ICES WKCOSTBEN REPORT 2016

SCICOM/ACOM STEERING GROUP ON INTEGRATED ECOSYSTEM OBSERVATION AND MONITORING

ICES CM 2016/SSGIEOM:12

REF. ACOM AND SCICOM

Report of the Workshop on cost benefit analysis of data collection in support of stock assessment and fishery management (WKCOSTBEN)

28 June – 1 July 2016

ICES Headquarters, Copenhagen, Denmark

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. 2017. Report of the Workshop on cost benefit analysis of data collection in support of stock assessment and fishery management (WKCOSTBEN), 28 June - 1 July 2016, ICES Headquarters, Copenhagen, Denmark. ICES CM 2016/SSGIEOM:12. 41 pp.

https://doi.org/10.17895/ices.pub.8654

Contents

Executive summary

WKCOSTBEN was established by the ICES Planning Group on Data Needs for Assessments and Advice (PGDATA). The need for a cost benefit framework is highlighted in the ICES document on "Implementing the ICES strategic plan 2014–2018" (ICES, August 2014). The work plan for the meeting addressed the following tasks:

- Establish what is meant by "cost–benefit framework for data collection", and who it is designed for;
- Identify scope of decisions about data collection and how they could be supported by objective and transparent methods appropriate to the scale of the issue;
- Develop some illustrative case studies around examples of regional data collection programmes (fisheries, surveys);
- Map out a longer term programme for development and implementation of the framework.

Case studies were developed for: age sampling for Kattegat cod on trawl surveys; and sampling of commercial catches for estimating the age and length composition of Norwegian Spring-spawning herring. These are intended as example case studies which could be developed further in subsequent WKCOSTBEN meetings, and should not be considered as definitive at this stage.

WKCOSTBEN also reviewed the results of questionnaires on work done on demersal and acoustic surveys in the Baltic and North Sea, and asking for views on potential for improvement in efficiency. The questionnaires were developed by PGDATA and circulated to four of the laboratories represented in the WKCOSTBEN meeting, to kick-off discussions on the cost–benefits of collecting survey data and provide a first overview of the variability that can be found in time and cost allocations but also protocols implemented in different surveys and MS. The questionnaires raised a wide range of issues that WKCOSTBEN will discuss within the ICES SCICOM/ACOM Steering Group on Integrated Ecosystem Observation and Monitoring (SSGIEOM), which is the parent body for all data collection related expert groups (including the expert groups dealing with research surveys), as part of the WKCOSTBEN series.

1 Terms of Reference and Agenda

The WKCOSTBEN meeting was hosted at ICES HQ, Copenhagen, and had 16 participants from eight countries, with part-time attendance by ICES Secretariat (Annex 1). The meeting addressed the following ToRs:

Propose options and analytical methods for an objective framework to evaluate the benefits vs. costs of datasets used to support stock assessment and fishery management advice, where the benefits are in terms of accuracy (bias and precision) of assessment results and derived management variables, and risks to stocks associated with management under uncertainty. This framework should be able to evaluate existing datasets, new data requests from end-users, and options for focusing elements of funding, survey design, spatial and temporal coverage, and sampling effort towards components of data collection that have greatest influence on quality of assessments and management decisions for particular stocks or groups of stocks.

Identify a range of stocks for detailed case studies, including those with full analytical age-based assessments and data-limited assessments, and contrasting stock status and biology. Describe the data used in the assessments, the design of fishery-dependent and fishery-independent sampling surveys providing the data, including hierarchical cluster sampling designs and analytical methods for quantifying precision reliably. Evaluate sampling rates and allocation for given survey designs that are required to derive estimates with adequate precision. Specify how simulations of the sampling schemes could be used to relate precision to sampling intensity and costs.

Develop a proposal for a longer term (3-year) project to develop a general methodological framework and open-source software to carry out cost–benefit analysis and provide proof of concept using the case study stocks. Identify potential sources of funding.

Identify the need for follow-up workshops in 2017 onwards in the event of no funding for a dedicated project.

ToR (c) was considered by PGDATA (ICES, 2016) to be not an appropriate approach at this stage, and they recommended a 3-year WKCOSTBEN workshop series to develop the cost–benefit framework and supporting case studies. It is recommended that the next WKCOSTBEN workshop be held in 2018, or 2019. This will provide time for National laboratories to develop a suite of programs in R for analysing data from catch sampling programs and scientific surveys to assess and optimize sampling strategies.

The Agenda and work plan for the meetings is in Annex 2. The work plan was based on the following tasks:

- Establish what is meant by "cost–benefit framework for data collection", and who it is designed for;
- Identify scope of decisions about data collection and how they could be supported by objective and transparent methods appropriate to the scale of the issue;
- Develop some illustrative case studies around examples of regional data collection programmes (fisheries, surveys);
- Map out a longer term programme for development and implementation of the framework.

The meeting was conducted through plenary discussions as well as breakout groups scoping out a range of case studies

2 Background to the workshop

The process of revision of the EU Data Collection Framework (Council Regulation (EC) No. 199/2008) has taken many years but has included as a fundamental principle the greater flexibility to meet evolving end-user needs for data. The Expert Working Group 13-02 of the Scientific, Technical and Economic Committee on Fisheries (STECF, 2013) recognized the need for objective criteria by which requests for new data or changes to existing data can be evaluated. Table 2.1 is extracted from the EWG 13-02 report and specifies seven criteria and who is responsible for the evaluation. These criteria include cost–benefit analysis. The STECF (2013) does not specify how the cost–benefit analysis should be done.

The ICES Planning Group on Data Needs for Assessments and Advice, which met for the first time in 2015 (PGDATA: ICES 2015a), has a 3-year work plan which includes development of a cost–benefit framework for data collection needed by ICES for its advisory roles. The first version of the 3-year plan included, in its second year, the 'planning and workshop to develop MSE-type tools for evaluating contribution of data quality to variance of assessment estimates and quality of advice, and evaluating relative impacts of data improvements'. This workshop, though not necessarily restricted to using management strategy evaluation (MSE) methods, would cover elements of a cost–benefit analysis. PGDATA in 2015 (ICES, 2015a) amended its 3-year work plan to include the formation of WKCOSTBEN and added a ToR to PGDATA 2016 to plan this workshop. At the same time, a theme session on this general topic of cost–benefit and optimization of data collection in marine science was submitted for the 2016 ICES Annual Science Conference (ASC) and accepted (Session O: "when is enough, enough?").

The need for a cost–benefit framework is highlighted in the ICES document on "Implementing the ICES strategic plan 2014–2018" (ICES, August 2014). This states that the main objectives of the ICES Integrated Ecosystem Observation and Monitoring programme includes the need to "Identify and prioritize ICES monitoring and data collection needs" and to "Implement integrated monitoring programmes in the ICES area". The implementation document identifies a need to:

- Identify monitoring requirements for science and advisory needs in collaboration with data product users, including a description of variables and data products, spatial and temporal resolution needs, and the desired quality of data and estimates.
- Develop a **cost–benefit framework** to evaluate and optimize monitoring strategies in the context of the capabilities of, and requests from, ICES Member Countries and clients.
- Allocate and coordinate observation and monitoring requests to appropriate expert groups on fishery-independent and fishery-dependent surveys and sampling, and monitor the quality and delivery of data products
- Ensure the development of best practices through establishment of guidelines and quality standards for: (a) surveys and other sampling and data collection systems; (b) external peer reviews of data collection programmes; and (c) training and capacity-building opportunities for monitoring activities.

The goal of WKCOSTBEN is to establish the basis and operation within ICES of a cost–benefit framework. The framework will provide a decision-support system ensuring that any requests by ICES for changes to data collections needed for its advisory role are fully transparent and objective, include clear evidence of how the data will be used and the benefits that are expected, and take the needs for cost-efficiency into account.

A cost–benefit framework is also needed by other groups tasked with evaluating data needs and delivery, such as STECF and the EU Regional Coordination Groups to be set up in 2017 to coordinate regional data collection under the revised Data Collection Framework. The RCGs will evolve from the current annual Regional Coordination Meetings, and will need to consider cost-efficiency in the establishment and implementation of national and regional sampling programmes funded by the revised DCF through the European Maritime and Fisheries Fund. The framework developed and demonstrated by WKCOSTBEN will support the decision-making responsibilities of RCGs in relation to data collection, as well as help national scientists and funding bodies to make objective decisions on investment in data collection programmes. The case studies chosen by WKCOSTBEN to demonstrate elements of the framework are highly relevant to the RCGs and national fisheries agencies.

Table 2.1. Proposed criteria for evaluation of proposed changes to dataseries in DCF (STECF 2013: EWG 13-02). (Note that "responsibility" for some topics other than need and relevance could lie with the end-user requesting the change in data collection).

Topic	Responsibility	Addition of new data series	Amendments to existing data series	Cessation of existing data series
Need and Relevance	End user	Reasons and legal basis for the need or relevance	Reasons for change to need or relevance	Reasons for change to need or relevance
Impacts	RCG; PGECON or Expert Groups	Expected improvements for end user purposes. Precision needed to deliver expected improvements. Impacts on ability to maintain existing data series	Impacts on data quality and end use Impacts on ability to maintain existing data series	Impacts on ability to respond to end user needs
Feasibility	RCG; PGECON	Feasibility of collecting the data, especially to required precision and accuracy	Feasibility of collecting the data, especially to required precision and accuracy	
Methods	RCG; PGECON	Sampling designs and data collection methods needed; Who will implement the schemes; Anticipated sampling rates in relation to desired precision.	Changes to sampling designs, methods, sampling rates and costs.	
Costs	RCG; PGECON	Cost – benefit analysis	Cost – benefit analysis	Cost – benefit analysis
Data quality	RCG; PGECON	Data archiving and quality assurance; Quality indicators for the data	Data archiving and quality assurance; Quality indicators for the data	
Data use	RCG; PGECON / Expert Groups	Process and methods for analysis of the data (models etc.) and application of the results	Process and methods for analysis of the altered data (models etc.) and application of the results	

3 ToR (a): propose options and analytical methods for an objective framework to evaluate the benefits vs. costs of datasets used to support stock assessment and fishery management advice

3.1 Terminology

Several terms are used widely in this context, including cost–benefit, cost-effectiveness and cost-efficiency. There is also a related concept of risk-benefit.

A typical cost–benefit analysis in the business world calculates the costs of different options for setting up and running a new project with the expected future revenues and profits for each option (adjusted for inflation to give the net present value, NPV). If the NPV of future benefits exceed costs, the project may be considered "cost-effective", but there would also be an evaluation of which options are most "cost-efficient", i.e. providing the greatest benefits for the same or lower costs. We have used the term cost–benefit-framework (CBF) in our report to represent the entire process of evaluating cost-effectiveness and cost-efficiency. See PGDATA 2016 report (ICES, 2016) for further discussion around terminology and concepts.

