WORKSHOP ON UNAVOIDABLE SURVEY EFFORT REDUCTION (WKUSER)

VOLUME 2 | ISSUE 72
Contents

1 Introduction ................................................................................................................................... 1
   1.1 Unexpected Survey Effort Reduction ................................................................................ 1
   1.2 The consequences of survey effort reduction ................................................................. 2
   1.3 Challenges when forced to make changes to survey effort ............................................ 2
   1.4 The need for different strategies ...................................................................................... 2
   1.5 Common needs for scientists and managers ................................................................... 3
   1.6 The need for the workshop ............................................................................................... 3

2 TOR Group Reports........................................................................................................................ 5
   2.1 TOR I. “Current Processes”: The current processes used in dealing with unavoidable change in survey effort and examine the existing coping strategies (e.g. spatial coverage, survey frequency, or sampling density) and their qualitative consequences ................................................................................................................................. 5
   2.1.1 Summary of current processes used in dealing with survey effort reduction .......... 5
   2.1.2 Summary of workshop presentations and common and unique components of dealing with survey effort reductions ................................................................. 9
   2.1.3 Best existing recommendations for dealing with survey effort reductions .......... 10
   2.1.4 What are current processes lacking? What are the major challenges? .......... 12
   2.1.5 Recommended future directions .............................................................................. 14
   2.1.6 References and articles on related topics .................................................................. 17
   2.2 TOR II. “Total Survey Uncertainty”: Develop key quality metrics that can be used to describe “total survey uncertainty” for survey derived indices of abundance for common survey designs ............................................................................... 18
   2.2.1 What is total survey uncertainty? ............................................................................. 18
   2.2.2 Why evaluate total survey uncertainty? .................................................................... 20
   2.2.3 How can we evaluate total survey uncertainty? ...................................................... 21
   2.2.4 WKUSER recommended actions based on available literature and presentations ...... 23
   2.2.5 Major gaps in existing information and future research priorities ......................... 24
   2.2.6 Roadmap for improving quantification of total survey uncertainty ....................... 25
   2.2.7 Near-term vs. long term priorities for improving quantification of total survey uncertainty ....................................................................................................................... 27
   2.2.8 Outstanding questions and caveats .......................................................................... 28
   2.2.9 WKUSER recommended actions to estimate total survey uncertainty ............... 28
   2.2.10 References ............................................................................................................... 29
   2.3 TOR III. “Survey Continuity”: Define changes to survey designs that require inter-survey calibration and what changes can be resolved by a model-based approach to index generation ............................................................................................................. 30
   2.3.1 Summary of existing examples of changes to survey design that required inter-survey calibrations ........................................................................................................... 30
   2.3.2 Summary of workshop presentations ....................................................................... 31
   2.3.3 Which survey data products are needed to be included in the inter-survey calibration? .......................................................................................................................... 32
   2.3.4 Best existing recommendations ............................................................................... 33
   2.3.5 Path forward .............................................................................................................. 33
   2.3.6 Recommended future directions. Collaborative research and tool development projects .......................................................................................................................... 35
   2.3.7 Summary of discussions ........................................................................................... 35
   2.3.8 References ................................................................................................................ 36
Executive summary

The Workshop on Unavoidable Survey Effort Reduction (WKUSER) reviewed available research, evaluated current practices, and recommended future directions on four key topics: Existing Decision Making Processes, Survey Uncertainty, Index Continuity, and Trade off Evaluation Tools.

Resource surveys are conducted worldwide to measure population trends of economically important marine species and characterize the state of marine ecosystems. Resulting information is used for stock assessments and management recommendations that contribute to sustainable fisheries and ecosystem management. Surveys are expensive and complex to execute and are vulnerable to unexpected reductions in effort expended due to funding shortfalls, vessel unavailability, weather, and other complications that require immediate or strategic actions.

Decision trees and tables were developed under each key topic largely corresponding to different groups of monitoring program actors, and an overarching concept of how the elements need to interact on different time-scales is developed to assist survey managers in decision-making in a variety of conditions and objectives on various time horizons. Together the trees can deliver best practice decision tools and provide assessments of the impact of survey effort reductions on data and advice quality through a series of questions linked to information tables.

Recommendations for best practices and future refinements of process are:

- Monitoring agencies are encouraged to routinely apply the developed approach to conduct survey evaluations. This will facilitate appropriate prioritization of monitoring tasks by examining its relation to objectives by exploring possible methods for gains in survey efficiencies (such as: reducing the number of biological samples, shortening tow duration, increase in catch subsampling while also considering station thinning, excluding areas, reducing survey frequency, or changing survey design).
- Continue further studies on estimation of total survey uncertainty by conducting research into the various subcomponents inherent in survey design and metric calculations, which include sampling design, sampling efficiency, spatial availability, density-dependence, vessel effects, timing, and environmental conditions. The interactions of these uncertainty components require studies to assess total survey uncertainty for appropriate weighting in likelihood-based assessments, provide greater insight into the impact of certain changes, and provide a long term strategy for improved surveys.
- Develop and expand simulation studies and research on model-based capabilities that can be used to define methods for survey effort reduction, aid in estimations of total survey uncertainty, and help with inter-calibration studies.
- Survey groups and assessment groups together should develop quantitative applications that can be used for any survey and assessment combination to determine the impacts of different monitoring strategies in terms of inputs (cost) and outputs (uncertainty). They should include functions to process abundance data, and to incorporate ecosystem data for use in model-based estimation and in process studies, multispecies/multi-objective optimization, and evaluation of trade-offs between different survey and estimation approaches.

Survey managers are recommended to intensify preparation for response to ecosystem changes, which are already underway in many areas. These preparations should include strategies for survey expansions into new areas (or reductions on other areas) to assure continued relevance of survey information to fisheries management and research.
## ii Expert group information

<table>
<thead>
<tr>
<th>Expert group name</th>
<th>Workshop on Unexpected Survey Effort Reduction (WKUSER)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert group cycle</td>
<td>Annual</td>
</tr>
<tr>
<td>Year cycle started</td>
<td>2020</td>
</tr>
<tr>
<td>Reporting year in cycle</td>
<td>1/1</td>
</tr>
<tr>
<td>Chair(s)</td>
<td>Dr. Stan Kotwicki, United States</td>
</tr>
<tr>
<td></td>
<td>Dr. Sven Kupschus, Great Britain</td>
</tr>
<tr>
<td></td>
<td>Wayne Palsson, United States</td>
</tr>
<tr>
<td>Meeting venue(s) and dates</td>
<td>13-17 January 2020, Seattle, WA, United States, (45 Participants)</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Unexpected Survey Effort Reduction

Regional surveys of marine resources are conducted in many parts of the world in order to provide fishery-independent indices and other measures of abundance and to characterize ecosystems. Surveys are conducted from research or chartered commercial fishing vessels and aim to adhere to standardized and repeatable scientific protocols. These endeavours inform valuable fisheries and ecosystem assessments, but in themselves can be costly to perform. Survey effort can be hindered or reduced by temporary or permanent funding reductions, vessel breakdowns, extreme weather, and other unexpected conditions. The COPV-19 pandemic started nearly at the same time as the January 2020 meeting and provides a stark example of how severe an unexpected reduction in survey effort can be. These reductions create a quandary to the scientists who design and execute the surveys. They also create challenges in how resulting survey data can be used in stock assessments to provide management advice. These short-term effort reductions can also compromise the long term objectives of having consistent survey time-series estimates for most of the populations in the ecosystem. Many survey programs are at one time or another asked to make substantial short-term changes to survey operations due to budget reductions, weather, and vessel breakdowns and unavailability or because new policy questions arise. Usually the required changes leave little time for planning and quantitative evaluation, so there is a need to develop methods that provide a better general understanding of the risks of different implementation options. There is a need for conducting research and development of methods to help scientists cope with unexpected survey effort reductions in the way that minimises loss of information from reduced surveys.

1.2 The consequences of survey effort reduction

Short- and long term reductions in survey effort can have consequences on many aspects of the information produced from surveys. These consequences affect stock assessments with resultant fisheries management, ecosystem indicators, and fisheries research and can include:

Survey Outputs:
- Increase in uncertainty in or lack of abundance estimates
- Increase in uncertainty of and lack of age and length composition data
- Decreased resolution or bias of spatial distribution data
- Change in survey bias, which could lead to detection of false trends in the population abundance and age composition
- Reduced ability to detect trends in population abundance
- Discontinuity of time-series
- Reduced ability to detect incoming year classes

Stock Assessments:
- Inability or reduced ability to check model fits against fishery-independent survey data. (Considered a good practice of stock assessments)
- Increased uncertainty in model outcomes
- Biased stock assessment outcomes
- Increased risk of overfishing
- Reduced allowable biological catch (ABC)
- Reduced ability to detect distribution shifts, which have been common in recent years
- Reduced ability to detect sudden changes in productivity and the ecosystem
Fisheries:
- Reduced total allowable catch (TAC)
- Loss of revenue and jobs
- Increased risk of overfishing
- Risk of going out of business
- Social hardships

Ecosystem Considerations:
- Loss of data on ecosystem indicators
- Loss of environmental data such as temperature, salinity, light intensity and attenuation
- Loss of non-target fish and invertebrate abundance data
- Loss of food habits information needed for multispecies ecosystem models
- Loss of other biological data

Fisheries Research:
- Loss of data for ecosystem modelling
- Reduced ability to conduct spatial and population dynamics studies
- Reduced ability for prediction of future scenarios due to climate or other factors
- Loss of platforms for novel studies

1.3 Challenges when forced to make changes to survey effort

The potential consequences of unexpected survey effort reduction create challenges in executing surveys:
- How to choose the survey data products to focus on? (e.g. abundance, composition)
- How to reduce the effort? (sampling density, spatial coverage, survey frequency)
- How to measure consequences of change in sampling effort?
- How to assess value or loss of information?
- How to minimise loss of information and optimize survey effort?
- How to assure continuation of the standardized time-series? How to assure constant catchability? Alternatively, how to deal with temporal changes in q (catchability)?
- What is the minimum survey effort required to provide useful information?
- How to weigh environmental vs biological data when deciding on effort changes?
- How to propagate increased uncertainty from survey to stock assessment outputs?

1.4 The need for different strategies

The future likelihood of unavoidable survey effort reductions may mean that agencies and ICES adopt new strategies for coping with survey changes in time of simultaneous changes to the ecosystem. Some considerations include:
- Design-based and model-based indices for coping with changes in surveys and changes in the ecosystem
Strategies employed could minimise these risks through spreading reductions as evenly as possible through the survey effort distribution.

- Increases flexibility in survey implementation but also increased risk of type I and type II errors.
- Not likely to mean the same thing for all resources or survey designs.
- Nor is it going to provide the greatest efficiency as per-sample costs are likely to be higher.

1.5 Common needs for scientists and managers

Given the consequences and challenges survey effort reductions may pose, survey and assessment scientists and managers likely have common needs:

- Methods to better understand the spatial and temporal variability of the “ecological” structure of our ecosystems and predictions for likely change. (The ecosystem drivers are likely to vary between ecosystems, but methods to characterize ecosystems are likely the same).
- Quantitative tools to predict the impact of changes on surveys for different data uses and different survey designs. (The implementations are likely to differ slightly, but the approaches should not).
- Prioritized lists of policy needs that allows survey scientists to work on survey effort optimization (the resources and the policy is likely to mean a difference in the preponderance of certain solutions, but not the range of solutions).

1.6 The need for the workshop

The International Council for the Exploration of the Sea (ICES) along with the National Oceanographic and Atmospheric Administration (NOAA) Fisheries conducted a Workshop on Unavoidable Survey Effort Reduction (WKUSER) in January 2020 that attracted scientists from Europe, Canada, and the United States. The workshop was initiated by the ICES Working Group on Improving Use of Survey Data for Assessment and Advice (WGISDAA) challenging survey and stock assessment scientists to investigate the nature, knowledge, and responses to unavoidable reductions of survey effort.

WKUSER participants examined methods that can minimise the amount of information lost and identified appropriate methods to accommodate the survey design and objectives. The outcome of the workshop identified tools to assist survey and assessment scientists to make better decisions when unexpected events force changes, to facilitate better contingency planning, and to convey the likely consequences to assessment scientists and policy-makers.
The specific topics were undertaken in terms of reference (TOR) groups and considered methods, tools, data gaps, and consequences of:

TOR1 - The “current processes” that address unavoidable reductions in survey effort and that examine the existing coping strategies (e.g. spatial coverage, survey frequency, or sampling density) and their qualitative consequences.

TOR 2 - Develop key quality metrics that can be used to describe “total survey uncertainty” for common survey designs and indices of abundance.

TOR 3 - Define “changes to survey designs” that require inter-survey calibration and what changes can be resolved by a model-based approach to index generation.

TOR 4 - Develop methods that can provide quantitative “decision-making tools” describing impacts on the quality of survey deliverables and advisory products.
2 TOR Group Reports

Oral presentations were made during the first two days of the workshop that introduced the problem and perspectives from North America and Europe and that addressed each of the four TOR objectives. The abstracts of these presentations appear in Section 4. On the third and fourth days of the workshop, attendees broke out discussion groups focused on the four TORs. The following were the results of each TOR group addressing each term of reference.

