory Services.
Linkages to other organizations	The work of this group is of direct interest to all organizations and client commissions with a interest in resource assessment

Annex 5: Recommendations

Recommendations	Adressed to
1. Assessment EG chairs should inform WGISDAA of priority stocks where improvments in survey information could be of benefit to the assessment proceedure.	WGCSE, WGNSSK, WGDEEP, NWWG, WGEF, WGWIDE, HAWG
2. Survey planning groups should refer problems with design or index calculation to WGISDAA.	Survey Planing Group Chairs (IBTSWG, SGNEPS, WGBEAM)
3. The WGISDAA meeting in 2013 should focus on the theme of developing indicators for change in distribution of target species in survey.	WGISDAA
4.WGISDAA Should move towards multiannual reporting of activites from next year.	SSGESST
5.WGISDAA should sponsor a theme session on surveys at the ASC in 2013.	WGISDAA and SSGESST
6. WGISDAA should review the benchmark planning for stocks in 2012 and 2013 and consult with stock coodinator where survey related weakness have been identified as an issue.	WGISDAA and Stock Coordinators (WGCSE, WGNSSK, WGDEEP, NWWG, WGEF, WGWIDE, HAWG).