3.2 Defining the structure of a cost–benefit framework and how it will operate

3.2.1 What is meant by a framework in this context

The cost–benefit framework for data collection supporting stock assessment and fishery management is essentially a set of guidelines to help people make objective decisions on initiating or changing data collection programmes or components of programmes, and the responsibilities, time-scales, tasks, and outputs needed to inform the decision-making process. The scale of the tasks and personnel numbers, skill sets and time involved depends on the magnitude of the issue, the anticipated impacts of any changes to the data collection on fishery management, and the likely risks involved. The table 3.1 below suggests extreme cases from minor evaluations of small changes to data to major evaluation of large and expensive changes to data collection that may have significant impact on the ability to manage fisheries more effectively. In between, there are simpler approaches, for example where the goal is simply to collect data more cost-effectively without making much change to the quality and use of the data. Also, simpler approaches are possible to examine impacts of improved data quality on stock assessments, such as sensitivity testing of an assessment, without carrying out a full MSE exercise.

Table 3.1. The effort of conducting cost-benefit analysis versus scale of impact and expected risks.

Scale of impact of the change	Expected risks	Scale of tasks for cost–benefit analysis	Example tasks
Low	Low	Days; single expert	"Rules of thumb" advice based on existing knowledge.
High	High	Months; multidisciplinary teams	Information gathering; end-to-end modelling using simulations of changes to data collection and impacts on assessment quality for range of options, with management strategy evaluations.

3.2.2 Summary of elements of the framework

In setting up a cost–benefit framework for data collection within ICES, the different elements of the framework need to be clearly defined. Specifically:

1) What are the drivers and objectives – long-term strategic and short-term reactive
2) Who it is targeted at and what they need to know
3) Operational elements of framework, and the information and analysis tools needed
4) Implementation: who – when – how
5) Outputs and process of communication
6) How to embed the framework alongside quality assurance frameworks within the ICES Expert Group, benchmark and Steering Group structures.

3.2.3 Drivers and objectives of the framework

3.2.3.1 Longer term strategic objective in the context of ICES

The ICES document on "Implementing the ICES strategic plan 2014–2018" (ICES, August 2014) contains an objective to:

"Develop a **cost–benefit framework** to evaluate and optimize monitoring strategies in the context of the capabilities of, and requests from, ICES Member Countries and clients."

ICES is currently focusing strongly on improving its ability to identify and evaluate data needs for regional integrated ecosystem assessments and for providing annual or multiannual advice to the Commission on fishing opportunities. A critical aspect is to improve efficiency of the entire process, both from a structural and process aspect (minimizing the burden of meetings and analytical work) and from a data aspect (prioritization of data needs; improving data quality; improving management and use of data). This is a process of improving the longer term benefits provided by ICES to clients and ICES Member Countries, while maintaining or reducing the costs of this through more efficient processes. Cost–benefit analysis may also identify cases where substantial improvement in accuracy (increased precision and reduced bias) can be achieved for minimal additional costs through improvement of survey sampling methods.

WKCOSTBEN considers that there is a series of longer term goals to meet this strategic objective in relation to data collection and use:

- Building the statistical and practical expertise in the components of sampling design, implementation and analysis in data collection. This includes fishery-dependent and fishery-independent sampling supporting stock assessments as well as collection of data for integrated regional ecosystem assessments and associated indicators. There is a need to attract existing experts with the necessary skills as well as building future capacity through training.
- Closer integration of experts in the fields of data collection, data management and stock assessment methods and other types of modelling needed

by ICES to fulfil its role, with a clear mandate and working process and making the most effective use of the existing expertise within ICES.

- A system for expert review of data collection programmes currently in place in each regional ecosystem, focusing on statistical design, implementation, data quality, data management, and analysis methods. This could be achieved as part of regional benchmarks.
- Prioritization of data needs within each region. This should take into account the cost–benefit relationships, linking costs of data, precision of derived estimates for input to models, and precision of assessments using the data. This will need data estimation procedures and assessment methods capable of this analysis.
- Development of databases structured to facilitate access to the required data as well as holding information on the sampling design and implementation needed to ensure unbiased estimates and associated precision where needed.
- Building libraries of software tools and routines for providing diagnostics of quality of data held in the databases, and for deriving estimates from data, including estimates of variances and covariances and other data quality metrics needed for use of the data and evaluating its impact on quality of the results of stock assessments and other modelling approaches.

ICES can build on ongoing methodological developments, for example the Norwegian REDUS project (www.redus.no) described in Section 4.3.

3.2.3.2 Shorter term reactive objectives

A major driver for improving cost efficiency is the conflict between expanding data needs at a stock/region scale, and stable or reducing budgets for data collection in individual countries. Objective procedures are needed to help national agencies make informed decisions about what data should continue to be collected and how it should be collected (Figure 3.1). There may be pressures to make decisions at relatively short notice in response to the national timetable for funding allocation. The DCF makes legal obligations for Member States to collect and supply specified data to end-users, and this is often a driver for prioritizing allocation of funds and perhaps limiting the funding to the minimum sufficient to meet the DCF requirement. The cost–benefit framework proposed by WKCOSTBEN, and the case studies, would prove useful to help national agencies evaluate options for improving cost efficiency.

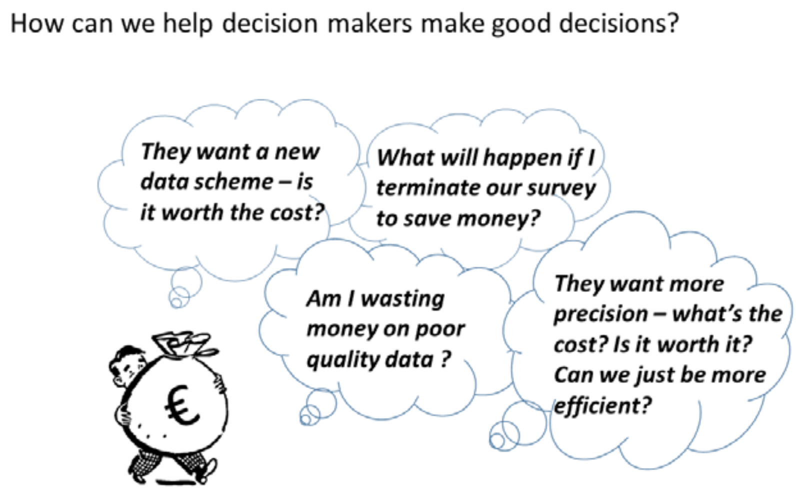

Figure 3.1. Why we need a cost–benefit framework

3.3 Broader picture of cost–benefit within a fisheries system

Fisheries and their management represent complex interrelated systems of costs and benefits. Fisheries management attempts to achieve sustainable fishing using a system of stock assessments and scientific advice such as provided by ICES, together with a regulatory system establishing fishing opportunities in line with (for example) advice based on the MSY approach, and a system of fishery surveillance and control to incentivise and monitor compliance (Figure 3.2). Fishers have to bear compliance costs of the control system, including alteration of gears, effort limitation or exclusion from new Marine Protected Areas, and have to face uncertainties in future catches and profits due to changes in stock size and market prices. The costs of the scientific monitoring and the regulatory system also have to be considered in relation to the benefits to the fishery and ecosystem by managing fishing impacts. There will be pressures to make surveillance and monitoring activities as cost-efficient as possible, but if there is a reduction in quality of scientific data, or less compliance due to reduced fishery inspection, there will be increased risk of overfishing.

There are therefore several dimensions to cost–benefit analysis of a fishery system:

- Profits vs. costs for the fishery, including compliance costs;
- The relationship between expenditure on fishery surveillance and control, and the level of compliance with fishery management measures imposed.
- The relationship between expenditure on scientific monitoring and assessment, and the bias/precision in estimates of management targets and thresholds (e.g. F_{MSY}, $MSYB_{trigger}$, B_{lim}, and F_{lim}) and of estimates of stock status relative to these.

The European Commission has developed guidelines for evaluating the cost–benefit of regulatory systems (EC tool #52: methods to assess costs and benefits http://ec.europa.eu/smart-regulation/guidelines/tool_52_en.htm which advises on how costs can be evaluated, and how benefits can be looked at for direct benefits, and indirect benefits. The website describes how different methodological approaches can be used to estimate costs and benefits *ex ante* (within impact assessment work) or *ex post* (in retrospective evaluation/fitness check work). The most appropriate choice will depend on several factors including the nature of the initiative and the availability of data.

Equivalent guidelines for evaluating cost–benefit of scientific monitoring activities do not exist, and WKCOSTBEN represents a first step in this direction within Europe. It is expected that this will be a long-term process.

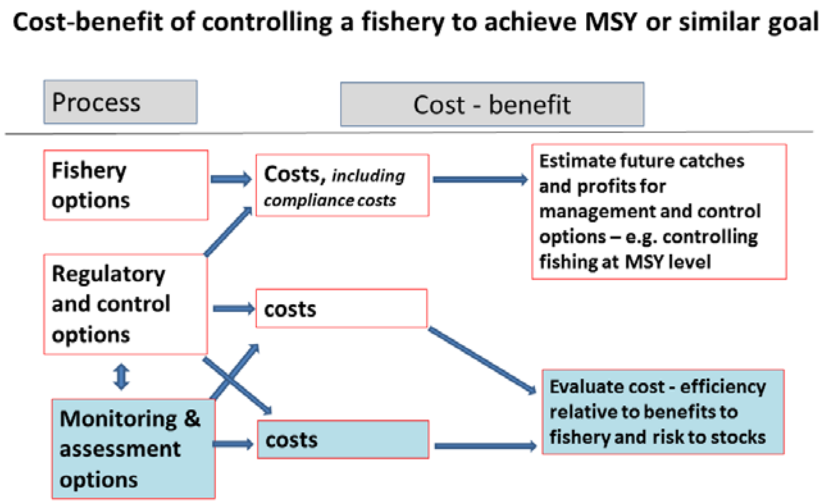

Figure 3.2. Managing a fishery to achieve a goal such as MSY. Benefits to the fishery are long-term profits (expressed as net present value) after costs of fishing are deducted. There are also costs associated with assessing the state of the stock and providing fishing opportunities advice, and implementing a regulatory and control system to ensure management is effective. There will be pressure to make this as cost efficient as possible, but reductions in effectiveness increase uncertainty in assessments, poorer compliance, and increased risk to stocks.

3.4 Key relationships for evaluating cost-benefit of scientific data used for management advice

The key relationships for the cost–benefit analysis are shown and explained in Figure 3.3. WKCOSTBEN focused on how to determine the relationships between precision of estimates from sampling schemes and the design and sampling intensity for the schemes, and the relationship between precision of the estimates from sampling schemes and the precision of estimates from the stock assessments using the data (i.e. Figure 3.3(a) and (b)). The case studies identified at WKCOSTBEN focused on the relationships in Figure 3.3(a), asking the question of how cost-efficiency could be improved through changes to design and sampling intensity. It is intended that future WKCOSTBEN meetings will implement analytical methods to quantify the relationships in Figure 3.3 (a) and (b) – i.e. looking at benefits in terms of precision of stock assessment results.

The system of data collection and supply to ICES assessment EGs, and use of the data in stock assessment models, don't currently allow an accurate evaluation of these relationships in the great majority of cases. Two conditions are necessary:

- Estimates of precision (including covariance structure) in data inputs such as multinomial age compositions, catch estimates and abundance indices should be available that correctly reflect the sampling design and sampling effort.