2.1 TOR I. “Current Processes”: The current processes used in dealing with unavoidable change in survey effort and examine the existing coping strategies (e.g. spatial coverage, survey frequency, or sampling density) and their qualitative consequences

Subgroup leads - Owen Hamel, Gwladys Lambert, Tien-Shui Tsou
Participants - Jennifer Blaine, Ingeborg Deboois, John Field, Ned Laman, Rick Rideout, Paul von Szalay

2.1.1 Summary of current processes used in dealing with survey effort reduction

Reasons that may lead to unavoidable survey effort reductions are numerous. Some are the result of a decreased budget, while others are due to unforeseeable and/or uncontrollable circumstances, such as weather, gear loss, new development of marine areas, regulations, government actions, loss of partner participation, etc. For example, the North Sea International Bottom Trawl Survey (IBTS) lost areas due to wind farms; the U.S. West Coast Bottom Trawl Survey (WCBTS) lost areas due to newly implemented marine protected areas, and 3 weeks of survey time in 2013 due to the federal government shutdown; Aleutian Islands and Gulf of Alaska surveys lost traditional sampling areas due to seal/sea lion haul-out no-transit zones under the U.S. Marine Mammal Protection Act; and the Puget Sound Bottom Trawl Survey lost the shallow water stratum in 2012 and 2013 due to U.S. Endangered Species Act (ESA) protection of juvenile Chinook salmon. How to deal with such reductions depends on the reason, the time available to plan in advance of the changes, and the objective of the survey. Regardless, short of cancelling a survey altogether, there are five general approaches currently used when dealing with survey effort reductions: improving survey efficiency, reducing spatial coverage, reducing station density, reducing survey frequency, and changing survey design.

Current processes

Improving survey efficiency

Tow length - Reducing tow time is a common approach used to reduce survey costs, as not only will the time spent at a site will be shorter, but also and less time will be required to process the subsequent catch and biological samples. Several bottom trawl surveys in different regions, such as the North Sea International Bottom Trawl Survey (IBTS), NOAA US West Coast Bottom Trawl Survey (WCBTS), NOAA Gulf of Alaska Bottom Trawl Survey (GOABTS), Washington Department of Fish and Wildlife (WDFW) Puget Sound Bottom Trawl Survey (WPSBTS) and Canadian
Department of Fisheries and Oceans Bottom Trawl survey (CDOBTS), have already implemented these reductions over the years. The majority of these surveys have reduced tows from an original 30-minute duration to a 15-minute duration or less. For most surveys, the reduction in tow time has been intentional in order to reduce survey costs, standardize tow length, increase spatial coverage, and/or to reduce waste. In some cases, however, it has occurred as a result of uncontrollable circumstances. For example, tow time for the California Current Recruitment and Ecosystem Survey (midwater trawl) had to be reduced from planned 15-minute tows to 5-minute tows in some instances due to high abundance of jellyfish or pyrosomes during some years that would otherwise have resulted in net or gear damage during a full tow. In this case, the reduction in tow length also has clear implications on the data, as towing a full net for longer than necessary will alter the density estimates. Further, it is arguable that, in any circumstances, reducing tow length by, say, a factor of 2 will not simply reduce abundance estimates by a factor of 2, e.g. because the catch efficiency of a gear may change over the time of a tow. Additionally, there will be implications on biodiversity estimates. Whatever the reason to employ shorter tows, it is also important to note that with a shorter distance, tows can be conducted in tighter spaces than longer tows; these tighter spaces might include more complex habitats that would result in different species being caught, or they might result in a larger overall trawlable area that will need to be included in estimate calculations or could change the catchability associated with the survey for many species. For surveys that capture larger, mobile species than can evade the net, the impact of tow duration on capture efficiency should be evaluated.

**Working hours** - Another way to increase efficiency is to work more hours in a day than has been typical in an effort to complete the survey in fewer field days. If going to regular days of more than 12-15 working hours, this approach might require a sufficient amount of survey and vessel crew on board in order to allow work shifts; without shifts, data integrity (as well as crew morale) will likely decline towards the end of the day or as survey progresses. However, this would often imply fishing during the night-time hours, so this approach might not be suitable if the target species is likely affected by daylight or circadian rhythms. The night-time hours might be more suitable to oceanographic sampling; for example, the Clean Safe Seas Environmental Monitoring Programme (CSEMP) surveys by the UK’s Centre for Fisheries and Aquaculture Studies (CEFAS) have looked at oceanographic sampling throughout the night in combination with other ongoing daytime fishing surveys.

**New technology** - Often new technologies can be implemented into catch processing procedures to make them more efficient. One common example in bottom-trawl surveys is the adoption of digital length boards and digital scales. Not only can these two devices greatly reduce the time it takes to process a catch, but they can also reduce the post-survey data processing time and improve the overall data quality; this is particularly true when a computer program is used to link them, as is the case on WCBTS. Another example of technology improving processing efficiency, the WPSBTS began using tablets with self-designed forms (iForms) with which to record biological samples; with this, barcodes, readable by the tablet are developed before the survey and used for all sample labels, thus reducing errors and the time it would normally take to hand-write a sample label.

**Catch processing** - Catch processing efficiency, and thus overall survey efficiency, can also be improved by subsampling catches and/or reducing the number of biological samples taken. Most bottom-trawl surveys already have a protocol of how and when to subsample large catches, which is generally done by weighing a certain number of baskets of mixed catch and processing, or identifying to species, a smaller number of baskets, implying that data must then be extrapolated. When processing time needs to be reduced, the number of fully processed baskets could be reduced and the data extrapolated to a larger degree, if implications in terms of data quality are understood. In addition to weights and counts, many surveys also take length measurements
and biological samples (e.g.: otoliths) from a subset of the processed catch. The amount of these measurements and samples taken can also be reduced in an effort to improve survey efficiency, provided the minimum amount needed by stock assessors or other end-users is still met.

Reducing station density

In this approach, survey scientists aim to cover the full survey area but reduce the number of stations that are sampled. If the survey design is stratified, often sites are dropped in the strata with the highest sampling effort, provided a minimum of two stations per stratum remain. Similar to the reduction in spatial coverage, stations that have been or that are near historically problematic sites (i.e. that might result in failed tows) are usually good candidates for removal. In another example, a list of prime stations may exist, as in a number of UK surveys (e.g. Irish Sea beam trawl survey) which have to be prioritized in case of lack of time. For the WCBTS, survey density has been reduced by 25% or 50% 3 times over 17 years. Twice due to budget and once (early on) due to other concurrent survey efforts. Since a full WCBTS survey consists of four vessels each with a set of stations selected through a stratified random selection process, this was achieved by reducing the survey to 2 or 3 vessels. In a final example, in the Washington State Pacific Herring spawning ground deposition survey, the historical low-use beaches were dropped out of the survey.

Reducing survey frequency

Another approach is to reduce the frequency of surveys. A number of surveys have historically done the opposite when recognizing the need for more sampling. For example, the GOABTS and Aleutian Islands Bottom Trawl Survey (AIBTS) went from triennial to biennial. The WPSBTS, in contrast, may become biennial (from annual) if the budget is cut further in the future. Surveys with different priority species or assemblages can rotate targets, such as yellowfin in one year, rockfish in another. The joint U.S. - Canada integrated acoustic and trawl survey of Pacific hake went from triennial to biennial in 2001. However, an additional survey was conducted in 2012 when the stock was near a low point in population, and catch limits had been limiting in previous years. When the stock is large, fishing limits are not constraining given the potential effort and the need for surveys is lessened. One could see a situation in which surveys could be reduced to triennial at peak abundances and increased to annual at low abundances in order to maximise the usefulness of the information to fisheries management while potentially decreasing overall costs.

Changing survey design

Several surveys in Washington State changed survey design due to major budget reductions. For example, the WPSBTS and Washington coastal nearshore rod-and-reel survey designs were changed from stratified random to fixed-station in the attempt to keep overall spatial coverage while reducing effort. Washington coastal Yelloweye rockfish set-line survey changed from adaptive survey design to fixed-station. The Puget Sound underwater video ROV survey is currently used to monitor recovery of ESA listed rockfish species, Boccaccio as endangered and Yelloweye Rockfish as threatened. Challenges for this survey are to “find a needle in a haystack.” Based on limited budgets, the survey design has changed from a habitat-focused random design, to a stratified-random design with strata defined by habitat suitability modelling vis a vis a maximum entropy modelling approach (MAXENT, Elith et al. 2011).
Cancelling surveys

When no funding is available or when the affordable effort level is too low to be meaningful, surveys have been cancelled. For example, NOAA Eastern Bering Sea Slope Bottom Trawl Survey, AIBTS, Puget Sound acoustic-trawl survey, and Puget Sound dive survey have been cancelled. This approach will lead to complete loss of information in a given year. Although from a data perspective this may feel as a loss, sometimes ‘no data’ provides more clarity on the use of a survey in a time-series than ‘limited data’. It is in the nature of people to use all data available, and when data is too limited the signals may be not representing the field situation.

Consequences

Consequences depend on the value of the information in the survey and the availability and quality of alternative data sources. Here we focus on stock assessments and advice. In this context, consequences depend upon how much the stock assessment relies on the survey index and associated length and age composition data, as well as on other variables considered for advice and management that are related to the wider ecosystem.

Interpretation of changes around the time-series can get challenging and lead to the need for decisions whether to create a new index for the assessment or to develop methods to incorporate changes in a single index (WCBTS example). Modelling tools such as Vector Autoregressive Spatio-Temporal models (VAST) (Thorson 2019) can be used to fill in gaps in spatial, or even temporal, coverage in the dataset, however this may require complex model validation analysis. Spatial modelling also comes with caveats and, for example, spatial variation in recruitment may be more difficult to fill in. Indices and compositional data developed for rarer species or those with a more specific geographical range (e.g. patchy eastern Bering Sea crab population) are likely to be more affected by reductions than the more widespread and abundant species. In terms of information derived from the survey, the CV is generally expected to increase as effort decreases. This in turn can be expected to have a direct impact on assessments which use sampling variance estimates, which is the case of most US stock assessments, while state-space models used mainly in Europe and parts of Canada do not rely on the sampling variance estimate as an input. On the one hand, if model framework and harvest rules remain unchanged, increasing uncertainty may result in a decrease in advised catches. On the other hand, losing data quality may mean changing assessment models/harvest rules (e.g. move down category in ICES framework), fixing more parameters and then reducing estimated uncertainty in the model outputs leading to potentially producing a less precautionary advice.

Advice may rely exclusively on surveys - e.g. some ICES data limited stocks under category 3, such as a number of elasmobranch stocks, are managed based on a ratio of indices over the last 5 years, which a reduction in survey frequency would have a significant impact (see https://www.ices.dk/sites/pub/Publication%20Reports/Advice/2019/2019/Introduction_to_advice_2019.pdf for the general context of ICES advise on stock assessment and categories). In that respect, changes in survey design should be carefully considered.

Where alternative survey or fishery-dependent data sources are available, that may mitigate the impact of survey reductions, especially in cases where it is possible to increase sampling for one or more of those data sources.

The consequences on a multispecies survey will be more challenging to predict than those on a single species survey, although it is worth noting that analyses of changes in biological community structure can be more robust to changes in sampling effort than analyses of changes at the
species level. For example, multivariate analyses of the WPSBTS showed that community structures among regions and depth strata remained consistent over time regardless of survey design changes. But this is very likely case specific, largely depending on the level of design change, sampling gear type and species considered.

A consequence that may be significant is that, in reducing survey effort, information may be lost that informs the ecosystem-based management approach and requirements to report on biodiversity and ecosystem structure. At present, not all information collected on the surveys may be used for management purposes, but the value of these data may not be yet known and is likely to become increasingly important. Along the same lines, reducing effort may lead to having less room and flexibility to address special requests for important research and development projects or to address new emerging priorities. Overall, a consequence of changing survey effort will be to trigger a dialogue between survey teams, assessors, and managers on priorities, whether political or scientific, and on existing research as well as to incentivise some new areas of research.

\section*{2.1.2 Summary of workshop presentations and common and unique components of dealing with survey effort reductions}

While the overview presentations and several other presented papers provided larger scale overviews of North American and European surveys and assessment systems, the majority of talks at the workshop represented case studies developed to address the trade-offs in the accuracy and/or precision of survey indices in the face of various reductions in survey effort, with a subset of these extending to the impacts on assessments. The common focus was on different metrics of increased uncertainty in both indices and stock assessments with declining information from surveys. Approaches varied from resampling existing empirical survey data under a suite of constraints (10 case studies), to simulating distribution and abundance data and subsequent surveys (11 case studies) to complete management strategy evaluation study of the impact of changing survey frequency on stock assessment results and subsequent fishery yields (1 study). In addition to the MSE study, seven of the other empirical or simulation case studies extended the analysis to evaluating the impacts on stock assessment at some level, the others focused on the impact on indices or compositional data that inform assessments. The majority of the case studies were based on bottom-trawl surveys of groundfish, however there were also case studies based on acoustic/midwater trawl surveys, invertebrate surveys, larval surveys and oceanographic indices, and several studies focused on age and length data quality rather than survey effort. Posters provided details on yet other surveys that were not part of case studies. The focus was on Northeast Pacific and Northeast Atlantic waters, specifically most cases concerned the Eastern Bering Sea (11 studies) and Gulf of Alaska (4 studies), with additional examples from the California Current (2), the North Sea (2), the Irish Sea (1) and Puget Sound (1). Highlights from individual studies follow.

The papers presented under the “Terms of Reference 1: Current Processes” section of the Workshop focused on practical and theoretical aspects of coping with unavoidable changes in survey effort. As indicated above, a common conclusion among these studies was that measures of uncertainty increased with decreasing survey effort as did relative bias in a number of cases. Another conclusion shared among these studies, and with the broader scope of presentations in the workshop, was that the type and magnitude of consequences from reduced survey effort are species-specific. There were some unique conclusions as well. Hamel et al. showed that species-specific responses to changing sampling effort could be informed from life history traits and
species distribution (i.e. rarity), and that compositional data quantity can be important in observing changes in status due to large recruitments. Laman et al. used a simulated species distribution developed by Ono et al. (presented in this Workshop) to demonstrate that the anticipated trajectory of an increase in precision and accuracy of biomass estimates for Gulf of Alaska species examined (arrowtooth flounder, Pacific cod, Pacific ocean perch, and sablefish) seemed to reach a maximum at around the 2 boat, 550 station effort level. Lambert showed that as much as a 60% reduction of samples from the fixed station design resulted in a slight increase in uncertainty but little overall change in plaice stock perception suggesting that the state-space model and surplus production models in use were somewhat insensitive to the reduction in survey effort. It was suggested that an MSE approach may lead to different conclusions. The presentation also provided insight on the challenges in deciding what stations to drop, which can be based on their effect on the outputs of the assessment but is not easily addressed when the survey has multiple objectives. This further highlighted the need to consider how reducing the number of stations sampled translates into actual savings in time. Finally, Marshall et al. were able to translate consequences and trade-offs associated with survey effort reductions (i.e. sampling frequency) using a Management Strategy Evaluation (MSE) into consequent management decisions that would lead to increased closures of Pacific hake fisheries in cases of reduced sampling frequency.

2.1.3 Best existing recommendations for dealing with survey effort reductions

There is a variety of options for dealing with survey effort reductions but there is not a single best approach under all circumstances, as highlighted in Section 2.1.1 under consequences. Among the factors that contribute to the choice of the most suitable strategy for a given scenario are (a) whether the reduction is temporary or long term, (b) the amount of advance notice of the need for survey reduction (e.g. months before survey start in the early planning stages vs. unplanned reductions during the course of the survey), (c) the extent to which the survey effort needs to be reduced, whether the survey is single or multi species, (d) the specific reason that necessitates survey reduction (e.g. budgetary considerations vs. uncontrollable loss of survey area) etc. Regardless of the specific circumstances, in most cases an attempt should be made to increase the survey efficiency while keeping most major aspects of the survey intact, those include the survey design, spatial coverage, survey frequency, and sampling density. Within the ICES context, this may have to be discussed within labs (e.g. in the US) or at the national level or within/between ICES survey coordination working groups (e.g. in the Northeast Atlantic with WGBEAM, WGMEGS etc.). Survey and assessment scientists should review the survey’s objectives and priorities before discussing the specific reduction scenarios and assessing their consequences. Modelling and simulations can offer further insight into how to weigh the potential outcomes of a particular decision, whether it has to be made on the spot in the field or it was planned ahead (e.g. how many stations could be dropped to still achieve a certain error rate etc.); such approaches will be discussed and introduced in subsequent TORs.

Short-term changes

When issues arise in the field requiring short-term changes to or reductions in sampling effort, the survey scientist determining how to proceed should have clear instructions on the priorities and objectives of the survey that will inform the decision to be made.

The first question that should be considered is whether the processing time can be reduced. In the common example of a bottom-trawl survey, processing time could be reduced by a number of approaches, with the major caveat that there is a risk to reduce data quality or value of information collected by the survey:
1. Reducing the number of biological samples or “bonus” samples taken from each haul - Efficiencies may be gained by reducing the number of processed biological specimens to a minimum required for stock assessment (e.g. because sex-determination of fish is often one of the most labour-intensive steps in catch processing, consideration should be given to sub-sampling the fish to be both sexed and measured, rather than just measured, similarly taking otoliths may be time-consuming and numbers can be optimized on a species by species basis). Considerations for reducing or eliminating the processing of low-priority species (e.g. invertebrates in the NOAA Aleutian Islands survey) and time spent on extraneous research projects that are not part of the core survey priorities should be made when these interfere with the collection of the next sample (e.g. closely spaced stations in stratified random surveys where catch processing time, rather than transit time between stations, limits survey progress).

2. Shortening the duration of the tows, provided the minimum duration is still met; reduction of tow duration not only reduces the time spent on sample collection, but also reduces the size of the catch to be processed, but may have consequences for capture efficiency of large, mobile species.

3. Subsampling the catch more than might be standard practice

Once all gains in processing efficiencies have been exhausted and the need for further survey effort reductions is evident, the highest priority is likely to be to maintain the survey design and spatial coverage in most cases. If there are clear instructions that entire spatial coverage is not essential, then altering the survey area by removing remote sites, whole strata or geographical regions - such as those that would take longer to reach or to fish, or those that pose additional challenges due to weather for example - may be considered.

If, according to the set survey priorities, the spatial coverage is essential to maintain, then the survey scientist should reduce the sampling density. The protocol for reducing sampling density in the field should include well-defined criteria for which stations to eliminate. For example, stations located in areas known to be difficult to sample for any reason may be good candidates for dropping, whereas certain high-priority (“prime”) stations may be identified in advance as stations that cannot or should not be cut from the survey. For stratified surveys, a minimum of two stations should be maintained in each stratum, and the priority for cutting stations should be in strata with a large number of stations. However, if reduction in spatial coverage is viable, consideration should be given to eliminate entire strata from the survey if they are particularly costly to sample (e.g. strata > 700 m depth in GOABTS). This may be a good option if there are other methods for sampling these strata (e.g. longline survey at depths > 200 m in Gulf of Alaska).

**Long term changes**

As in the short-term reduction scenario, the first consideration should be whether processing times can be reduced and/or made more efficient. While the options suggested in the short-term section, i.e. subsampling, reducing biological samples, and reducing tow durations, are also options for long term reductions, here they can be implemented more systematically for the entirety of the survey. Another potential option is to research new technologies that might aid in more efficiently processing the catch and subsequent data, such as electronic length boards or other electronic data collection devices. If it is identified that improving the efficiency of processing times will not be sufficient to address the reduction, then the next step is again to ask whether the spatial coverage is essential to maintain. If not, then spatial coverage can be adjusted in the same manner as described above, by removing sites that are challenging to sample, geographical
areas or strata, but the decisions can be informed by further research on underlying ecological or biological processes that might offer justifications on whether to drop or keep particular areas in the long run.

If spatial coverage has to be maintained, among the remaining options are reductions of sampling density, survey frequency, and the number of passes through each station in case of surveys with replicate samples. Of these, the latter is likely to be the best option if available. Otherwise, careful consideration of the trade-offs between reductions of survey frequency and sampling density is required where both options are available. In this scenario the status of a species may have to be considered into the priorities for single-species surveys (e.g. stocks in good shape may not need as much surveying as marginal stocks and long-lived stocks that have slower changes in abundance can be monitored with less frequent surveys). However, in some cases (e.g. NOAA Eastern Bering Sea Shelf where catch quotas of crabs are based on annual estimates of absolute abundance), the only viable option may be a reduction of sampling density.

Finally, in cases where it is essential to preserve the status quo of all major options for survey reduction (spatial coverage, sample density, and survey frequency), it may be necessary to consider a change of survey design, or even cancelling the survey altogether. Particular circumstances will apply and risks should be considered when dealing with international collaborations (mostly in the EU). In any case, this decision should not be taken lightly, particularly for surveys with a long time-series. Survey objectives and priorities should be discussed again and all possible decision scenarios should be re-examined. If a full design change is considered to be the only remaining option, a resource such as ICES CRR #347 (de Boois, 2019) might prove valuable in the redesign process as it offers insight into considerations that should be incorporated into survey designs.

**Conclusive remarks**

Ultimately, there needs to be a dialogue between the survey and assessment scientists, managers and possibly stakeholders, when applicable, in deciding how to reduce survey effort. Data end-users need to identify their priorities for the data and survey scientists are responsible for providing advice on logistics, field considerations as well as present coefficients of variation for quality control even if not directly used in assessments, as a criterion for validating the use of a survey index. In the ICES context, WGISDAA is a working group that aims at facilitating some aspects of this dialogue, it is aimed as a platform for stock assessment and survey scientists to engage to help address a number of challenges related to survey data products. Some more specific survey and assessment teams already communicate outside this WG, e.g. coordination of time and location meetings between WGMEGS (egg survey) and WGWIDTH (assessment WG that uses the egg survey).

**2.1.4 What are current processes lacking? What are the major challenges?**

While case studies provide very helpful guidance on likely relative impacts of survey effort reductions on many stocks and species, the actual effect of any change in effort on a survey index will depend on the interaction between the target species abundance, range, patchiness (spatial aggregation) patterns and the sampling design. Thus, some effort reductions may have modest impacts for abundant, broadly distributed species, but cross a threshold beyond which indices cannot be developed or are uninformative for rare or very patchily distributed species.
Ideally, the informational needs of all assessments that rely on a given survey to inform trends and demographic structure would be considered in effort reduction scenarios.

Workshop participants recognize the use of survey information beyond single species assessments, to better inform multispecies and ecosystem modelling efforts, which in turn help inform ecosystem-based fisheries management. However, as these are generally emerging disciplines, rigorous evaluation of the consequences of effort reduction on ecosystem science products is difficult. However, ecosystem research needs and priorities should be evaluated relative to other priorities to survey teams by stakeholders, researchers and assessment analysts.

The lack of understanding of underlying processes may hold back progress towards designing and adapting surveys that capture spatial and temporal changes and variations. Such processes may be related to habitat drivers of species distribution, drivers of recruitment and migrations such as the presence or absence of the cold pool in the Bering Sea or the effect of El Niño vs La Niña years on Pacific Hake. Research on drivers and predictability of those drivers may facilitate the design of effort reduction schemes or even lead to adaptive sampling approaches. As species distribution patterns and the understanding of drivers that alter distributions are better understood, a more informed approach for effort reduction could be to use such information to rank the information content of given stations (or regions) to indices, and use this information to quantitatively evaluate the multispecies impacts of dropping “low-priority” stations (or more sparsely sampling “low priority” regions or strata) to inform both short and long term effort reductions (as in the Jorgensen et al. case study). Further, better understanding a system may lead to increasing predictability and reduce required sampling.

In many of the simulation and empirical (resampling) effort reduction studies, sampling locations were randomly removed. However, in most cases reducing sampling density relatively evenly throughout a survey region would require that a greater fraction of vessel time be spent transiting, and less fraction sampling. Future simulation or resampling studies should attempt to account for this non-linear relationship between effort time reductions and sampling density.

There needs to be more research and guidelines on the implication of changing the full survey design. For example, choosing to sample a few fixed stations has been used in the context of having limited funds to sample a large area, where random sampling would likely provide highly uncertain estimates that would not capture temporal trends. Fixed survey designs may also be maintained in order to not disrupt a long time-series (example of Isle of Man scallop survey in the Irish Sea). However, a number of problems are recognized with this design, such as concerns over local depletion (for example for bottom dwelling rockfish), or distrust from the industry as the stations sampled for historical reasons may not be located in the most productive areas. Depending on the coverage and density, this design may also not offer valuable information on spatial and temporal changes. Impacts of reductions that would lead from a random sampling design to a systematic or fixed stations design, or how to deal with/calibrate the time-series, should be carefully evaluated.

It is worth noting that although the focus here was on survey effort reductions, challenges can also arise when effort is increased, particularly if areas formerly not sampled are sampled partway through a time-series. This was a challenge when the NWFSC hook and line survey began sampling within a large closed area (the Cowcod Conservation Area) to better inform the stock assessment for this rebuilding species. Site effects in the index model were used to account for the survey expansion, and as compositional data indicated differences in size and age structure between the historical and expanded survey locations, time blocks were used to allow selectivity in the assessment model to vary between the periods.
In planning and allocating survey effort and resources, improvements in communication processes and procedures among survey groups, stock assessment analysts and other users (e.g. ecosystem analysts) would improve efficiencies. Examples were discussed during the workshop of changes in age sampling needs at AFSC to develop more achievable targets and logistically feasible sampling approaches. Improving such information transfer can result in better informed decisions regarding changes in survey protocols and an improved evaluation of the trade-offs associated with alternative means of addressing effort reductions.

It may be worth considering situations in which partnerships with industry or citizen scientists can allow for increased sampling at lesser cost, while still involving scientists to ensure that such programs maintain scientific standards and rigor. For example, the California Cooperative Fisheries Research Program, a hook and line survey to evaluate the effectiveness of state marine protected areas uses chartered fishing vessels and volunteer anglers alongside paid scientists to reduce the expense of the survey (Yochum et al. 2011).