- The assessment model should include appropriate precision metrics for sampling data, and the confidence intervals for assessment model outputs should accurately reflect the combination of sampling errors in the input data along with the additional random process error due to assumptions in the model.

In many statistical assessment model applications, these conditions are not met, for example where catch estimates derived from sampling (discards; recreational catches) are treated as exact, or where covariance structure in data such as age compositions is not accounted for. It is also inevitable that other assumptions about model parameters such as natural mortality, selectivity, catchability and growth cause additional residual error in model fits to input data over and above the true sampling error. However the key requirement for the cost–benefit analysis is that, if the model is re-run with different versions of one or more input datasets with different magnitude of sampling error, the effect on precision of assessment model outputs is accurately represented (i.e. Figure 3.3b).

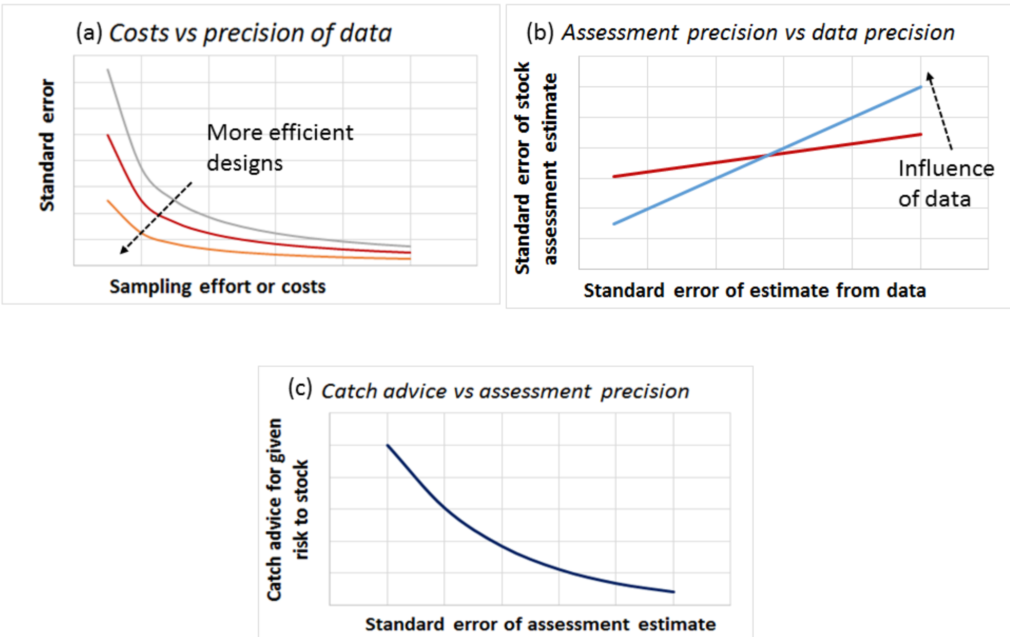

Figure 3.3. Theoretical examples of the relationships important for evaluation of cost efficiency of data collection:(a) Precision (standard error) of an estimate such as fleet discard quantities as a function of the sampling effort or costs, for several sampling designs of varying efficiency; (b) How the precision of a stock assessment estimate such as current SSB might vary according to the average precision of an input dataset. Increasing the precision of relatively influential datasets is likely to have greater impact on the assessment precision, indicating a greater cost–benefit; (c) curve showing that the advised catch which leads to a risk of X% of an undesirable outcome (e.g. risk of SSB falling below the limit reference point) will be higher for a more precise assessment, indicating a benefit that can be compared with the cost of improving the precision of input data.

The WKCOSTBEN case study on Norwegian spring-spawning herring (Section 4.4) describes new software developed in Norway for design-based estimation of variances and covariances in input data from sampling schemes, and a new statistical assessment model configuration that can correctly handle these data precision metrics. <u>WKCOSTBEN strongly supports the continued development and wider dissemination of these methods.</u>

3.5 Costs and benefits of reducing bias vs. improving precision in input data for stock assessment

Precision is an easier metric to deal with than bias in cost–benefit analysis as it is (for a well-designed probability-based sampling scheme) directly related to the survey design and sampling intensity (number of sampling units) which can be converted to economic costs of data collection. This emphasizes the need for sound statistical design of data collection schemes to minimize bias and allow correct estimation of random sampling error.

Bias in assessment datasets can in many cases be a greater issue to address than precision. For most sampling programs employed in fisheries science the true values of key parameters and statistics (e.g. the exact number of fish by age class caught annually from a stock in a commercial fishery) will never be known exactly. Thus, it is not possible to quantify the bias by comparing estimates to the true population values. If the design and implementation of a sampling scheme leads to substantial bias, increasing the sampling intensity will improve precision, but may not reduce bias. The total error may still be dominated by the bias and it is impossible to correctly define a relationship between sampling intensity and accuracy of estimates such as shown in Figure 3.3(a). Many potential causes of bias exist in data collection (ICES 2008). A few examples include:

- Incomplete frame coverage, or strata missed during implementation;
- Non-access to fishing vessels or catches due to refusals;
- Incorrect target strength relationships in acoustic surveys;
- Use of ad-hoc or non-probability-based sampling designs;
- Observer effects in at-sea sampling;
- Unaccounted-for changes in catchability in surveys or commercial cpue.

Since bias generally cannot be quantified, it is essential that all aspects of the data collections and estimation be conducted in accordance with best scientific practice to minimize sources of bias. It is important that potential for bias be evaluated through review of the design, implementation and analysis of data, rather than looking only at trends in residuals in an assessment model as these may be driven by incorrect assumptions in the model such as assuming constant fishery selectivity, growth or natural mortality when it is changing non-randomly over time.

The important questions on bias for cost–benefit analysis are then:

- What is the potential magnitude of the bias?
- What is the potential impact on the assessment accuracy?
- How can the bias be mitigated through better sampling design, and what is the cost of this?

The potential impacts of bias can to some extent be evaluated using sensitivity analysis of the assessment using plausible alternative scenarios for potentially biased datasets. Stock assessment EGs also commonly carry out sensitivity analysis where individual datasets are down-weighted or even removed for all or part of the time-series to examine the relative effect on assessment trends and precision. This is usually done one at a time, underestimating the overall sensitivity of the assessment to data choices. More appropriate methods such as Global Sensitivity Analysis (Saltelli *et al.*, 2008) are available that explore sensitivity to all possible combinations of data choices. This could help prioritize bias issues for investment in methods to reduce bias where this is needed.

3.6 Options for improving cost-efficiency

The cost–efficiency of scientific monitoring of stocks (fisheries-independent surveys) and catch sampling programmes (fisheries-dependent surveys) could be viewed in relation to single-species assessments, multispecies assessments, mixed fishery assessments or integrated regional ecosystem assessments where these influence the annual decisions on fishing opportunities. The WKCOSTBEN has focused this year mainly on single-species assessments, but the proposed framework should be able to consider the multispecies dimension in relation to data collection methods such as trawl surveys and fishery sampling where a survey will deliver data from more than one species or stock.

Cost efficiency can be improved in many ways (these are not exhaustive):

- Better sampling design and implementation (e.g. improved coverage; move to more statistically sound designs that can be optimized to give the same precision with less sampling effort; procedures to minimize bias; reduced tow duration on trawl surveys if this can provide more accurate estimates due to time for additional tows or more accurate sampling of catches);
- Reducing staff time in obtaining data (e.g. using electronic data entry; automated data collection or sample processing; optimized RV survey tracks; fisher self-sampling schemes);
- Improving accuracy of measurements (e.g. better training and protocols; quality assurance framework; age calibration studies etc.)
- Reallocation of resources between different data types according to the contribution of the data to overall benefits. E.g. relative investment in fishery-independent surveys vs. sampling fish catches at sea or on shore.
- Development of databases and software tools to streamline the tasks of data quality assurance, data analysis and reporting.
- Making use of previously unused data that might be available from other sources, where considered appropriate (e.g. data on discards collected during fishery inspections, such as the "last haul" scheme; size compositions of retained fish collected during discard observer trips).

In all cases, a thorough evaluation of such options, using (where possible) analytical approaches to demonstrate the expected improvements in cost-efficiency, should be conducted. The WKCOSTBEN case studies provide some examples of this for the first example given above.

3.7 Skills and analytical tools needed for cost-benefit analysis of stock assessment data

Even simple exercises to improve the cost-efficiency of data collections require a fundamental understanding of how data are being collected, the statistical soundness of the designs and analysis, data quality problems arising during implementation and data archiving, and the human and financial resources being spent on the different elements of the data collection system, including any lab work such as otolith processing and reading. These skills should be present in any laboratory conducting data collection programmes.

Improving the design of a data collection programme can be done most rigorously by developing simulation models of the underlying processes that affect how data should be collected to meet specified objectives. For example, the distribution and scales of patchiness of species in a sea area can be simulated to examine different trawl or acoustic survey designs that are optimal for providing the most precise estimates for particular species or as many species as possible. Simulations can also be carried out to evaluate alternative schemes for collecting biological data (length, age, maturity etc.) from individual sampling stations on a survey – this allows an evaluation of the relationship between overall precision and the balance of sampling between primary sampling units (e.g. numbers of trawl hauls sampled in each stratum) and numbers of fish sampled at the PSU, and also the way in which fish are sampled at a PSU (e.g. random or length-stratified). The WKCOSTBEN case study on Kattegat cod examines these relationships.

For fisheries sampling, the patterns of landings of different species and sizes of fish, and the magnitude of landings, into all the ports in a region can be simulated using historical data from logbooks and trip-reports to evaluate stratified random designs for optimized collection of data across a range of species. Simulations of this nature, though not extended to include length and age sampling, were carried out for the recent EU project "Strengthening regional cooperation in the area of fisheries data collection" (EU MARE grant MARE/2014/19[1]).

3.8 Calculating the costs

A detailed examination of cost-efficiency requires a breakdown of time and costs at each stage of the data collection. For a full sampling programme, such as an observer scheme or a port sampling scheme, this should include all stages from the implementation of the sampling through to lab processing, data archiving, quality assurance, data analysis and management of the programme such as ongoing monitoring of the sampling. A change in design or intensity of sampling will affect the costs of each of these stages to differing degrees.

[1] http://www.masts.ac.uk/media/36045/fishpi_report_-final-4-8-16.pdf

4 ToR (b) Case studies

4.1 Case study 1: Evaluating costs and efficiency on research vessel surveys

WKCOSTBEN reviewed the results of questionnaires on the breakdown of tasks on demersal and acoustic surveys in the Baltic and North Sea, and which also asked for views on potential for improvement in efficiency.

The questionnaires were developed by PGDATA and circulated to four of the laboratories represented in WKCOSTBEN, to kick-off discussions on the cost–benefits of collecting survey data and provide a first overview of the variability that can be found in terms of time and cost allocations but also protocols implemented in different surveys and MS.

A brief analysis of the questionnaires is available on the WKCOSTBEN SharePoint site. It raises a wide range of issues that WKCOSTBEN will discuss within the ICES SCICOM/ACOM Steering Group on Integrated Ecosystem Observation and Monitoring (SSGIEOM), which is the parent body for all the data collection groups (including the expert groups dealing with research surveys), as part of the WKCOSTBEN series. It is therefore a work in progress that will be presented at future meetings following input from the SSGIEOM.

4.2 Case study 2: Kattegat cod sampling

Background

In this case study, we will use a cost–benefit approach to evaluate the contribution of different data sources to the uncertainty in the assessment of the Kattegat cod stock. With this information it will be possible to identify the areas where gains from optimization of data collections would be most cost-effective. Since we in this study only consider as end-users the stock assessment scientists that are responsible for management advice, we will not address how the data could be used for other purposes, such as ecosystem analysis or for assessing good environmental status.

An important question to be addressed is if the quality of the assessments and advice could be maintained or even improved within a fixed cost of data collections by (for example) improving the accuracy (i.e. reduce bias and increase precision) in estimates of catch-at-age data, or abundance-indices by age. We will assess how a decrease or increase in the sampling effort in some datasets would affect precision in key estimates.

In some cases, the changes in survey design and sampling effort may have a major effect on assessment quality, whereas other changes may have only a minor impact. If the costs of these changes in data collections can be quantified, then cost-efficiency metrics can be provided to identify where resources would best be allocated to improve the assessments of multiple stocks. However, it can be a very challenging task to get the full overview of all data used in a given assessment, and to quantify the uncertainties of the different input data for all the relevant stocks.

As a first step in the cost–benefit framework we set up some simple diagnostic tools to evaluate if it is possible to decrease sampling effort (to free time for other tasks) without losing essential information, i.e. for assessment purposes. To be able to improve the assessment, additional analysis will also have to be conducted to determine were data need to be improved.

An aim with this case study is to produce some simple analysis on data variability depending on the amount of sampled data. The analyses are conducted in R using the DATRAS format, thereby producing scripts that can be applicable for many other surveys currently used in the ICES system. We plan to share the computer code with the scientific community through the ICES repository GitHub, where it could be further developed to include different data sources in future.

Data sources

The Kattegat cod stock was used for the case study and the first exercise was to analyse current surveys to assess how the precision (CVs) in key estimates depends on the numbers of hauls (primary sampling units, PSU) vs. the number lengths or ages collected by subsampling fish from each PSU.

Survey data

The survey data used in the preliminary analyses comes from the Fishermen's Research Survey. This annual survey targeting cod in the Kattegat has been carried out since 2008 with the exception of 2012. The survey is conducted in November-December by four commercial trawlers from Denmark and Sweden. The survey design has been largely fixed during the years, but a fourth strata representing the closed area in Southern Kattegat was added year 2014. The survey is designed for cod, but the total catch and lengths of all fish species and Norwegian lobster is recorded. Age sampling is only done for cod; the original instructions were to collect two otoliths per cm class and haul, up to five otoliths per cm class and area (North and South, see Figure 4.1). Since then, the instructions for Swedish vessels has been changed to sample more otoliths, and from 2016 otoliths are sampled from all hauls.

Survey design

The survey is designed as a stratified random bottom-trawl survey. Data are raised by strata allowing for re-stratification between years if necessary. The survey area was stratified into three geographic strata during 2008–2013: (1) a stratum with expected high density of cod, (2) a stratum with medium density and (3) a stratum with low density of cod based on information from the fishers. In 2010 and 2011 there was a minor re-stratification to adapt the areas to the catch information collected during the former years. In 2014 a fourth stratum was added to better ensure that data be collected from the area closed for fisheries. Each stratum is further subdivided in 5*5 nm squares (sections). The high density, medium density and closed area stratum has been allocated relatively more stations than the other strata (Figure 4.1).

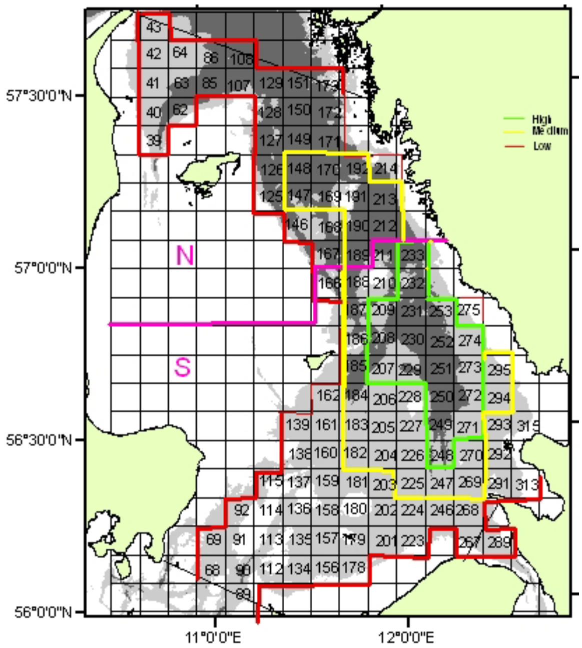

Figure 4.1. The stratified survey area (2011) with section numbers. Green High density of cod. Yellow Medium density. Red Low density. N and S Northern and southern area, respectively.

Station (tow) location

The survey is planned with 20 hauls in 6 days for each of the 4 vessels, i.e. in total 80 trawl hauls. The hauls are allocated randomly to the 5*5 nm squares within strata and each vessel will fish in 20 different squares (Table 4.1). In the closed area, several vessels are allowed to fish in the same square within the high and the medium density strata. The low density area is divided in a Southern and Northern area, and only one haul is allocated in each square.

Table 4.1. Showing planned number of stations by vessel, stratum and area. In 2013 were only 2 Swedish vessels participating in the survey.

Year	No of vessels	high density	medium density	low density	closed area	**total hauls by vessel**	**total haul survey**
2008	4	6	8	6		**20**	**80**
2009	4	6	8	6		**20**	**80**
2010	4	6	8	6		**20**	**80**
2011	4	9	6	5		**20**	**80**
2013	2	15	10	10	5	**40**	**80**
2014	4	6	5	7	2	**20**	**80**
2015	4	6	5	7	2	**20**	**80**

Preliminary analyses

To date, the following issues have been examined using simulation and bootstrap analysis.

- Number of stations/hauls – can sampling effort be reduced while providing abundance age composition and abundance indices that do not significantly differ from estimates based on actual sampling effort. How is precision affected by reduction in sampling effort?

- Number of length samples - can sampling effort be reduced while providing the same overall length composition?
- Number of age samples - can the number of samples per length class be reduced or the size of the length class increased while providing the same overall age composition?

Simulations

To investigate the number of length samples needed we used cod length data from the Fishermen research survey. So far we have not included the actual stratification from the survey design in the simulations; instead we collapsed strata by year and country for simplicity, under the assumption that bias in overall estimates will be the same under varying sampling effort. For the simulation, data were resampled according to a stratified 2-stage random sampling design. At the first stage; *n* hauls within each stratum were randomly sampled with replacement, at the second stage; *m* fish within each selected haul were randomly sampled with replacement. At the second stage, *m* was not allowed to be larger than the actual number of fish in the haul. Figure 4.2 illustrates the uncertainty in length composition (measured as relative standard error of mean length in percentage) in relation to the number of hauls and the number of fish sampled.

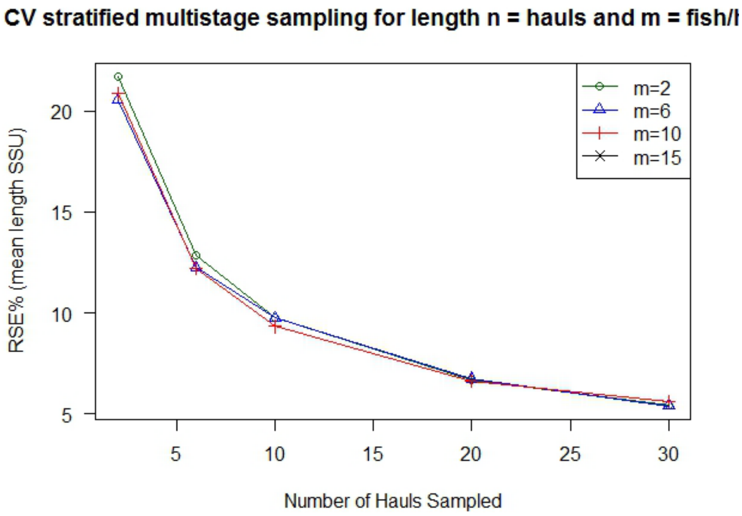

Figure 4.2. Relative standard errors (%) of mean length estimated from simulated trawl surveys making 2–30 hauls. From each haul a sample size of *m* fish were sampled for length. 1000 replicates were simulated for each haul – sample size combination.

A similar approach was used to investigate how sample size influences the precision of the estimated age composition. To account for the fact that age sampling usually is carried out by length (e.g. *m* samples per 1 cm length class) the second stage of the sampling design was modified so that within each selected haul a stratified sampling design was used to sample *m* fish within each length class. Note that for this part, the number of length classes vary by haul and the number of fish per length class within each haul; *m* cannot be larger than the number of fish in the length class. It should be noted that in the present simulation this limitation will result in an underestimate of the variation in length classes with few individuals to sample from.

Figure 4.3 illustrates the result from 1000 simulations. In this example 2 fish per length class were sampled for age. Total number of number of samples was varied by changing the length class width from 1, 2, 5, 10 cm with sampling probabilities proportional to the frequency of the original 1-cm length classes (i.e. whereas the first setting samples 2 fish per cm, the final setting samples 2 fish per 10 cm). Again, it is clear that the number of hauls has greater influence on the uncertainty than number of fish per length class* haul.

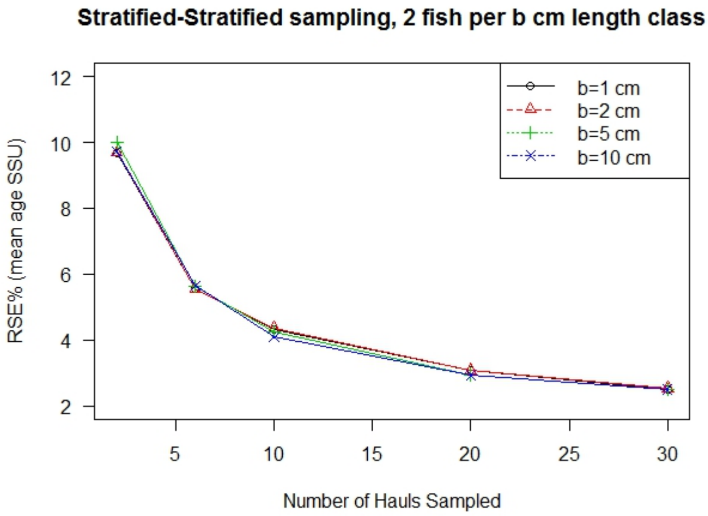

Figure 4.3. Relative standard errors (%) of mean age estimated from simulated trawl surveys making 2–30 hauls. From each haul 2 fish per length class were sampled for age. In the simulations 1, 2, 5 and 10 cm length classes were used with sampling probabilities proportional to the frequency of the original 1-cm length classes. 1000 replicates were simulated for each haul – length class combination.