### 2.1.5 Recommended future directions

Survey reductions can be seen as short-term, i.e. in response to a temporary cut in budget or a patch of bad weather or a broken vessel, or long term, i.e. in response to a permanent change in budget or a new protected area. We, as a collective of survey and assessment scientists, have identified some of the best and most common practices for accommodating both short-term and long term reductions, as detailed in 2.1.3. The type of decisions that can be made ahead of a survey or during a survey both rely on clear objectives and priorities that must be informed by dialogue and research before change is required.
Long term (ahead of survey)

Figure 2.1.1. Decision tree for considering long term changes to survey effort and design.
Short term (in the field)

Figure 2.1.2. Decision tree for considering short-term changes to survey effort and design.

Research and tools development

Responses to short- and long-term unexpected survey effort reductions by the collective survey community (i.e., survey and stock assessment scientists and resource managers) will benefit from the informed perspective on considerations and consequences of various effort reduction strategies presented here. Additional benefits can be realized from directed research efforts designed to elaborate or improve these strategies. A list of process and empirical research needs has been developed in section 2.1.4 above, which may not be comprehensive but covers important aspects identified by the group. Here we highlight some of the avenues for additional research that we believe hold the greatest potential to preserve survey quality and avoid unintended consequences when required to reduce survey effort. A first response to either anticipated or unanticipated survey effort reductions should be a review of process-related tasks to identify if efficiencies can be gained internally. Risk assessments and optimization are fields of statistical endeavour that provide a framework for approaching questions of process efficiency. In emergent situations (typically while already at sea) when reductions in survey effort must be affected quickly (e.g., in cases of extended mechanical failures or poor weather conditions), the resulting ad hoc effort reduction can negatively affect survey goals and objectives when not well considered. In these cases, early deliberation or applied research directed at anticipating either “prime” stations that cannot be dropped or “low priority” stations that can be dropped can facilitate at-sea scientists achieving a more objective effort reduction response. Specifically, emergent situations at sea would benefit from the development of programmatic tools to aid in the spatial delineation of
optimized sampling effort changes (increases or decreases) taking into account transit times, consequences for abundance estimators, and losses or gains in estimator precision.

Since the collective survey science community is in agreement that survey effort reduction is inevitable, be it emergent, short- or long term, we can also agree to longer term research activities designed to anticipate and address our effort reduction concerns. For instance, in multispecies fisheries, reducing survey density, frequency, or survey area will always result in species-specific impacts to data collections and assessments; some species being impacted more than others. These “trade-offs” in the repercussions to multispecies surveys form an arena of study where additional research efforts will better inform our responses to reducing survey effort. Jorgensen et al.’s paper (presented at this Workshop) provided one potential method for assessing spatial priority in multispecies surveys utilizing information criteria in the eastern Bering Sea. Ongoing research in the fields of species distribution modelling and essential fish habitat could also be expanded to inform decisions regarding survey effort reduction from a habitat- or process-linked view of distribution and abundance.

2.1.6 References and articles on related topics


Blaine, J., B. Pacunski, and D. Lowry. In progress. WDFW Southern Salish Sea scientific bottom-trawling: sampling design changes and consequences. WDFW FT-20XX-XX.


Huret, Martin, Bourriau, Paul, Doray, Mathieu, Gohin, Francis, Petitgas, Pierre. 2017. Survey timing vs. ecosystem scheduling: Degree-days to underpin observed interannual variability in marine ecosystems. Progress in Oceanography, 166: 30-40


2.2 TOR II. “Total Survey Uncertainty”: Develop key quality metrics that can be used to describe “total survey uncertainty” for survey derived indices of abundance for common survey designs

Subgroup Leads- Lewis Barnett, Kristin Marshall, Kotaro Ono
Participants- Jason Conner, Peter Munro, Wayne Palsson, Eric Ward

2.2.1 What is total survey uncertainty?

Total survey uncertainty is a comprehensive representation of the accuracy and precision of a survey data product. Here, we focus on the application of this concept to an abundance index. Total survey uncertainty could be expressed as a value that incorporates all sources of error (Table 2.2.1): analytical sources (i.e. from choice of estimation approach, model structure and assumptions), biophysical sources (i.e. from individual and population processes and how they vary with environmental conditions), and observational sources (i.e, from the sampling process). This differs from the typical representation of uncertainty in abundance indices, which often only includes a single aspect of observation error (e.g. the sample coefficient of variation in catch rates from a design-based estimator).

Table 2.2.1. Sources of Uncertainty and their Characteristics.

<table>
<thead>
<tr>
<th>Uncertainty source category</th>
<th>Uncertainty source component</th>
<th>Description of the problem</th>
<th>Possible consequences for abundance index</th>
<th>Possible ways to mitigate problems or evaluate their influence on the index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biophysical Density-dependence</td>
<td>Survey catch efficiency is influenced by fish density-dependent behavior (could be influenced by environmental variables) Is influenced on physical characteristics of the sampler (size, speed, duration, etc.)</td>
<td>Mean: possible bias Variance: underestimation</td>
<td>Conduct lab experiment Conduct field experiment (parallel sampling using two different instruments) Develop estimation model using acoustic and trawl data</td>
<td></td>
</tr>
<tr>
<td>Biophysical Distributional (spatial and vertical)</td>
<td>Is the catch representative of the fish population’s entire distributional range (both horizontal and vertical) Does the spatial distribution of the population vary with changes in abundance or population structure (e.g. ontogenetic shifts in habitat)</td>
<td>Mean: possible bias Variance: unknown</td>
<td>Conduct simulation e.g. mechanistic species distribution model (or population dynamics model) Develop estimation model using acoustic and trawl data Deal directly in the assessment model</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Description</td>
<td>Mean: Possible Bias</td>
<td>Variance: Underestimation</td>
<td>Methodology</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Biophysical</td>
<td>Behavioural response to gear</td>
<td>Mean: possible bias</td>
<td>Variance: underestimation</td>
<td>Conduct simulation e.g. IBM (Thorson et al. 2011)</td>
</tr>
<tr>
<td></td>
<td>Survey catch efficiency can be influenced by fish herding behaviour. Related to density-dependence.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biophysical</td>
<td>Environmental variables</td>
<td>Mean: possible bias</td>
<td>Variance: Unknown</td>
<td>Conduct simulation e.g. species distribution model, comparing models with different covariates, forms (e.g. linear, quadratic), and scales</td>
</tr>
<tr>
<td></td>
<td>Do environmental covariates influence the species distribution?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Are we modelling their effect correctly?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>What is the appropriate spatial and temporal scale for defining the covariate?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(e.g. point estimate at start vs end of haul)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observational</td>
<td>Survey progression</td>
<td>Mean: possible bias</td>
<td>Variance: unknown</td>
<td>Conduct simulation e.g. IBM of survey vessel operation</td>
</tr>
<tr>
<td></td>
<td>Does survey duration affect catch efficiency? Some portion of catch happens during the descent and ascent of the trawl net (portion ignored in the analysis) and its influence on total catch depends on tow duration. Also longer tows can tire and then catch faster, mobile species</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observational</td>
<td>Measurement error (equipment tolerance and data processing)</td>
<td>Mean: possible bias</td>
<td>Variance: underestimation</td>
<td>Conduct simulation e.g. IBM of survey vessel operation</td>
</tr>
<tr>
<td></td>
<td>Does the gear opening change during the tow duration? If so, how can it affect fish escape-ment?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Are some data removed before the analysis? Is it reasonable?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>How much sampling error is introduced by sub-sampling methods?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observational</td>
<td>Gear-capture efficiency</td>
<td>Mean: possible bias</td>
<td>Variance: underestimation</td>
<td>Conduct simulation e.g species distribution model</td>
</tr>
<tr>
<td></td>
<td>Do we consider differences in gear catch efficiency?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the gear catch efficiency change over space and time?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observational</td>
<td>Vessel effect</td>
<td>Mean: possible bias</td>
<td>Variance: Unknown</td>
<td>Collect vessel characteristics and include in estimation model</td>
</tr>
<tr>
<td></td>
<td>Some technical characteristics of the vessel influence catchability (horse power, boat speed, skipper, sound)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deal with it directly in the assessment model</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Each model comes with its own assumption and how does this affect the estimate of uncertainty? Should the uncertainty estimate be based on confidence interval or should it include all sources of uncertainty (e.g., posterior predictive interval)?

---

**Issue of temporal scale for data analysis:** time lapse between the data collected at the beginning and end of survey

**Issue of spatial scale of data analysis:** are the covariate information collected representative of the survey haul?

**Issue of spatial scale in prediction:** is it reasonable to assume that all covariates values are the same within prediction grids

**Treatment of the random effect during prediction:** Which vessel ID to choose when making prediction? Is that a fair representation?

---

### 2.2.2 Why evaluate total survey uncertainty?

It is critical to quantify total survey uncertainty to accurately represent the quality and reliability of estimated survey data products, which in the context of this TOR may include point estimates of abundance, their variance, trend, scale, and perhaps projections. For example, the quality of an abundance index can be affected by the method of standardization and consistency of sampling methods. It is important to communicate the total survey uncertainty because abundance index estimates and variances can influence the interpretation of products made by users of abundance indices, notably in the context of stock assessments.

The most relevant example of the importance of propagating uncertainty in survey estimates of abundance comes from their potential influence on biomass estimates from stock assessments. Abundance index variances are themselves random variables (reviewed in Kotwicki and Ono 2019), whose own uncertainty is not typically estimated (with some exceptions, e.g., Schnute and Haigh 2003, Cadigan 2011), or estimated with high uncertainty. Despite this uncertainty, abundance index CVs are often used as data weights in stock assessments—where estimates with higher CVs are down-weighted—potentially affecting the influence of the abundance estimates among years and data types (e.g., compositional data). Abundance index variances are commonly derived from design-based approaches, however, many surveys violate at least some statistical assumptions of randomization due to practical constraints. Such design-based abundance index variance estimates also fail to account for process error from e.g. variable catchability. While some components of abundance index variance can be estimated with model-based approaches to standardize abundance indices (e.g., variable catchability in certain situations), it is likely impossible to estimate all such components.
Table 2.2.1 summarizes the primary components contributing to total uncertainty in survey-derived abundance indices, their effects on accuracy and precision, and potential solutions for assessing or mitigating such effects. Examples of these components include the patchiness of species distributions, species’ abundance, proportion of habitat that can be sampled (e.g., due to limitations of gear), selectivity (gear efficiency, saturation of hooks, etc.), sampling coverage, sampling intensity and design, estimation model assumptions, and spatial availability (fish movement and habitat selection). The availability of a species to a fisheries survey is a key component of abundance index uncertainty. When the spatial distribution of a species changes, perhaps shifting due to environmental or biological pressures, an established survey extent may not include enough of the population to robustly estimate an abundance index. As an example, yellowfin sole in the eastern Bering Sea shelf may delay seasonal migration in response to cold seawater temperatures, which can cause interannual variability in the overlap of their distribution with coverage of the bottom-trawl survey in this region (Wilderbuer et al., 2015). Total survey uncertainty may also vary over time with management changes (spatial closures, untrawlable habitat) and species distribution shifts, or between surveys of different types in the same area (fishery-dependent vs. fishery-independent surveys, and among gear types: acoustic, trawl, visual, etc.).

2.2.3 How can we evaluate total survey uncertainty?

Total survey uncertainty could be estimated as an “additional survey variance” parameter within a stock assessment model, or independently (Maunder and Starr 2003, Maunder and Punt 2004, Mauner and Piner 2015). The reliability of such estimates of total survey uncertainty within assessments is difficult to quantify as its estimation performance depends on many other data inputs and parameters of a stock assessment model. Therefore, here we focus on estimating total survey variance when standardizing the abundance index (outside of the assessment model).

To quantify total survey uncertainty and how it varies with survey effort and design, we need metrics that quantify the difference between abundance estimates and the true value of abundance. Examples of such metrics include the variance of mean CPUE among surveys of the same population (e.g. Spencer et al. WKUSER), but also the variance of the variance of the mean CPUE based on multiple realizations (e.g. MCMC draws) of the underlying spatial distribution of abundance (e.g. Kotwicki and Ono 2019; Conner et al. WKUSER), and the relative difference (e.g. relative error) between the true variability around the abundance index and the sample variance (Kotwicki and Ono 2019). The former measures of survey uncertainty are sometimes referred to as the “true” coefficient of variation (CV) of an annual abundance index. Further, there are component metrics that are useful for assessing the performance of estimates of abundance and their uncertainty, including those that evaluate accuracy (e.g. RMSE or root mean square error), precision (e.g. “true” survey CV), and both accuracy and precision (e.g. REE or relative estimation error).

Although it is possible to directly estimate individual components of survey uncertainty using empirical approaches (Table 2.2.1), it is highly implausible that such approaches alone could be used to quantify the combination of all the factors contributing to total survey uncertainty. To do this empirically would require replicate surveys of a population at a given time and place. Thus, to more comprehensively quantify total survey uncertainty we will likely need to estimate the components of survey uncertainty within an estimation model or perform simulation studies with known “true” abundance generated by an operating model.

Statistical models are often used as the backbone of operating models, because simulating from them allows researchers to design experiments and evaluate the performance of estimators or management strategies while knowing the “true” state of the system. There is a wide range of
model types that can be adopted as operating models to evaluate effects of model parameters on total uncertainty in the estimated index of abundance. These operating models can be divided into non-spatial approaches versus spatial or spatio-temporal models. Non-spatial approaches might include sensitivity analyses to components of design-based estimators, as implemented in parametric statistical models (Schnute and Haigh 2003) or parametric bootstrapping (Smith 1997). The resolution of either of these approaches is determined by the area of the strata being used and confidence intervals from bootstrapping or other resampling approaches that are limited by distributional assumptions (e.g. parameters of interest are approximately normally distributed). Other techniques for evaluating uncertainty in indices of fish abundance have used operating models that have been hybrids between the parametric resampling approaches and full geostatistical models of abundance (Thorson et al. 2015). One example of a stratified operating model comes from Conn (2011), who used parametric models to evaluate the effects of sampling intensity and spatial coverage for spatially structured populations of reef fish. A second approach is the spatially gridded metapopulation model used by Ono et al. (2015) which combined simple population growth models (e.g. logistic) with movement rules governing between-cell dispersal. A third approach is to adopt an individual- or agent-based operating model, such as that used by Thorson et al. (2012) to evaluate the effect of fish shoaling on index standardization.

More recent approaches to evaluating uncertainty in standardized indices of abundance have used spatial or spatio-temporal operating models. Spatial models may be used to simulate spatial heterogeneity in the density of a species, while spatio-temporal models may be useful in simulating populations that are both patchily distributed over space, and whose spatial patterns change through time (randomly or with persistent trends). The majority of recent spatial operating models adopt the same general structure and assumptions as the estimation model (e.g. VAST and INLA, Thorson 2019, examples in Mormede et al. 2020, Thorson et al. 2019a, Zhou et al. 2019, Bryan and Thorson WKUSER). These approaches all model fish density via Gaussian random fields, which generate spatio-temporal variability in density, but lack the mechanistic underpinnings of alternative models which may be desirable for certain applications. Approaches incorporating such mechanistic drivers have been explored, including adding explicit spatial movement of stocks (Kerr et al. 2017).

The utility of any of the above approaches for simulating variation in fish density is partially related to how well the operating model captures reality. For species that are infrequently encountered or occur in approximately the same densities throughout their range, the full spatial or spatio-temporal models described above may not be needed. The interaction between operating model and estimation model will likely affect the relative importance of different sources of uncertainty, such as inferring whether biological processes (shoaling, density-dependence, migration) contribute a larger source of variation than sampling processes (variable catchability, density and extent of sampling over space). As the approximations between the operating model and truth or interactions between operating model and estimation model are never known a priori, a general strategy is to perform sensitivity analyses of operating model parameters both within and across operating models. For example, an operating model designed to evaluate the effects of population movement outside of the survey area would likely need to include scenarios representing (1) a constant fraction of the stock remaining outside the survey area, (2) random interannual variability in the fraction of the stock beyond the survey boundary, and (3) temporal trend(s) in the mean and variance of the fraction of the stock outside the area.
Research presented at WKUSER begins to increase our understanding of uncertainty in abundance index means and variances derived from fishery-independent survey data, with a particular focus on Alaskan groundfish. One key difference between these approaches and examples in existing literature (e.g. Liu et al., 2009) is the use here of multiple surveys replicated in time, and in some cases (Conner et al. WKUSER), multiple simulated “true” distributions of fish density for any given time. This form of replication and representation of uncertainty in the true and sampled densities is a positive step towards estimating total survey uncertainty and understanding generally how survey changes will affect the quality of abundance indices. These simulation studies demonstrate the influence of sampling design and choice of estimators on the accuracy and precision of abundance indices. However, these studies typically focus on sources of variation within the survey itself, and do not include sources of error outside of the observation process. It will be valuable to extend these analyses to explore the generality of strategies for mitigating reductions in survey resources by incorporating multiple operating models (structure or parameterization), in addition to implementing a broader variety of sampling designs, species distributions (spatial and temporal correlation structures, and spatial arrangement), and sources of uncertainty (Table 2.2.1).

Several simulation studies were conducted based on the spatio-temporal operating model developed by Kotwicki and Ono (2019). Laman et al. (WKUSER) used this operating model to simulate the current strategy of survey effort reduction in the Gulf of Alaska, reducing sampling intensity and eliminating sampling from the deepest strata. They found that this increased uncertainty and increased errors and biases in the abundance index, particularly for deeper-dwelling species (sablefish, Anoplopoma fimbria), regardless of estimation approach. von Szalay et al. (WKUSER) compared the response of design- and model-based estimators among sampling designs for a few key species in the Gulf of Alaska. They found that even model-based estimators were sensitive to sampling design, as abundance index variance estimates differed greatly between systematic and random sampling designs while mean abundance estimates were similar. They concluded that model-based estimation (with VAST, Thorson 2019) with a simple random sampling design typically performed best, yet indicated that more exploration is needed to test the generality of this result. Systematic sampling designs produced biased design-based estimates of means and variances of population density for multiple stocks in the eastern Bering Sea shelf, where the magnitude and direction of bias in the abundance index depended on differences in spatial distributions among species (Conner et al., WKUSER).

Kotwicki and Ono (2019) demonstrate the impact of one component of catchability, sampling efficiency and its variability, on estimates of survey uncertainty. Their work reinforces the necessity of catch efficiency studies and highlights the value of addressing catchability in the estimation model. The key metric they used to evaluate the total survey variance of an annual abundance index was the mean squared difference between the true mean abundance in a given year and the estimated mean abundances from multiple simulated surveys.

In addition to sampling efficiency, another major component of catchability is spatial availability, or the extent to which the spatial distribution of a stock overlaps with the spatial coverage of a survey. The challenge of understanding how spatial availability can affect the total variance of an abundance index needs further evaluation. However, some studies have evaluated the influence of vertical availability to a trawl survey and its contribution to the variance of an abundance index (e.g. Kotwicki et al. 2009).

Spencer et al. (WKUSER) addressed how variance propagates from survey estimates of abundance to stock assessment outputs in a simulation study. Their operating model simulates a trend
in biomass determined from past assessment output and distributes biomass spatially according to a variogram fit to past survey data. They then simulate surveys to generate a design-based abundance index and CV. These are passed to a simple stock assessment model where the only input is abundance index and the biomass is estimated from a generalized linear mixed-effects model fit to the design-based index. They found that the propagation of survey uncertainty depends on observation error, process error variance, and the magnitude of the underlying abundance trend. As expected, sample CVs increased with decreasing sampling intensity, and this inflation of CVs reduced the accuracy of estimated biomass and biomass trends from the stock assessment model when these CVs were used to weight the abundance estimates among years.

2.2.5 Major gaps in existing information and future research priorities

In an ideal world, all sources of uncertainty (Table 2.2.1) would be incorporated in an integrated estimation model to calculate the total uncertainty associated with the derived abundance index. Nonetheless, this is seldom possible as some sources of uncertainty cannot readily be addressed in an estimation model. In addition to the sources of uncertainty prevalent in many fisheries datasets around the world, we highlight several additional data gaps and needs for future research activities.

1. Non-stationary trends may be present in biological, physical, and sampling processes that may create variability or trends in biases or uncertainties through space or time. Examples include the effects of environmental variables on species’ behavioural responses to sampling gear, shifts in their vertical distribution, spatial shifts or migrations, and any spatial or temporal variation in sampling effort or coverage (e.g. due to creation of protected areas or offshore wind farms), measurement error (e.g. technological improvements in tolerances), and catchability (e.g. sampling efficiency). Ultimately, trends in any of these processes and their specific consequences are generally unknowable from the survey data alone, and can only be estimated or validated from external data sources and simulation approaches.

2. The majority of advances in index standardization modelling has been in application to single-species stock assessment models. The impacts of common uncertainties (Table 2.2.1) are less straightforward in their effects and mitigation in the context of multispecies surveys or multispecies models for informing optimal monitoring and management decisions. Examples may include interactions between species, or species-specific responses to variation in sampling intensity and design (e.g. common and diffusely distributed species versus rare and patchily distributed species). Despite these challenges, multispecies surveys offer the opportunity to leverage information from one species to inform another (this may be achieved via shared model parameters, or estimating the covariance matrix between species occurrence probability or densities).

3. Despite rapid advances in spatio-temporal estimation models, few studies have attempted to validate the robustness of estimation approaches with different operating models. This presents a need for the development of general operating models (individual or agent based models, structured or unstructured spatially explicit population and community models, etc.) independent of the Gaussian random field approach used in INLA and other software (Rue et al. 2009, Anderson and Ward 2019). More realism in future operating models may be achieved via incorporating additional data streams (e.g. telemetry-derived migration or movement rates, experimentally derived threshold responses to temperature).

4. While approaches for abundance index standardization with spatio-temporal models have accelerated over the last decade, there are many opportunities for improving on
these methods. Recognizing that not all uncertainties can be included in estimation models for example, in which case parameters with uncertainty uninformed by empirical evidence may be bounded or informed through prior distributions in a Bayesian framework. Other techniques may include meta-analysis via literature review (such as the recent approach used to construct steepness priors for rockfishes (Sebastes spp.) on the west coast of the USA, Thorson et al. 2019b).

5. Our description of major sources of uncertainty in estimates of relative abundance (Table 2.2.1) may be useful for identifying research projects to understand the contribution of these sources of variation to total survey uncertainty. However, the magnitudes of the contributions of these factors to the total uncertainty, and their own magnitude of variation in time and space, is likely to vary across species within a survey, as well as between surveys (those with different gear, location, design, etc.). We recommend a simulation approach with operating models designed to represent the system of interest. This approach would allow sources of variability to be ranked on a relative scale, and facilitate evaluation of the robustness of these rankings to operating model assumptions (e.g. the degree to which the spatial distribution of a given stock is within the survey boundaries, or whether any components of the model have temporal trends).

6. Fisheries scientists and managers would benefit from better communication about which sources of uncertainty are being quantified when relative estimates of abundance are presented. Recent advances in spatio-temporal modelling with Gaussian random fields have largely been developed in a maximum likelihood setting for computational efficiency (e.g. VAST, Thorson 2019). These approaches allow for uncertainty intervals to be generated, as is done in many applications to fisheries stock assessment. A key point is that we differentiate the use of confidence intervals from prediction intervals. Confidence intervals represent uncertainty in predictions of the mean, and incorporate only those uncertainties associated with residual or observation errors (measurement, sampling errors), while prediction intervals represent bounds of uncertainty for predicting future data. Predictive intervals include the same sources of uncertainty as confidence intervals, but also include uncertainty in model parameters. Using simple linear regression as an example, the prediction intervals would incorporate uncertainty in the slope and intercept parameters whereas the confidence intervals would not. As a result of these differences, predictive intervals are wider than confidence intervals. The majority of model-based approaches to abundance index standardization in fisheries present estimates with confidence intervals. While we recognize that part of this choice is because of computational limitations, advancing these methods in a Bayesian framework (e.g. by sampling from the posterior predictive distribution as in the operating model of Kotwicki and Ono (2019)) or other approaches to generate predictive intervals would be useful. Such approaches will be particularly critical towards representing uncertainty beyond that which can be represented by past observations, given that global change is likely to generate novel environmental conditions and fish population states in the near future.

2.2.6 Roadmap for improving quantification of total survey uncertainty

In this section we aim to provide practical guidance to practitioners to more fully account for uncertainty in their surveys and downstream data products, such as abundance indices (Figure 2.2.1). The first question to consider is whether a complete estimate of abundance and its total uncertainty exists from another survey, which could be used for calibration. For example, new or complementary technologies such as camera or ROV surveys may provide these estimates.
If this option does not exist currently, it may be worth revisiting in the future as technology advances because this is one of the only ways to empirically measure the total uncertainty associated with a given survey.

Absent a complete estimate of abundance and its total uncertainty, the next step is to build an understanding of total survey uncertainty from the bottom up, considering all potential contributing factors (Table 2.2.1). We recommend listing all potential components under the categories above (e.g. observation error, sampling efficiency, spatial availability, estimation error, structural uncertainty) that would likely be relevant to the specific survey/species/region in question. As mentioned above, traditionally abundance index uncertainty only considers the sample variance as a measure of observation error, and as such, special attention should be given to the other types of uncertainty. Expert knowledge from biologists, survey operators, and end-users (e.g. stock assessors) should contribute to creating this list of potential factors and prioritizing their likely relative importance for a given survey and species.

Figure 2.2.1. Roadmap for Quantifying Abundance Index Uncertainty.

For each contributing factor, further consideration is needed as to whether the uncertainty due to that factor can be informed by published studies (or metaanalysis), or if additional research such as capture efficiency is needed (and it is possible) to quantify it. For example, see Table 2.2.1 for guidance on lab and field studies and their contribution to quantifying capture efficiency. If lab and field studies are not possible or not applicable to the type of uncertainty (e.g. species distribution shifts are occurring), another possibility is accounting for the uncertainty during the development of the index (within an estimation model).

For any remaining components of uncertainty that are prioritized as important and cannot be addressed with any of the above approaches, simulation studies are needed. Simulations can be used to define plausible upper and lower limits of uncertainty components and to refine the
prioritization of uncertainty components by quantifying the sensitivity of estimates to variation in each component. When simulation studies are undertaken, the guidance in earlier sections can inform the choice of operating model structure most appropriate to characterize the specific set of concerns (spatial, temporal, etc.).