The results clearly suggest that the best efficiency is obtained when more hauls are sampled rather than more fish per haul. Further evaluation is needed to check how representative the age/length distributions from smaller samples are and if the loss of information from the tails of the distributions is within acceptable limits. One possible continuation would be to repeat the simulations for other potential target species and construct a compound measure for evaluating minimal effort allocation among and within hauls.

To evaluate the effect of sample size on the abundance indices used for assessment it is possible to simulate different sampling scenarios and calculate the index of interest. In Figure 4.4, the result of such an exercise is presented. Catch-at-age was estimated from simulated trawl surveys making 10–80 hauls from 2008 to 2015. The resampling of data was made with replacement according to the actual survey design and inclusion probabilities. From each simulation the estimated number of cod aged 3 or older (3+) was calculated as a proxy for the spawning stock size. 1000 replicates were simulated for each survey design to generate and estimate of the variation, measured as relative standard error, RSE in percent (Figure 4.4).

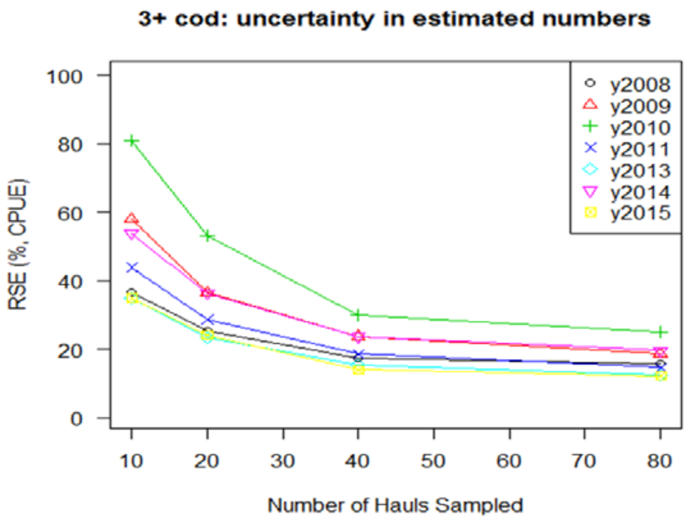

Figure 4.4. Relative standard errors (%) in estimated numbers of 3+ cod estimated from simulated trawl surveys making 10, 20, 40 or 80 hauls. 1000 replicates were simulated for each survey design.

The result of this preliminary simulation suggests that, for this specific area and target species, little improvement in precision is achieved by sampling more than 40 hauls. If this result is validated by further analysis the sampling effort in the Fishermen survey could probably be reduced with minimal increase in the level of uncertainty in the selected index highlighting the potential that these kind of simulation exercises have. However, to fully analyse the effect on the assessment, the whole chain from resampling surveys and commercial catch data to estimation and assessment needs to be simulated. This work has been initiated and will be continued. For surveys where estimates are required for multiple species, which is mostly the case for trawl surveys, more complex extensions to the simulation process are needed to develop better optimized and cost-effective survey and sampling designs.

Further development

As the code produced in this case study is based on DATRAS exchange format, the code can be used for many other surveys as well. We therefore recommend that ICES hosts the code (through the ICES repository GitHub or equivalent) and thereby make it available to the wider audience. We also propose that the Data Centre and the Data and Information Group (DIG) guide the WK in interfacing the proposed tools to the ICES databases. This should include a script to download (or loop over) several surveys and provide a framework for running the case study script produced by the WK. As the code produced in this case study is only the first step in a more advanced cost–benefit framework, an important aspect is to allow survey scientist as well as other users to write their own diagnostics tools, and, the ICES community will over time crowd source the further development. The ambition is that this will become a standard tool that all the survey groups will use to evaluate the results from the surveys conducted in their group and that the code is further developed so that it can be used during the benchmark process to evaluate more closely the data quality, survey design and associated estimators, and if data are fit for purpose.

4.3 Case Study 3: The Norwegian REDUS (ReDuced Uncertainty in Stock assessment) Project

Summary

There are unknown uncertainties surrounding stock assessment and the whole fishery advisory process. The lack of knowledge of these uncertainties is costly in terms of suboptimal management, monitoring and advice. ICES, the Norwegian Government, the Norwegian Research council and others have therefore identified more precise and accurate stock assessment as a key development challenge within fisheries. Recent model and data infrastructure developments make it realistic to quantify uncertainty from observations through stock assessment to advice. The REDUS project will provide a seamless and generic framework for uncertainty estimation and analysis that will allow for more optimal fisheries management (e.g. potential higher long-term quotas) and better prioritization of fisheries monitoring and research. Coupling measures of sampling and observational uncertainty in input data with statistical stock assessment models provides the basis for (1) optimizing monitoring programs, and (2) for quota advice and long-term harvest strategies that takes uncertainty into account. REDUS will build a generic Open Access toolbox founded on the StoX package and the S2D format that has a potential universal application in renewable marine resource management.

The main aim of REDUS is achieving reduced uncertainty in stock assessment and advice for Norway's most important fish stocks. The objective is to develop and implement the ability to quantify and communicate the trade-offs and risks caused by varying levels of uncertainty of stock assessment and management advice from: i) observations, ii) stock assessment modelling, iii) management strategy evaluation (including harvest control rules), and iv) real-world implementation in practical fisheries management. The REDUS project will use both a bottom–up approach (from observations through management) and a top–down (society's uncertainty requirements implications for observations). Thus, REDUS will provide society with knowledge of how uncertainty affects stock assessment and hence quota advice.

Relevance of the REDUS project

Sustainable and precautionary management is widely acknowledged as a central pillar of ecosystem-based fisheries management (EBFM, see Pikitch *et al*., 2004; Link, 2010). Understanding how different human factors and actions affect the ecosystem is a key to sustainable fisheries management, but without knowledge of the uncertainty associated with each factor alone, or the compounded total uncertainty, it is fruitless to compare different management options. Perceived differences might be statistically insignificant, or the effect of changes in fishing, monitoring or management might be impossible to evaluate, all because the uncertainties are too high. The sources of uncertainty in fisheries advice and management are complex and compounded, but it is essential to be able to estimate the magnitude of this uncertainty and how it affects the various steps in the advisory process, leading up to management decisions: observation of stock status, stock assessment, stock projection, and stock advice. Once this uncertainty can be pinpointed and quantified it will be possible to estimate the total uncertainty associated with a particular advice, discuss this with stakeholders and decision-makers to reach a common understanding of the management implications of the given levels of uncertainty leading specifically to how much precaution must be included in the quota setting to avoid detrimental stock or ecosystem impacts. Conversely, a comprehensive understanding of the uncertainty allows society to *a priori* set thresholds for uncertainty of advice, and from this managers and scientists can infer which monitoring and assessment efforts and tools are necessary to reach the target set by society. Understanding uncertainty will make the fisheries management process more transparent and understandable for stakeholders and experts alike, identifying and prioritizing areas for improvement that will lead to direct benefits for the fishing industry in terms of more predictable stock advice. Therefore, both the Norwegian Government and the Norwegian Research Council place high priority on improving the methods for fish stock assessment (see: "Masterplan for marin forskning" and "Marine Ressurser og Miljø - MARINFORSK. Programplan 2016–2025").

Background and state-of-the-art

The premise for reducing uncertainty in stock assessments is that sources of errors in input-data and modelling errors can be quantified. Stock assessments that form the basis for scientific advice for many commercially important stocks in Europe often provide point estimates of stock parameters (e.g. SSB, F), with no measures of uncertainty, and annual catch recommendations that consist of a single number for each option given (Dankel *et al.*, 2016). Since stock assessment outputs and predictions used for quota advice always will be subject to different types of uncertainty, it is important to (i) quantify sources of errors in input-data, observation methods and sampling regimes and how they propagate to stock assessment and advice, (ii) test the robustness of the quota advice, harvest control rules and management strategies to different types of uncertainty and (iii) have a clear and transparent dialogue with the resource users and decision-makers to communicate uncertainties and risks. ICES is asking for the next generation of assessment models that can incorporate uncertainty in complex input data from multistage sampling surveys and quantify how errors propagates to stock assessment and advice (e.g. WCSAM 2013, World Conference on Stock Assessment Methods for Sustainable Fisheries).

Analytical stock assessments in ICES are based on data from fisheries-independent as well as fisheries-dependent sampling surveys, with inherent uncertainty due to sampling errors and various sources of bias. Historically, yearly point estimates of abundance-indices and catch by age class have been used as input-data to VPA type of models. The uncertainty in estimated catch-at-age has generally been ignored and errors in input data have been assigned solely to abundance indices in various methods for tuning the VPA (e.g. Shepherd, 1999). Publications by Gudmundsson (1994), Quinn and Deriso (1999), Aanes *et al*. (2007), Gudmundsson and Gunnlaugsson (2012), and Nielsen and Berg (2014) (and references therein) provide alternative statistical assessment models that can provide measures of uncertainty in estimated stock-parameters. To account for spatial variability of demographic rates and population variables Thorson *et al*. (2015) has developed a delay difference model that helps explain large portions of parameter variance and hence reduce model uncertainty. In REDUS we will focus on the further development and parameterization of statistical assessment models that can integrate data with varying accuracy (bias and precision) from multiple sources.

Scientific problems, hypotheses and research approach

Accuracy of input-data to stock assessments

Time-series derived from combining biological sampling from commercial fisheries with official surveys and scientific abundance surveys are critical to stock assessments and quota advice. Such long-term monitoring is costly, and it is therefore crucial to employ cost-effective survey designs and efficient estimators to minimize errors (i.e. maximize precision and minimize bias). The overall aim for a design-based sampling strategy (e.g. Særndal, 1978; Gregoire, 1998) is to:

- Collect data in a way that accuracy (bias and precision) can be reliably assessed at national and regional level;
- Ensure that sampling intensity is allocated in a way that would minimize bias and maximize precision at the level where it matters most in the assessment of stocks and fisheries.

We follow Jessen and Jessen (1978) and use the term accuracy as a measure of the proximity of an estimate to the true value (i.e. high accuracy signifies high precision and low bias). It is useful to evaluate the accuracy of input-data to stock assessments within the framework of "total survey design," which is defined as the attempt to control the total error in the estimates derived from survey data (Lessler and Kalsbeek, 1992). In sampling theory, the total error is generally divided into variable errors and bias (e.g. Cochran, 1977). Bias refers to systematic errors that cause the average survey value to deviate from the true population value under a specific survey design. Bias can be reduced by employing unbiased statistical estimators, improving the survey designs and coverage, and by reducing measurement errors. In this project we will try to control the total error in input-data to stock assessments, and choose the combination of sampling design, measurement procedure, and estimators that will minimize the total errors within the resources available for monitoring. In the literature, this has been referred to as total survey design (e.g. Lessler and Kalsbeek, 1992).

It is useful to separate sources of errors in input-data to stock assessments into three broad categories: (1) Sampling errors, (2) Coverage errors, and (3) Measurement errors.

Sampling error reflects the degree to which a survey statistic (e.g. estimated number of fish per age class in the total yearly catch of a species) differs from the "true" value due to the fact that one particular survey only realized one of near infinitely many possible survey samples. Small sampling errors for a survey statistic signify high precision; i.e. the estimate from one sample is close to the average over repeated samples (Jessen and Jessen, 1978).