After the most important components contributing to total uncertainty have been considered and quantified, they will need to be combined into a single uncertainty metric (e.g. variance, CV, confidence or prediction interval) to accompany the resulting abundance estimates for each period. How to combine multiple sources of uncertainty is an area of ongoing research, encompassing fields much broader than this discussion focusing on survey uncertainty. The simplest approach would be to take the sum of the variance components; however, this assumes that the component uncertainties are independent and identically distributed, which is almost certainly violated in many cases. Under other assumptions and conditions, perhaps the component uncertainties can be aggregated by taking their product. Guidance on the best approach will likely be case-specific.

How to use improved estimates of total survey uncertainty will depend on the context and downstream user of the data product (e.g. abundance index as an input to a stock assessor’s model). In all cases, documentation of the considerations and approaches to quantifying the total survey uncertainty should accompany the abundance index (not only point estimates of the variances). Commonly, abundance indices inform stock assessment models, and communication between the creators of the index and stock assessors will be necessary to ensure uncertainty is propagated appropriately and not double-counted. It is a common practice in some regions to estimate and add an additional variance to an abundance index within a stock assessment model (Mauner and Punt 2004). However, if total survey uncertainty can be better characterized outside the assessment model, then adding additional uncertainty attributable to unknown sources may not be necessary or appropriate. In an integrated stock assessment framework, abundance indices are sometimes weighted by their respective uncertainties. In this context, more completely accounting for uncertainty would lead to larger estimates of uncertainty and potentially down-weighting the abundance index in the assessment. A better approach would be to more fully represent uncertainty in all data sources contributing to an assessment in the way we have outlined here for survey uncertainty. More holistic consideration of uncertainties across data sources would lead to more fair comparison and weighting of data in relevant assessments.

### 2.2.7 Near-term vs. long term priorities for improving quantification of total survey uncertainty

- **Short-term**
  - Prioritize component sources of uncertainty by their likely influence on the data product (Table 2.2.1).

- **Medium-term**
  - Conduct empirical research to directly estimate uncertainty components (e.g. catch efficiency).
  - Develop operating models tailored to quantify the contribution of specific sources of uncertainty to total variance.
  - Improve estimation models to partition sources of uncertainty.
  - Develop relative metrics of total survey variance to facilitate comparisons among species, surveys, regions and uncertainty components.

- **Long term**
  - Test the influence of new survey design implementations on estimates of total survey uncertainty.
○ Evaluate the trade-offs of implementing adaptive survey designs, e.g. short-term forecasting to inform changes in survey design (e.g. coverage).

### 2.2.8 Outstanding questions and caveats

- What does the magnitude of total survey uncertainty mean for existing approaches to downstream users (e.g. data weighting in stock assessment models)?
- Factors outside of the scope considered here:
  ○ Estimation of, and incorporation of survey variance within assessment models.
  ○ Interactions with changes in compositional data (i.e. length and age structure) uncertainty.
- What is an acceptable level of total survey uncertainty? What is the change in total survey uncertainty caused by current processes used to mitigate changes in survey effort (see TOR 1)?
  ○ Provide objective advice about the acceptable magnitude of uncertainty by quantifying e.g. the probability of detecting an abundance trend as a function of the magnitude of the underlying trend, the interannual variation in abundance, and the true uncertainty of the abundance estimate (Spencer et al. WKUSER).
  ○ Consider soliciting subjective advice from policy-makers about the acceptable magnitude of uncertainty given their regional and species-specific priorities and their risk tolerance.
  ○ Moving beyond the context of a single abundance index for a given species:
    ■ Multiple species
      ○ Consider the influence of the reliability of abundance index on stock assessments in the context of differences among stocks in the availability of additional data sources that can be included in the stock assessment model (e.g. compositional data).
    ■ Multiple surveys
      ○ Intercalibration (see TOR 3)

### 2.2.9 WKUSER recommended actions to estimate total survey uncertainty

We outline several high-level recommendations for future studies to help quantify total survey uncertainty.

First, we emphasize the continued need for simulation-based research. To advance this simulation-based research, our advice is to develop operating models tailored to address specific questions, adding complexity only when needed to optimize one or more of these elements: inference, computational efficiency, model fit, realism, or generality (e.g. to what extent are the results species- and system-dependent). Probably the most powerful feature of the simulation-based approach is that the various sources of uncertainty outlined in Table 2.2.1 can be ranked by their relative (per unit) influence on an abundance index, and robustness of such inferences may be examined by developing operating models with different assumptions.

Second, we recommend the expansion of estimation models to fully propagate uncertainty from multiple sources (e.g. uncertainty in the sampling or observation process, in addition to uncertainty in the spatial distribution, and any other associated model parameters). Some of these
model parameters should also be informed by external data when available, such as field or laboratorial experiments.

Third, there may be an opportunity to bridge the two recommendations above to extend statistical approaches in existing estimation models to quantify the absolute influence of each uncertainty component on estimated abundance index means and variances, given each component’s magnitude of uncertainty.

Fourth, while spatio-temporal models for estimating the abundance of single species have become mainstream, there is a great need and many opportunities for expanding inference from these models to a multispecies or multi-survey context.

2.2.10 References


2.3 TOR III. “Survey Continuity”: Define changes to survey designs that require inter-survey calibration and what changes can be resolved by a model-based approach to index generation

Subgroup Leads- Jim Ianelli, Michael Martin, David Stokes
Participants: Patrik Borjesson, Walter Ingram, Cole Monnahan, Robert Pacunski, Chris Rooper, Kresimir Williams

For the purposes of this discussion, we defined “changes to survey designs” to include:

- Spatial changes
  - Planned reductions (dropping areas)
  - Thinning
  - Unplanned missing stations
  - Missing regions or strata
  - Merging two surveys with overlapping extents
  - Expanding ranges (e.g. Northern Bering Sea addition for pollock and Pacific cod)

- Temporal changes
  - Planned skipping years
  - Vessel or gear changes (abrupt)
  - Technology creep (e.g. improved net mensuration)
  - Improved knowledge of sampling process (e.g. trawl selectivity estimates improved through experimentation)

- Sampling design
  - E.g., Systematic to stratified random
  - Adaptive response to species distribution (also spatial, preferential sampling issue (Conn paper))
  - Dynamic area based on habitat (for area surveyed)
  - Changing tow duration / station effort
Furthermore, we considered “inter-survey” to mean an index developed for assessment purposes that may depend on multiple vessels/countries or simply a standard survey approach that has changed somehow (e.g. new nets and/or sampling protocols).

Based on discussions, model-based approaches for inter-survey calibration situations were suitable for the issues identified above except for situations where data from a “new” survey (e.g. new vessel) is non-overlapping and without any actual inter-calibration data. For assessment purposes, depending on the species, as “new” survey data accumulate the break in the survey time-series may become less important and the relative catchability can be reasonably estimated. A summary figure describing conclusions from our discussions indicates that for a number of situations calibration issues can be handled well by modelling (Figure 2.3.1).

### 2.3.1 Summary of existing examples of changes to survey design that required inter-survey calibrations

We provided references to the literature on survey calibrations in the “references” section below. A few examples on the inter-survey calibrations are listed below:

- Northeast US case of NOAA Fisheries RV Albatross versus RV Bigelow.
- Gulf of Alaska acoustic-trawl survey (NOAA RV Miller Freeman and to “quieter” vessel, RV Oscar Dyson).
- IBTS (with guidelines for overlap experiments for calibration). Swedish case of a new vessel being phased in following these guidelines.
- NOAA intermittent Northern Bering Sea Bottom Trawl Survey
- International Pacific Halibut Commission (IPHC)

For the NOAA Northern Bering Sea Bottom Trawl Survey and how it was dealt with in Eastern Bering Sea (EBS) pollock and Pacific cod assessments we noted that temporal correlation in the spatial distribution was used to estimate areas missed in each year. In the 2019 EBS Pacific cod assessment, a suite of 9 models with alternative spatial hypotheses and data applications were used in a weighted ensemble.

For the International Pacific Halibut Fisheries Commission (IPHC) case they abandoned design-based longline survey estimates of biomass in favor of a “space-time” model (Webster 2015). This had features that stabilized regional variability (for apportionment purposes), improved uncertainty estimates, extended estimates into infrequently surveyed areas (which also likely have low “leverage”), and allowed evaluation of sensitivity to thinning, chopping, or adding stations.

### 2.3.2 Summary of workshop presentations

The workshop presentation abstracts written by von Szalay, Williams, Conner, and Blaine provided some cases of inter-calibration approaches. In some cases, changes in survey designs were evaluated. On the US west coast, the group noted a dedicated workshop was held on how to best use different historical surveys (Anon. 2007).

In the Eastern Canada case, the presentation noted large areal gaps in survey coverage. It was unclear the extent to which calibrations were used but that some areas were dropped due to constraints on funding/ship availability while other areas were sampled due to different priorities (i.e. covering an area that was important for shrimp resources in favour of more traditional groundfish / cod regions).
Jason Conner investigated alternative sampling designs (systematic vs stratified and simple random) and this was an example where calibration for some species seemed evident. We agreed with his conclusion to further investigate why there were biases apparent from the simulated data using systematic sampling compared to stratified-random or simple random sampling. The implication is that if the design was switched from systematic to SRS then a calibration would be necessary because of the shift from a biased (high or low depending on species) to unbiased design. It was noted that perhaps sampling more MCMC draws for replicate purposes would be preferred over repeat sampling within an MCMC draw.

Elaina Jorgensen presented analyses dropping stations by relative leverage for a few species. The differences varied significantly by species and showed trade-offs in that some were completely different from others.

**2.3.3 Which survey data products are needed to be included in the inter-survey calibration?**

The subgroup reviewed information needed for calibration purposes, including for species-specific abundance indices and size/age composition, ecosystem indicators. A case for alternative station design having impact on the collection of physical variables showed that the impact on broad indicators would be relatively minor (e.g. Cynthia Young’s talk on the EBS BTS data collections). Gaps here could be filled in with alternative platforms as well (including remote sensing). The extent to which physical variables are used to “tune” oceanographic models was noted as being important.

Composition impacts (within and across species) could be impacted if catchability/selectivity changes. As noted elsewhere, model-based approaches such as VAST with dynamic factor analyses and multiple species could contribute to vessel effects (provided there is sufficient overlap). It was noted by a reviewer that the work by John Wallace and others on the WCBTS concluded that four replicate surveys with similar but uncalibrated vessels provided a good index because inter-vessel differences just became part of the overall variance (Helser et al. 2004, Keller et al. 2017).

![Figure 2.3.1. Pathways for evaluating survey continuity.](image-url)
2.3.4 Best existing recommendations

Our overall conclusion is that the best scenario is to have data and samples from which calibrations can be done, regardless of estimation approach. The group concluded that when possible, it would be preferable to thin existing area coverage and provide at least some sampling/data collection activities in the region that is dropped. For temporal breaks in surveys (e.g. a new vessel), after a period before and after the break is established, we recognized that stock assessment modelling may provide estimates of how the series has changed.

An example from NOAA EBS acoustic-trawl survey in 2018 occurred due to a vessel breakdown in which a number of transects were missed and represented about 6.1% (~6,016 nmi$^2$) of the normal core survey. The amount of pollock in the vicinity of the three unsampled transects at the survey end was estimated using ancillary acoustic data collected in that area by the chartered vessels conducting the bottom-trawl survey. The biomass estimate in this area was 178 kt or about 7.1% of the total survey biomass estimate (and 11% of total abundance). Had time permitted, a model-based approach might have been possible to combine these data more simply and obtain estimates to cover the missing area.

An example of how experiments can evaluate past (and current) sampling assumptions was provided from a midwater trawl gear configuration (de Robertis et al. 2017). This showed that assumptions about gear performance, specifically selectivity, can affect (and effectively recalibrate) a survey time-series. Here a set of experiments was performed to directly estimate length-dependent escapement (i.e. mesh selectivity) of the survey trawl gear using recapture or “pocket” nets and back-propagate an escapement correction factor to historical data to account for missing smaller fish that were observed in the echo sign/backscatter. Re-analysis of these surveys then provided a corrected time-series as input for the stock assessment model.

2.3.5 Path forward

For this section the subgroup was asked to address the following questions:

- What are current methods lacking?
- What are the major challenges?
- What research can be done to answer some questions?

In cases where change is required but money is unavailable for proper calibration experiments we noted that developing proxy approaches for inter-calibrations are needed. For example, in the Swedish case discussed, they will operate with other IBTS vessels that worked with the Danish research vessel so relative “overlap” observations will be available. A modelling approach should work for this case. A similar situation may help with the new West Coast Canadian vessel that can overlap with the RV Bell Shimada (and the Bell had previously overlapped with the RV Ricker). The group recommended that, to the extent practical, calibrations using such a proxy approach should be done using design-based and model-based approaches for comparison.

Regarding information on what research is lacking, we noted that further studies on model-based capabilities to help with inter-calibration are needed. For example, such features of spatio-temporal structured modelling can:

- Smooth across space to handle missing stations/areas
- Smooth across time to handle missing years, also spatio-temporal smoothing is possible
- Estimate “catchability” covariate effects which can be used on different vessels/gears
- Use “habitat” (environmental) covariates to improve predictions in unsampled regions
- Use “spatially varying” covariate to predict density from annual environmental index, including into unsampled regions
- Fit survey and commercial CPUE data simultaneously (experimental)

The group identified a key activity to help with calibration issues as applying model-based approaches to existing experimental data and estimates within the stock assessment. For instance, comparing “vessel-effect” from VAST with coefficients from estimators highlighted in Pelletier (1998) and Miller et al. (2010). Using this approach one could directly compare the ability of model-based methods to calibrate against physical measurements in real surveys. This approach could be extended to do so in different time-series configurations in order to delineate when model-based methods are reliable vs when physical calibration would be required. For instance, consider these temporal configurations:

**TIME/SPACE**

**Scenario A**
- Vessel 1
- Vessel 2

**Scenario B**
- Vessel 1
- Vessel 2

**Scenario C**
- Vessel 1
- Vessel 2

Analogous work on simulated data would also be informative, and these simulation models could be tailored to specific situations and used in power calculations for how much effort, and how to optimally allocate it, is needed for physical calibration studies.

In combination, this research would provide insight into how evaluate and approach a potential change in survey design and the value of a physical-based calibration experiment. We hypothesize that different types of changes could be divided into roughly three groups: (1) no advantage over model-based approaches; (2) advantageous if highly valued species or difficult circumstance; (3) absolutely required because a model-based solution fundamentally unable to provide reliable intercalibration estimates. This grouping would provide a broad baseline against which survey changes and the resulting need for physical vs. model-based calibration would be required.

The group noted that VAST may have issues with detecting a shift from systematic to stratified-random sampling. Further explorations here are needed. Changing from systematic to stratified-random station locations may be problematic in a practical sense. For example, in the EBSBTS 20 mile distances between stations correspond well with sample processing times and a stratified random sampling might result in more travel time between stations or shorter times when the vessel is waiting for the catch to be processed before making another set. An idea to test might be to compare a resampled survey systematically (e.g. use historical data subsampled in a way to make systematic and “random” (but probably not preferred over simulation approach).
2.3.6 Recommended future directions. Collaborative research and tool development projects

The highest priority would be to evaluate when a model-based approach can avoid or minimise expensive calibration experiments. Steps/tools needed include:

- Development of robust simulation tools including the diverse nature of different species and priorities.
- Identify modelling approaches (e.g. spatial or time-varying) most suitable (links to TOR 1).
- Collating test cases on past experiments.
- Identify highest priority survey reduction types (spatial, temporal, etc.).
- Identify conditions where species characteristics are most well suited calibration studies.

2.3.7 Summary of discussions

The discussion started around whether “calibrations” are most appropriately done within a stock assessment framework or outside. We reviewed gears used for surveys to highlight if there are differences in inter-calibration approaches. The gears and applications included:

- Trawl
- Longline / line
  - Alaska Fisheries Science Center, Southeast Fisheries Science Center, Southwest Fisheries Science Center, Pacific Islands Fisheries Science Center, CDFO
- Trap/pot
  - West Coast CDFO, crab stocks in Alaska
- Acoustic-trawl
  - Walleye pollock (GOA and EBS), Pacific hake
- ROVs / submersibles
  - Cases for Salish Sea and outer US West Coast
  - “Mosaic” approach used for relatively stationary species (yelloweye rockfish) in SE Alaska
- Cameras (stationary)
- Towed bodies (e.g. HABCAM for Northeast Fisheries Science Center for scallops)

The group noted that issues related to calibrating surveys when there are spatial gaps differs from when there are gaps or changes in a time-series of data. We noted that time-series issues might be more easily dealt with within an assessment model. The group then discussed issues in common for some bottom-trawl surveys including:

- Year-effects (for age-specific indices, but also across species).
- Whether index uncertainty estimates are defensible, to inform potential rampant data weighting within assessment models (e.g. estimating additional variance or assuming the stock assessment model knows best).
- VAST applications may be inappropriate for surveys that are adaptive (e.g. Lauren Rogers) or based on habitat (Pacific saury surveys conducted by Japanese scientist which bound area covered based on temperature).
• Combining gear types to address fish availability
  ○ Behavioural patterns that affect availability within the water column (e.g. Walleye Pollock, Boarfish).
  ○ Alternative gears to evaluate untrawlable habitat (e.g. cusk, rockfish).
  ○ Combining US-Russian for stocks and surveys straddling the maritime boundary.
  ○ Pacific hake example (MSE case) but less of a survey case (RV Ricker-RV Shimada case; calibrated acoustic data are combined (treating it as a single vessel collecting the data), often overlapping in length-based strata are combined and often transects are interleaved in areas of southern British Columbia and northern Washington State).
  ○ In Celtic sea surveys are combined to cover a broader range (age and spatially) of the stock.
  ○ Adriatic hake combines data over many countries.

• Composition data
  ○ Identified as being important for detecting changes in selectivity and availability. Can be treated similarly to multispecies issue.

• Calibration when reductions occur
  ○ Spatial coverage vs thinning (chop or thin)
    ■ Balance consideration needed for transit times etc.
    ■ Chopping inadvisable, best to have at least some samples to help w/ model-based approach.
    ■ Thinning better, but see 1st bullet.

The group discussed design issues, e.g. converting from systematic to random or other sampling design. The extent that calibration experiments are needed likely varies on a case-by-case basis and should be evaluated. The model-based simulations may provide an appropriate way to evaluate needs.

2.3.8 References


2.4 TOR IV. “Decision-making Tools”: Develop methods that can provide quantitative decision-making tools describing the impacts on the quality of the survey deliverables and advisory products.

Sub-Group Leads- Curry Cunningham, Sven Kupschus, James Thorson,
Participants: Lyle Britt, Meaghan Bryan, Elaina Jorgensen, Lauran Rogers, Paul Spencer, Nicola Walker, Cynthia Yeung

2.4.1 Summary of existing methods and literature

For over 100 years, fisheries science has developed quantitative tools to produce survey products and information to inform fisheries management. Advice to management typically involves applying a sequence of models, where the output from one step is then one of several inputs to the next. Examples of modelling applications and tools include tools to plan field sampling operations (optimal design models), interpret field-sampling operations (i.e. gear performance and area swept mensuration models), process data (index standardization models), population dynamics (stock assessment models), and evaluate management actions (harvest control rule models). This fisheries-science enterprise has resulted in a large number of tools that can be used to predict the consequences of changing survey effort (whether increases or decreases) and design (i.e. changes in footprint, sampling density, and frequency).

In Table 2.4.1, we list some of the methods that were presented at WKUSER, or published elsewhere, that could be used to provide information regarding the likely consequences of changing survey effort and design upon survey products and management information. We categorize methods based on their inputs (what information is used by the model) and outputs (what information is then derived or estimated from those inputs). We specifically use the following five categories to define inputs and outputs:

1. Cost/logistical constraints and survey designs;
2. Samples and operations-level data arising from samples;
3. Survey productions such as abundance indices or annual physical conditions;
4. Population/ecosystem dynamic models;
5. Management advice and consequences (see section 3.6 below);

We apply this functional classification to approaches, while noting that a given approach may have been implemented independently by multiple WKUSER speakers, or have multiple software implementations that implicitly make different assumptions about the data collections used. Consequently not all applications within a method are equally suitable in all situations.

For each of these five levels of analysis, we highlight the multiple units of analysis that are applicable for each level. For example, models used to evaluate changes in sampling design on survey products may illustrate impacts upon resulting abundance indices, abundance-at-age indices, compositional data, physical data, or other products. We therefore classify each model by the general units of analysis that they involve for a given level of analysis. This approach for categorizing approaches/models therefore allows future research to be conducted by linking different approaches in a modular fashion to the desired output level (i.e. potentially combining different candidate models to span from survey product to management performance). This modular approach provides the additional advantage that where there are regional differences in the monitoring strategy or management metrics, the cross-regional utility of the tools can be
maximised by sharing development where possible, while retaining flexibility for regional applicability.

**We use this classification to help highlight the following conclusions:**

Different models are either targeted to a single level, e.g. a spatio-temporal operating model for understanding how changes in level 2 (number of samples) affects level 3 (abundance indices), while others are expansive and cover multiple levels, e.g. management strategy evaluation for understanding how changes in level 3 (survey products) affects level 5 (likely management consequences).

Estimation models can often be used as operating models to explore impacts upon data quality on a given level. Similarly, most analyses can also be implemented by resampling or subsetting data, i.e. without specifying an operating model structure. The conclusions from either analysis type should consider the difference in assumptions between the approaches and their consequences for risk, and ability to meet specific analysis objectives.

The proliferation of spatio-temporal estimation models has allowed for a concurrent increase in spatio-temporal operating models, which can be used to explore the transition from operation-level samples to survey products. Similarly, the long history of assessment modelling allows a wide variety of methods for exploring the impact of changing survey products on the variance of stock-assessment estimates. However, there are fewer examples of estimation and operating models for exploring the impact of changing stock-assessment variance on management consequences; the existing literature on this topic has focused on harvest control rule development rather than survey design.

We also note that some approaches (and associated model implementations) may be more or less appropriate for evaluating a given change in survey design. For example, evaluating the likely consequences of expanding the spatial footprint of a survey cannot be easily evaluated simply by subsampling or subsetting available data, and instead will likely require an operating model and explicit assumptions about the data-generating processes outside the current range of data. These assumptions would then be developed based on results from process research (i.e. special collections on field surveys, meta-analyses, laboratory experiments, etc.), where the design of future process research could similarly be optimized by expanding the modular methods discussed here.

In addition to the ecosystem conditions under consideration, the sampling characteristics (including the existing survey sampling design, the collection of age and length data etc.) have impacts on the appropriateness of the analysis methods, which should be used to evaluate changes to surveys. For example, surveys where the inter-sample distance is greater than the scale of variability or where the sample probability is disproportional to the area such as fixed-station designs do not lend themselves to evaluation by operating models parameterized only using historical data, but could again be informed by process research.
2.4.2 Summary of the workshop talks and new methods.
Which survey data products need to be included?

WKUSER participants have included short summaries of talks in tabular format (Table 2.4.1), and we have also classified each presentation following categories defined above (see Table. 2.4.2 column: “WKUSER Presentation”).

A wide range of talks were presented during the workshop, describing the effects of alterations in survey effort or design on various quantities and analyses, including (for example) environmental indices and single-species biomass indices through management strategy evaluations (MSEs) and high-level management decisions on the relative priority of surveys within geographic regions. Some important analyses on alterations of survey effort on short time-scales were not included in the workshop and may largely exist in the grey literature, such as operational analyses on the feasibility (both logistical and financial) of survey designs, and analyses to evaluate the real time impact of dropping of survey stations in response to field conditions. Closed loop MSEs are a common fisheries science tool but only a few have focused on survey effects (e.g. Mutniczak et al., 2018; Harford et al., 2015); however, a MSE evaluating survey frequency was represented in the workshop by an application to the Pacific hake. Several workshop talks used simulation models to evaluate the effects of survey effort alteration on the accuracy and precision of assessment output. Finally, spatial-temporal models have become a standard tool for characterizing the spatial dynamics of marine populations, and several presentations in the workshop used these models to simulate various spatial strategies for altering survey effort.