Coverage error is the degree to which statistics such as estimated abundance indices in number of fish per age class, based on yearly scientific trawl surveys, are off due to the fact that the available biological samples from the trawls do not properly represent the entire target population.

Measurement error is the degree to which a survey statistic differs from the targeted population value due to imperfections in the way the data from each sample is collected. In scientific abundance surveys, sampling is typically conducted using trawls, plankton nets, and acoustic methods.

REDUS will develop analytical methods for an objective framework to evaluate the benefits vs. costs of datasets used to support stock assessment and fishery management advice, where the benefits are in terms of accuracy (bias and precision) of assessment results and derived management variables, and risks to stocks associated with management under uncertainty. This framework will be used to evaluate existing datasets, new data requests from end-users, survey design, spatial and temporal coverage, and cost-effective allocation of sampling effort towards the components of data collection that have the greatest influence on quality of assessments and management decisions for particular stocks or groups of stocks.

Management Strategy Evaluation (MSE)

A key goal of the REDUS project is to answer the critical questions concerning the utility of knowledge of uncertainty in data (both fisheries and survey) for long-term stock management. Given the precautionary principle one would expect that reducing and understanding such uncertainties would allow for higher long-term yield while retaining a low probability of overfishing the stock. All of these questions are suited for investigation via Management Strategy Evaluations (MSEs), given a sufficiently flexible software tool. A Management Strategy Evaluation combines an "operating model" (simulating the real world) with an "assessment model" to simulate the assessment and advice giving cycle as accurately as possible. "Data" on surveys and catches are taken from the operating model, and have errors applied. These are then used to tune an assessment model, which gives an estimate of stock size. This estimate is then applied to the HCR in order to produce a quota, which in turn is input into the operating model (possibly with implementation errors), and the cycle repeated. Such a system allows for HCRs to be evaluated in a realistic setting, but also allows for the performance of the assessment model and the impact of various sources of error to be investigated. In many cases around the world today a slightly simplified version of this procedure is used, with the "assessment model" replaced by exact knowledge of the stock (to which errors can be added). Currently IMR uses this procedure, employing a tool called "PROST" to evaluate harvest control rules (HCRs), but this tool does not offer full MSE (PROST does not include an assessment model). Such simplified procedures are simpler to develop than "full MSE simulations", and are able to evaluate HCRs provided that the assessment model in use is simple. However they apply errors to output of the assessment model, not to the inputs (e.g. survey indices, fisheries data, age determination), and hence are not suited to evaluating more complex assessment models or evaluating how such models behave when given knowledge of uncertainties in different input datasets. Nor do the tools currently in use allow for multispecies or ecosystem operating models, and therefore cannot incorporate uncertainties that arise from multispecies interactions. Consequently, there is a need at IMR to develop such a tool. A particular challenge is to allow for complex correlated input-data resulting from multistage sampling. This will be addressed in close cooperation with WP1, where variance-covariance matrices for catch-at-age and abundance indices at age will be estimated and included in the statistical assessment modelling.

Data infrastructure

The Sea2Data project has developed an application interface (API) to the main data sources for the project. This includes fisheries independent data from scientific surveys, including interpreted acoustic nautical area scattering coefficients (NMDechosounder) and biological data from trawl samples (NMDbiotic). ICES has adopted similar structure for trawl samples (DATRAS) and acoustic data (AtlantOS project). A thin software client on top of the API can be accessed through http://tomcat7.imr.no:8080/DatasetExplorer/v1/html/main.html. Fisheries-dependent survey sample data are currently being uploaded to NMDbiotic, and will be available through the course of the project. Landing data are currently available via flat text files, updated every second month, but there is a proposal to include this in the NMD API to streamline data access. Intercatch is a web-based system where data providers upload aggregate estimates of catch-at-age from their countries that are then combined by stock coordinators to represent the total catches of a stock. The aggregated output files can then be downloaded to be used as input to ICES stock assessments. In general these estimates are not provided with associated variances, and since they are aggregated, the variances cannot be estimated for total catch-at-age.

The Sea2Data project has developed the StoX software and R-libraries (R-StoX and R-ECA) for calculating statistics based on data from the surveys, and these are interfaced with the NMD API described above. These libraries are released under GPL and fully versioned under SVN. The software packages are being adjusted and implemented in ICES under the AtlantOS project. The output from the StoX software are used as input to the stock assessment models (NSS Herring). The framework will extended to other modules like the assessment models and HCR offering an efficient infrastructure for the whole REDUS data and estimation processing pipeline.

Test case species

The REDUS analytical framework will be generic, in essence being applicable to all species the IMR gives advice on. For development two test case stocks, the Norwegian spring-spawning (NSS) herring and the Northeast Arctic (NEA) cod, will be in focus, but REDUS methods will be applied to other stocks as soon as they have been tested and verified. Cod and herring are chosen because of the extensive and long time-series of data that exist from catch and scientific surveys. Also, these two species form the basis for the two most valuable Norwegian fisheries. Finally, the NSS herring stock assessment has been under intense public scrutiny and debate, while the NEA cod stock is currently declining from a record high abundance. Therefore, both these species are in need for improved stock assessment and advisory processes where uncertainty should be dealt with explicitly and openly.

Norwegian spring-spawning herring

Norwegian spring-spawning herring (*Clupea harengus*) constitutes the largest herring stock in the world and supports a highly valuable fishery. The NSS herring stock is assessed annually using a virtual population analysis (VPA) type model applied to estimates of fishery catch-at-age data and fishery-independent indices from research surveys for calibration (‘tuning’). The NSS herring stock is characterized by occasionally large year classes that dominate the fishery for many years. Combined with its wide-ranging pelagic behaviour and high-biomass, makes the stock assessment particularly prone to uncertainties in the data. As such it represents a perfect case study for the REDUS project.

northeast Arctic cod

The northeast Arctic cod (*Gadus morhua*) supports the largest cod fishery in the world and is the commercially most valuable demersal fish species globally. The NEA cod stock is well studied with long high-quality time-series of catch and scientific survey data, which are shared between Norway and Russia that jointly manage the stock. The stock abundance and biomass are assessed yearly by the ICES Arctic Fisheries Working group using a virtual population analysis (VPA) type model applied to estimates of fishery catch-at-age data and fishery-independent indices from research surveys for calibration (‘tuning’). Cod is also a key predator in the ecosystem (including cannibalism on young cod). Multispecies interactions are critical in managing this stock. The interaction between uncertainties in the cod data and the interacting species are different from uncertainties for NSS herring, and the two together cover a wide range of uncertainties on which to test REDUS methods.

Coupled species (ongoing activities relevant to REDUS)

Although NEA cod and NSS herring will constitute the test species, the REDUS project will keep abreast and cooperate with other ongoing stock assessment developments at the IMR. In particular REDUS will establish cooperation with the capelin assessment model development carried out by Sam Subbey and Hiroko Kato Solvang, and the planned saithe catch sampling and stock assessment project being developed by IMR in accordance with best scientific practice from ICES WKPICS 2013 and the EU FishPi project.

For more detailed information we refer to the REDUS website: http://www.redus.no/

4.3.1 The Norwegian Spring-spawning Herring case study under the REDUS project.

Background

A main objective of WKCOSTBEN is to assess the trade-offs between sampling efforts distributed to different survey sampling programs (fisheries-independent surveys and fisheries-dependent surveys) with respect to precision in stock assessment outputs (e.g. SSB, F) used in quota advice. One challenge in general in ICES is that there are relatively few cases available where sampling variances are being estimated for the different data sources (e.g. abundance-indices and catch in numbers-at-age) that are key inputs to stock assessments simultaneously, and where there is a statistical modelling framework in place to evaluate these trade-offs. The newly started four-year REDUS project at IMR offers one such case, and we propose to use this also as a case study for WKCOSTBEN.

The REDUS project has identified the NSS Herring as one of three focus stocks. The quotas for Norwegian spring-spawning (NSS) herring (*Clupea harengus*) are shared among Norway, Iceland, Russia, Faroe Islands, and EU. The Norwegian fishing fleet landed 254 658 tonnes out of a total quota of 419 000 tonnes in 2014, and the total quota for 2015 was reduced to 283 013 tonnes, with 172 638 tonnes allocated to Norway. Until 2015, the International Council for the Exploration of the Sea (ICES) has assessed the NSS herring stock annually using Virtual Population Analysis (VPA), and has reported estimates of fishing mortality (F) and SSB without any measures of precision. The most important input data to the stock assessments comes from the ICES coordinated international ecosystem survey in the Nordic Seas (IESNS) and biological sampling of catches from the commercial fishery. Time-series derived from biological sampling of commercial catches and scientific abundance surveys are critical to stock assessments and quota advice. Such long-term monitoring is costly, and it is therefore crucial to employ cost-effective survey designs and efficient estimators to minimize errors. In this case study under the REDUS project we use sampling theory and the statistical assessment model chosen by ICES WKPELA 2016 to quantify how precision in estimates of SSB and F estimates depends on the precision of input data on catch and relative abundance indices by age (cohorts) from monitoring programs. Using empirical self-sampling data from Norway we estimate how many catch samples and otoliths per catch that are required across vessels and fishing operations to achieve sufficient estimates of SSB and F for the stock.

Sampling errors in input-data to stock assessment

Sampling errors for Norwegian catch-at-age for the years 2008–2015 is estimated using ECA (Salthaug and Aanes, 2015, Hirst *et al.*, 2012). The fisheries independent indices with associated measures of uncertainty will be estimated using the StoX software, with estimators of abundance indices by age based on Stenevik *et al.* (2016). The estimated abundance indices and catch in numbers-at-age with associated variance-covariance matrices will provide input to the XSAM model (Aanes, 2016 a, b, c) which since 2016 is used to estimate spawning-stock biomass (SSB) and fishing mortality (F) that are basis for quota advice. Example summary of sampling errors in input data for stock assessment is given in Figure 4.5.

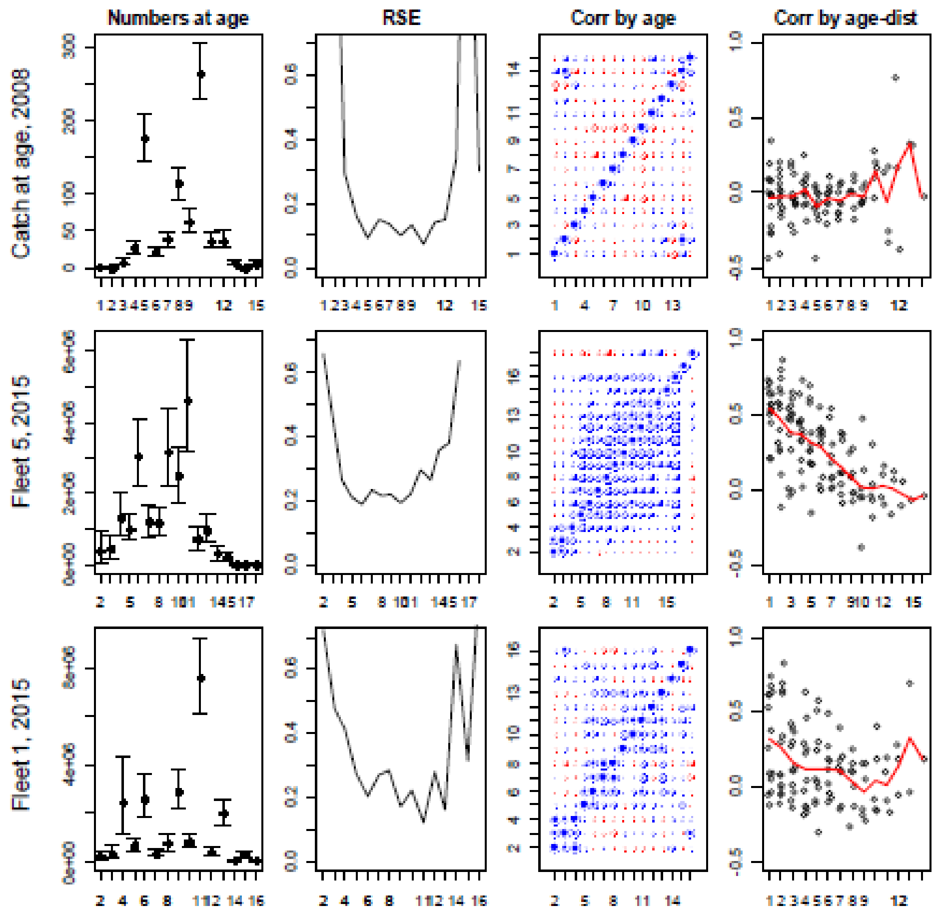

Figure 4.5. Example of estimates of errors in input data to stock assessment: Norwegian catch-at-age (2008), abundance indices at age from the May survey of Norwegian Spring-spawning (NSS) Herring (Fleet 5) in 2015, and Norwegian acoustic survey on spawning grounds in 2015 for NSS herring.

Statistical assessment model

The statistical assessment model XSAM (Aanes, 2016 a, b, c) is used for Norwegian Spring-spawning Herring from 2016. XSAM is accounting for errors in input data from catch sampling programs and acoustic-trawl surveys (Figure 4.5). The model is based on a state space model and structural time-series models for fish stock assessment (inspired by Gudmundsson, 1994). It includes other statistical assessment models including the DTU Aqua SAM model (Nielsen and Berg, 2014) as special cases, and can utilize the sampling distributions derived from analysis of sample survey data by giving appropriate weights to data points. XSAM is coded in TMB (R library) which is efficient for parameter estimation for non-linear models with latent variables. Figure 4.6 is an example run of XSAM for Norwegian Spring-spawning Herring, showing assessment results with measures of uncertainty that takes into account uncertainty in input data.

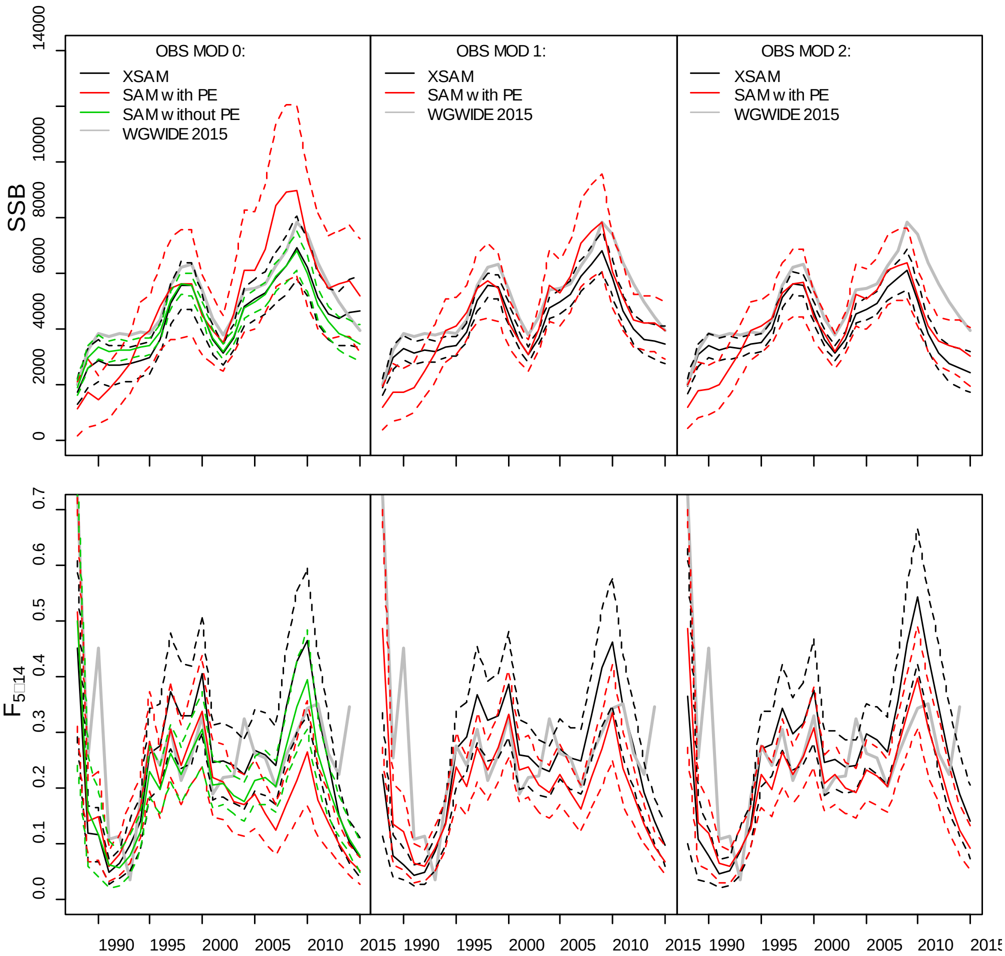

Figure 4 6. Estimates of spawning-stock biomass (top row) and fishing mortality (bottom row) for 3 different formulations of the observation model in XSAM: assuming iid errors and constant variance across ages and time for each data source (OBS MOD 0), using variance structure from estimates of the data sources (OBS MOD 1), and adding estimated correlation structure for the observations (OBS MOD 2) for the XSAM model (black), the SAM model with process error (red) and the SAM model without process error (green). The estimates from WGWIDE 2015 are included for comparison (grey line). The data used for fitting XSAM and SAM are restricted to catch data and fleet 5 (ages 3–15). For OBS MOD 1 and 2, the results from SAM (Nielsen and Berg, 2014) without process error are nearly identical with XSAM and therefore not shown. Broken lines are approximate 95% confidence intervals.

The REDUS project will develop a simulation framework that can be used to evaluate the effect of changing sampling efforts between the different surveys and the effect on the stock assessment model results. This framework will also be used to evaluate effects of various sources of bias in the abundance indices from acoustic surveys.

4.4 Case Study 4: Belgian commercial at-sea sampling

This case study is identified for inclusion in future WKCOSTBEN analyses – a description of the fishery and sampling is given along with the cost–benefit questions to be addressed.

The beam trawl fishery is the most important commercial fishery for Belgium. It is a demersal fishery targeting sole and plaice, with bycatches of other commercially interesting species. The vessels use two (starboard and port) bottom-trawl nets, each attached to a steel beam fixing the net opening. The fleet is rather small (approx. 80 vessels) and operates in the North Sea, the English Channel, the Irish Sea, the Celtic Sea, South of Ireland, and the inner part of the Bay of Biscay. The beam trawl vessels, active in the Southern North Sea and the Eastern English Channel comprise of vessels with a maximum power of 221 kW (coastal fleet segments and "euro cutters") and vessels with a capacity of more than 221 kW.

Objective: estimate Numbers at Length and Ages at Length for the discarded and landed fraction of a selected list of species.

In 2015, the 'Statistically Sound' sampling scheme (4S, design-based sampling) was applied, stratified by fleet segment (large and small fleet segment of the TBB_DEF métier). This strategy allows to randomly select, with known equal probability, individual vessels from a population. The refusal rate of the vessels was documented. Up until now, several changes to the initial design were implemented due to logistic issues. The pool of vessels has been steadily decreasing and for the small fleet segment proved too small to ensure random selection. Therefore, the sampling strategy was adapted and two vessels (euro cutters or coastal vessels) were sampled ad hoc every quarter.

In most cases, a trip (PSU) covers more than one division. Every second haul (SSU) is sampled by an observer so sampling takes place around the clock to reflect typical working conditions. The observer sorts all the discarded species of commercial importance and determines the total weight in a haul for each species. For a selected set of species (14 species), the observer also takes length measurements. Usually, the length of all individual fish in the discarded part of the tow is measured. Only when a species is extremely abundant, a smaller representative subsample is measured. The ratio of the total weight and the subsample weight is used to estimate the total number of discards per cm-size class per species in the sampled tow. The retained part of the catch is treated in the same way as the discarded part of the catch. Non-commercial species are not sampled.

In each trip, otoliths from 3–5 (species dependent) fish per cm-size class per species per area, are collected for age estimations. Otoliths are taken throughout the whole trip (several hauls) until the quota of otoliths is achieved.

What to investigate: Can the Numbers at Length and Ages at length be optimized? E.g. in 2015 the ILVO observers measured 12 600 discarded fish and 6500 landed fish (stock: plaice in 7f and 7g). 540 individuals of the discarded fraction and 410 individuals of the landed fraction were aged. Can this sampling effort be reduced without information loss (quality assurance)? The time and money gained could be used to sample more species (e.g. invertebrates).

5 ToR(c) Develop a proposal for a longer term (3-year) project

PGDATA (ICES, 2016b) decided that a three-year EU project was premature until sufficient groundwork was completed within the ICES system, and has proposed a 3-year series of WKCOSTBEN workshops. This does not rule out future project proposals.

6 ToR(d) Identify the need for follow-up workshops in 2017

It is recommended that the WKCOSTBEN2 be held in 2018 or 2019 to allow time for the further development of computer software in R for simulations studies. The Terms of reference will be draftet during 2017 after the WKCOSTBIO in collaboration with PGDATA.

7 References

Aanes, S., Engen, S, Sæther, B.-E. and Aanes, R. 2007. Estimation of fish stock dynamical parameters from catch-at-age data and indices of abundance: can natural and fishing mortality be separated? Canadian journal of fisheries and aquatic sciences, 64: 1130-1142.

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. ICES J. Mar. Sci. 60: 297–303.

Aanes, S., and Vølstad J.H. 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. Canadian journal of fisheries and aquatic sciences 72 (6): 938-953.

Aanes, S. 2016a. A statistical model for estimating fish stock parameters accounting for errors in data: Applications to data for Norwegian Spring-spawning herring. WD4 in ICES (2016).

Aanes, S. 2016b. Diagnostics of models fits by XSAM to herring data. WD12 in ICES (2016).

Aanes, S. 2016c. Forecasting stock parameters of Norwegian spring spawning herring using XSAM. WD in ICES (2016).

Bjørkvoll, E., Grøtan, V., Aanes S., Sæther, B.-E., Engen, S. and Aanes, R. 2012. Stochastic Population Dynamics and Life-History Variation in Marine Fish Species. The American Naturalist, Vol. 180, No. 3 (September 2012): 372-387.

Cochran, W. G. 1977. Sampling techniques. 3rd ed. John Wiley & Sons.

Dankel, D. J., Vølstad, J. H., Aanes, S. (2015). Communicating uncertainty in quota advice: a case for confidence interval harvest control rules (CI-HCRs) for fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 73(2): 309-317.

Fenton, F. 1960. The sum of log-normal probability distributions in scatter transmission systems. IRE Transactions on Communication Systems, 8:57–67.

Gregoire, T. G. (1998). Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research, 28(10), 1429-1447.

Gudmundsson, G., 1994. Time series analysis of catch-at-age observations. Journal of the Royal Statistical Society: Series C (Applied Statistics) 43, 117–126.

Gudmundsson, G., Gunnlaugsson, T., 2012. Selection and estimation of sequentialcatch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences 69,1760.

Hirst, D., Storvik, G., Rognebakke, H., Aldrin, M., Aanes, S., and Volstad, J.H. 2012. A Bayesian modelling framework for the estimation of catch-at-age of commercially harvested fish species. Can. J. Fish. Aquat. Sci. 69(12): 2064– 2076.

Hrafnkelsson, B., and Stefansson, G. 2004. A model for categorical length data from groundfish surveys. Can. J. Fish. Aquat. Sci. 61(7): 1135–1142.

ICES 2008. Report of the Workshop on Methods to Evaluate and Estimate the Accuracy of Fisheries Data used for Assessment (WKACCU). ICES CM 2008/ACOM:32, 41 pp.

ICES. 2014. Report of the third Workshop on Practical Implementation of Statistical Sound Catch Sampling Programmes, 19–22 November 2013, ICES HQ, Copenhagen, Denmark. ICES CM2013/ACOM:54. 109 pp.

ICES. 2016a. Report of the Working Group for the Celtic Seas Ecoregion (WGCSE), 12–21 May 2015, Copenhagen, Denmark. ICES CM 2015/ACOM:12. 1432 pp.

ICES. 2016b. Report of the Planning Group on Data Needs for Assessments and Advice (PGDATA), 30 June–3 July 2015, Lysekil, Sweden. Eirini Glyki. 103 pp.

ICES. 2016c. Report of the Benchmark Workshop on Pelagic stocks (WKPELA), 29 February–4 March 2016, ICES Headquarters, Copenhagen, Denmark. ICES CM 2016/ACOM:34. 106pp.

ICES. 2016d. Report of the Working Group on Widely Distributed Stocks (WGWIDE), 31 August-6 September 2016, ICES HQ, Copenhagen, Denmark. ICES CM 2016/ACOM:16. 500pp.

Jessen, R. J., Jessen, R. J. 1978. Statistical survey techniques. New York: Wiley.

Kristensen, K. 2014. TMB: General random effect model builder tool inspired by ADMB. R package version 1.1.

Lehtonen, R., and Pahkinen, E. 2004. Practical methods for design and analysis of complex surveys. John Wiley & Sons.

Lesser, V. M., Kalsbeek, W. D. 1992. Non-sampling error in surveys. John Wiley.

Link, J. (2010). Ecosystem-based fisheries management: confronting tradeoffs. Cambridge University Press.

Nielson, G.A. 2014. Cluster Sampling: a pervasive, yet little recognized survey design in fisheries research. Trans. Am. Fish. Soc. 143: 926–938.

Nielsen, A., and Berg, C.W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158: 96–101.

Pikitch, E., Santora, C., Babcock, E. A., Bakun, A., Bonfil, R., Conover, D. O., Dayton P. *et al.*, (2004). Ecosystem-based fishery management. Science, 305(5682), 346-347.

Quinn, T. J., Deriso, R. B. (1999). Quantitative fish dynamics. Oxford University Press.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola. 2008. Global Sensitivity Analysis - The Primer. John Wiley & Sons. Chichester.

Salthaug, A., Aanes, S. 2015. Estimating the Norwegian catch at age of blue whiting, mackerel, North Sea herring and Norwegian spring-spawning herring with the ECA model. Working document in the Report of the working group on widely distributed stocks ICES (WGWIDE). ICES CM 2015/ACOM:15.

Scientific, Technical and Economic Committee for Fisheries (STECF) –The 2013 Annual Economic Report on the EU Fishing Fleet (STECF-13-15). 2013. Publications Office of the European Union, Luxembourg, EUR 26158 EN, JRC 84745, 302 pp.

Shepherd, J. G. (1999). Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. ICES Journal of Marine Science: Journal du Conseil, 56(5), 584-591.

Stenevik, E. K., Vølstad, J. H., Høines, Å., Aanes, S., Óskarsson, G. J., Jacobsen, J. A., Tangen, Ø. (2015). Precision in estimates of density and biomass of Norwegian spring-spawning herring based on acoustic surveys. Marine Biology Research, 11(5), 449-461.

Särndal, C. E., Swensson, B., Wretman, J. (2003). Model assisted survey sampling. Springer Science & Business Media.

Thorson, J. T., Ianelli, J. N., Munch, S. B., Ono, K., Spencer, P. D. (2015). Spatial delay-difference models for estimating spatiotemporal variation in juvenile production and population abundance. Canadian Journal of Fisheries and Aquatic Sciences, 72(12), 1897-1915.

XSAM. Working Document at WGWIDE in ICES (2016).

Annex 1: List of participants

Name	Institute	E-Mail
Jon Helge Vølstad (chair)	Institute of Marine Research	jon.helge.voelstad@imr.no
Jose Rodriguez	Instituto Español de Oceanografía	jose.rodriguez@st.ieo.es
Josefina Teruel	Instituto Español de Oceanografía	josefina.teruel@vi.ieo.es
Josefine Egekvist	DTU Aqua - National Institute of Aquatic Resources	jsv@aqua.dtu.dk
Kirsten Birch Håkansson	DTU Aqua - National Institute of Aquatic Resources	kih@aqua.dtu.dk
Liz Clarke	Marine Scotland Science	e.d.Clarke@marlab.ac.uk
Lucia Zarauz	AZTI-Tecnalia	lzarauz@azti.es
Marie Storr-Paulsen	DTU Aqua - National Institute of Aquatic Resources	msp@aqua.dtu.dk
Mary Christman	University of Florida	marycchristman@gmail.com
Mike Armstrong (chair)	Centre for Environment Fisheries and Aquaculture Science (Cefas)	mike.armstrong@cefas.co.uk
Niels Hintzen	Wageningen IMARES	niels.hintzen@wur.nl
Nils Olav Handegard	Institute of Marine Research	nils.olav.handegard@imr.no
Nuno Prista	Swedish University of Agricultural Sciences	nuno.prista@slu.se
Patrik Börjesson	Swedish University of Agricultural Sciences	patrik.borjesson@slu.se
Sofie Nimmegeers	Institute for Agricultural and Fisheries Research (ILVO)	sofie.nimmegeers@ilvo.vlaanderen.be
Sofie Vandemaele	Institute for Agricultural and Fisheries Research (ILVO)	sofie.vandemaele@ilvo.vlaanderen.be
Cristina Morgado	ICES	Cristina@ices.dk

Annex 2: Agenda

Location: ICES HQ, Copenhagen

Meeting start: 9:00 am. Tuesday 28 June 2016

Meeting close: 13:00 Friday 1 July 2016

Background to agenda:

WKCOSTBEN is the first step in a longer process over several years. The goal is to flesh out the operational details of a possible framework to prioritize allocation of sampling efforts within and among fishery-dependent and fishery-independent sampling schemes for collection of data for assessments ("cost–benefit framework"), and identify future steps in the process. The meeting will be most successful if there is a good mix of experts in design and implementation of data collection from surveys and fisheries, and people who have a good understanding of end use of data and how different types of data impact the quality of assessments and management advice and how this can be evaluated. The meeting will address the Terms of Reference given at the end of this document.

PGDATA 2016 decided that a series of annual workshops would be preferable at this stage for further development of the topic rather than a 3-year study contract, hence the current ToR c is no longer being addressed.

Draft agenda

Tuesday 28 June start 9:00 end 18:00

1. Introductions and Terms of Reference
2. Work plan for meeting
3. Overview of PGDATA 2016 outcomes on cost benefit framework for data collection
4. ICES Theme session on "when is enough, enough" – scope of submitted papers.
5. Interactive plenary session: defining how a cost–benefit framework would be structured and implemented for fishery-dependent and fishery-independent sampling schemes:
    - Objectives – long-term strategic and short-term reactive
    - Who it is targeted at and what they need to know
    - Operational elements of framework, and the information and analysis tools needed
    - Implementation: who – when – how
    - Outputs and process of communication
    - How to embed the framework alongside quality assurance frameworks within the ICES Expert Group and Steering Group structures.
    - Plenary will include some brief presentations to lead in to discussion sessions for each topic.

- Depending on the mix of attendees, some stocks, fisheries and surveys from one or more European sea areas will be identified in advance and used by subgroups as case studies to explore how a cost–benefit framework could be applied. Some specific data types, for example allocation of effort to length and age sampling, where work has already been done internationally, will be used to help develop ideas on the framework. Some data on activities and costs will be sought prior to the meeting.

6. Identify subgroups for Wednesday and Thursday:
    - Survey subgroup
    - Fishery sampling subgroup
7. Develop meeting notes for subgroups to refer to.

Wednesday 29 June start 9:00 end 18:00

<u>Morning</u>:

- Work in subgroups to explore in detail the data currently being collected and used by ICES for assessments for each case study, and how the bullet points in item (4) above on a cost–benefit framework would be applied operationally to research vessel surveys and fishery sampling schemes providing data to ICES. This will cover data collection in general from the case studies, but also with additional focus on some specific data types such as age-length sampling. This will form the basis for the report text, with modifications following plenary discussion, so must be suitably documented with this in mind. Where possible, some example data on sampling activity and costs will be sought prior to the meeting.

<u>Afternoon</u>:

- Subgroup presentations and discussion in Plenary. During this discussion we will put together the proposed components and processes of a cost–benefit framework for fishery-dependent and fishery-independent data for assessments into some form of flowchart / map that will be a key output from the meeting.

Thursday 30 June start 9:00 end 18:00

<u>Morning:</u>

- Continue with Plenary discussion from previous day – finalize the proposed framework elements and processes.
- Future developments: propose Terms of Reference for a WKCOSTBEN 2.
- Chairs and subgroups start drafting report sections.

<u>Afternoon</u>:

- Continue drafting report sections.
- Later in day – review what has been completed so far and any outstanding issues to be discussed.

Friday 1 July start 9:00 end 13:00

Morning:

- Review text written for WK report.
- Allocate responsibilities and deadlines for any text not completed by end of meeting

Close meeting

Annex 3: Recommendations

Recommendation	Adressed to
1. Adapt ICES survey protocols to provide some additional types of information e.g. subsampling rules; age subsampling protocols, relevant to evaluating options for improving cost efficiency	SSGIEOM
2. Consider the results of the preliminary survey made by WKCOSTBEN on costs and benefits of research surveys and establish a roadmap for future developments of the cost/benefit framework in that field	PGDATA; SSGIEOM
3. Plan a follow up WKCOSTBEN2 that focuses on case studies.	PGDATA