To estimate the effects of altered survey effort or design on survey products, two approaches dominated: either (1) to simulate an assumed population distribution using models conditioned on past survey observations from which “new” samples were subsequently generated under different spatial allocation of simulated survey effort, or (2) through subsampling or resampling of existing data. The former is conditional on the specification of the model used to simulate data, while the latter restricts the exploration of survey options to the existing survey realm, and to the suite of past environmental, demographic, and species distribution conditions, and as such limits the scope of application and inference.
### Table 2.4.1. Summaries of oral presentations and literature and relevance to quality metrics.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Data type</th>
<th>Factor affected by survey reduction</th>
<th>Approach</th>
<th>Metrics</th>
<th>Conclusion</th>
<th>Recommended actions</th>
<th>Research gaps identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kotwicki, S and Ono, K. (2019).</td>
<td>Abundance indices from fishery-independent trawl surveys, for EBS pollock</td>
<td>Abundance index variance, as influenced by mean and variance of sampling efficiency</td>
<td>Simulation: Model-based with spatio-temporal operating model and design-based estimators</td>
<td>Relative errors (true vs sample mean and SD); CV (sample SDs)</td>
<td>Mean and variance of sampling efficiency affects accuracy of abundance index variance estimates, so the reliability of indices depends on how well assumptions of design-based estimators are met</td>
<td>1) More prevalent and thorough assessment of total survey uncertainty; 2) Refinement/re-evaluation of the use of sampling CVs as weights in assessment models</td>
<td>1) Few studies evaluating how catchability contributes to uncertainty in abundance indices, and how this influences stock assessments and comparisons of survey designs/effort levels; 2) Need studies evaluating how spatial availability influences uncertainty in abundance indices, and consequences of this</td>
</tr>
<tr>
<td>Ono, K. WKUSER Talk.</td>
<td>Abundance indices from fishery-independent trawl surveys</td>
<td>Sampling intensity and design, as it affects abundance index mean and variance (with potential use for measuring effects on estimating changes in species distribution and so forth)</td>
<td>Simulation: Model-based with spatio-temporal operating model</td>
<td>Bias, CV, RMSE, relative error</td>
<td>Spatio-temporal models can generate realistic representations of spatial and temporal variation in population densities of groundfish around Alaska</td>
<td>Use a full simulation framework with operating and estimation models so that you can assess the quality of estimates relative to &quot;true&quot; values (e.g. precision and accuracy of abundance indices)</td>
<td>1) Multispecies models; 2) Diagnostics for assessing model fit and realism; 3) Simulation of future environmental scenarios</td>
</tr>
<tr>
<td>Marshall et al. WKUSER Talk.</td>
<td>Abundance indices and compositional data from fishery-independent acoustic surveys for hake</td>
<td>Survey frequency and it's influence on estimates of SSB</td>
<td>Closed-loop simulation for MSE (incorporating annual stock assessments, harvest control rule), but empirical approach to evaluating the effect of</td>
<td>Relative error (true vs estimated SSB)</td>
<td>With less frequent surveys, increase in variability and decrease in median SSB and catch (also increase in frequency of fishery closures)</td>
<td>MSE is a useful framework for evaluating consequences of survey reduction, but need to consider how features of operating model affects results</td>
<td>1) Translate into monetary consequences</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
<td>Methodology</td>
<td>Results</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamel et al. (in prep)</td>
<td>Abundance indices and compositional data from fishery-independent trawl surveys (for several groundfish)</td>
<td>Survey frequency and intensity (annual and seasonal scales, dropping passes/years/vessels: amounting to 50% reduction) and it's influence on estimates of SSB</td>
<td>Simulate different assessment models but with empirical approach to dropping input data from surveys (bootstrapping from past survey data)</td>
<td>Relative change in estimated biomass and SSB from status quo, percent change in OFL (management advice)</td>
<td>Effect on SSB varied among species, but can be quite substantial (varying from ~20%-200%). Also a large effect on recruitment signal. Best option was to drop a pass. Results are species and assessment specific, difficult to draw general conclusions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laman et al. WKUSER Talk.</td>
<td>Abundance indices from fishery-independent trawl surveys, for GOA sablefish, Pacific cod, arrowtooth flounder, POP</td>
<td>Sampling intensity as it affects abundance index mean and variance</td>
<td>Simulation: Model-based with spatio-temporal operating model and design- and model-based estimators (given stratified random sampling)</td>
<td>Bias, CV, RMSE</td>
<td>Uncertainty increases with reduced sampling intensity, but specific results were species-dependent: Largest drop in performance was from 1 to 2 vessels (for POP) Mostly TBD for specific recommendations, but generally the implication is that decreasing effort increases CV at rates that vary by species More evaluation of model-based estimators; further exploration of confounding of survey coverage and sampling intensity, and recommendations will depend on guidelines for acceptable level of abundance index CV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambert WKUSER Talk.</td>
<td>Abundance indices from fishery-independent beam trawl surveys, for plaice in the Celtic Sea</td>
<td>Sampling intensity and location as it affects age-specific abundance indices and subsequent assessment outputs</td>
<td>Simulate different assessment models but with empirical approach to dropping input data from surveys (jackknife, dropping stations with lowest leverage on SSB estimate)</td>
<td>Leverage, SE of SSB from assessment</td>
<td>Removing up to 60 percent of sites had little impact on assessment Need guidance on whether to stop with biomass index, or go all the way through to assessment, translate dropped sites into savings in terms of time ($$$); through discussion-need multispecies perspective to evaluate trade-offs in information provided by stations among species/questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Thorson et al. 2019 (in review)</strong></td>
<td>Age composition</td>
<td>Ageing intensity</td>
<td>Analytical framework: (1) bootstrapping age samples, simulation testing with VAST, (2) simulation testing with a pop dyn model and assessment sample size ratio (n input to n effective), where n effective follows a monotonic saturating function (Michaelis-Menten relationship)</td>
<td>One can calculate the diminishing returns of ageing for assessments</td>
<td>Proposed workflow for the value of ageing analyses</td>
<td>Explore theoretical relationship between n effective and stock assessment variance; apply to specific species; identify general conclusion</td>
<td></td>
</tr>
<tr>
<td>Jorgensen et al. WKUSER Talk</td>
<td>CPUE from fishery-independent groundfish bottom-trawl surveys in the Eastern Bering Sea</td>
<td>Sampling spatial coverage (funding and vessel breakdowns)</td>
<td>Empirical (dropping stations with lowest information to minimize RMSE of design-based estimate of abundance)</td>
<td>Information criterion: mean squared + variance of CPUE; Performance metrics: bias, RMSE, CV; look at contribution of each station to each species</td>
<td>Threshold relationships between number of stations removed and changes in bias, cv, and RMSE; number of stations acceptable to remove varies among species (most for species with a smaller footprint of distribution like yellowfin sole), and perhaps over time with changes in bottom temperature</td>
<td>If faced with forced reduction in survey effort, can possibly drop stations, mostly in the northeast portion of the eastern Bering Sea shelf</td>
<td>Perform more formal multispecies synthesis/prioritization of stations, may need to expand information criterion to include spatial covariance</td>
</tr>
<tr>
<td>Munro et al. WKUSER Talk</td>
<td>CPUE from fishery-independent groundfish bottom-trawl surveys in the Eastern Bering Sea</td>
<td>Ability to detect abundance trends</td>
<td>Simulated trends and empirical: Non-parametric bootstrapping, comparing estimators of trend in density (mean, median, geometric mean, Kappenman)</td>
<td>RMSE of annual change</td>
<td>Species-dependent results: for pollock, all estimators have similar efficiency, but Kappenman is slightly better; Kappenman may be much better for species with more constrained distributions within the survey area (e.g. yellowfin sole) and/or those with more skew in CPUE; sample mean</td>
<td>Can achieve unbiased estimator of trend even from biased CPUE estimators</td>
<td>Need to generalize across species/regions</td>
</tr>
</tbody>
</table>
Spencer et al. WKUSER Talk.

<table>
<thead>
<tr>
<th>Biomass and CV of Biomass</th>
<th>Sampling intensity as it affects abundance index mean and variance, and how this influences SSB estimates in stock assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation: operating model with true trend in biomass determined by past assessment output, distributed spatially according to semi-variogram fit to past survey data, simulate SRS surveys and pass design-based abundance index and CV to stock assessment where only input is abundance index (tier 5 stock)</td>
<td></td>
</tr>
</tbody>
</table>

1) Propagation of survey uncertainty depends on observation error, process error variance and underlying abundance trend; 2) Sample CVs increase with decreasing sampling intensity; 3) This inflation of CVs reduces accuracy of estimated biomass and trend from stock assessment model when these CVs are used as weights; 4) At low sampling intensity, biomass estimates from the assessment track the survey more closely than they probably should (as sample CVs are underestimates of true survey CV)

Preliminary implication is that abundance index CVs would need to be below 0.2-0.3 to achieve one measure of acceptable accuracy of biomass and biomass trends in simple assessments (for pollock, Pacific cod, and Pacific ocean perch)

von Szalay et al. WKUSER Talk.

<table>
<thead>
<tr>
<th>Abundance indices from fishery-independent trawl surveys, for GOA sablefish, Pacific cod, arrowtooth flounder, POP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abundance index mean and variance</td>
</tr>
</tbody>
</table>

Simulation: Model-based with spatio-temporal operating model and design- and model-based estimators (given simple random sampling)

Bias, CV, RMSE

Model-based estimation with simple random sampling may provide best abundance estimates for key species of interest, but simple random sampling with design-based estimator does okay too

1) Interpret variance of abundance indices for some species with caution, as uncertainty is high; 2) Model-based estimators should be implemented

1) Explore influence of assumed response distributions in model-based estimators
Williams et al. WKUSER Talk.

Length distribution data from camera vs midwater trawl methods

loss of weight and age data with open codend tows

resample survey data [remove 50% or 66%], removing weight, age, sex, maturity information (keep length)

Proportional change in empirical biomass estimates, age-specific relative abundance, and sex ratio relative to status quo sampling

for smaller survey (GOA) Age-specific relative abundance estimates are sensitive to sampling intensity, for larger survey (EBS) with more intense baseline sampling, less sensitivity

1) camera could provide "emergency backup" role to traditional survey methods in the case of reduced sampling resources

1) survey trackline optimization; 2) combining multiple sources of uncertainty to calculate total survey uncertainty;

Conner et al. WKUSER Talk.

Abundance indices from fishery-independent trawl surveys, for Eastern Bering Sea groundfish

sampling intensity as it affects abundance index mean and variance

Simulation: Model-based with spatio-temporal operating model and design-based estimators (comparing sampling designs to mitigate effects of survey reduction)

relative estimation error; RMSE, relative bias

1) Systematic designs are most biased and have most error in their estimate of biomass and its variance, where magnitude and direction of bias depends on species (except for diffusely distributed species like pollock, where the systematic performs slightly better). Systematic also introduces the most variation in bias. 2) Stratified random design can reduce variance in some cases (performs well for patchily distributed species like yellowfin sole). 3) Trade-off of bias and variance in bias

1) Do not use systematic design when index standardization is primary goal (generality of this to be further explored)

1) Evaluate how differences in spatial distribution properties among species influences these outcomes; 2) application of cost estimation related to transit times between stations

Blaine et al. WKUSER Talk.

Abundance indices from fishery-independent trawl surveys, sampling intensity

bootstrapped biomass densities, used NMDS to compare survey designs

percent difference in biomass estimates and cvs

1) higher error rates and more variability, and this increases with a finder scale (stratum-purpose of survey is important to consider (if you don't need biomass estimates, the
<p>| Rogers and Mier | Larval CPUE from fishery-independent midwater trawl surveys in the Bering (pollock) | sampling intensity (density, coverage) in its influence on abundance index | 1) remove stations in alternative ways and compare to full dataset, 2) simulation study to evaluate new adaptive design | CV, RMSE between sampling alternative designs and &quot;all data&quot; | 1) perform adaptive sampling to extend coverage when indicated (for pollock) | 1 more exploration of model-based estimators | 2,4 |
| Bryan and Thorson | CPUE from fishery-independent groundfish bottom-trawl surveys and bottom temperature or cold pool extent in the Eastern Bering Sea | simulation of spatio-temporal operating model, where all random fields are recreated each iteration; model-based estimation model with same structure | MAE on the spatially varying coefficient of the environmental covariate, and biomass | 1) spatially varying effect of cold pool on groundfish abundance/distribution can be well estimated by model-based tools (VAST), even given less data than status quo | flexible tool to evaluate other survey designs | 1) Evaluate alternative survey designs | 4 |
| Cunningham | biomass indices, SSB, reference points | bootstrapped data inputs from past survey data, use model-based estimator for abundance index, see how changes propagate through stock assessment model | differences in index, index CV, change in SSB/depletion/b40/F40 among scenarios | Results depend on way the data is removed (whole strata vs random stations removed) and the species considered. But, in general: 1) index estimate is not affected, CV increases with less data, 2) more variability in the SSB (with 50 and 75 percent reductions) and B40 than for F40 and F35 | 1) optimal strata to drop depends on species, will require prioritization or multispecies optimization | Need to consider implications of loss of age comp data in addition to biomass index; use operating model with known true values | 2,4 |</p>
<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
<th>Samples/Intensity</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kupschus</td>
<td>Biomass indices sampling intensity</td>
<td>bootstrapped data, reduce samples, reduce number of individuals sampled</td>
<td>failed assessment models when tow duration reduced, conflicts in the data, method of index calculation matters more than the number of samples. Multivariate methods provide ecological structure for strata, weight this against other options like getting better info recruitment indices 1) investigate conflict between catch and survey data resulting in different abundance trends; 2) more multivariate methods; 3) more case studies to generalize</td>
</tr>
<tr>
<td>Walker, N.</td>
<td>Biomass indices sampling intensity (seasonality)</td>
<td>1) drop one season from all years or 2) drop one country's data, and evaluate effect on stock assessment model output</td>
<td>Dropping the season-specific index did not affect results for 2 of 3 species, but did have large effect on a key stock; However, indices for recruits were estimated with more error (?); increased observation and reduced process error estimates, cascades to biomass estimate from assessment to influence recommended catch by ~6%</td>
</tr>
<tr>
<td>Richar</td>
<td>Abundance indices and patch size estimates from fishery-independent trawl surveys, for Eastern Bering Sea crab</td>
<td>empirical, dropping stations, dropping patches, dropping even/odd years</td>
<td>Uncertainty increases, but point estimates not strongly biased except for minor stocks; reducing survey could hamper rebuilding of stocks?</td>
</tr>
<tr>
<td>Yeung, C.</td>
<td>Bottom temperature data from fishery-independent trawl survey coverage, its effect on an oceanographic index (the cold pool index)</td>
<td>empirical, spatial effort reduction by 50% in different ways (random)</td>
<td>Error in bottom temperature estimates with even 50% reduction is ~ +/- 2 degrees celcius Effect of change in precision and accuracy of cold pool index on the products it is used to predict (e.g. abundance and dis-</td>
</tr>
</tbody>
</table>
survey in Eastern Bering Sea, checkerboard, grid column

Table 2.4.2. Characteristics and relevance of survey quality metrics.

<table>
<thead>
<tr>
<th>Name /Description</th>
<th>Description</th>
<th>Input¹</th>
<th>Output¹</th>
<th>Type of cost or sampling design constraint</th>
<th>Type of samples (see also²)</th>
<th>Type of survey products (see also³)</th>
<th>Type of population/ecosystem model estimates (see also³)</th>
<th>Type of management consequence (see also³)</th>
<th>Number of species²</th>
<th>Existing Tools³</th>
<th>Citation</th>
<th>WKUSER Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vessel operations submodel</td>
<td>Cost–benefit analysis of survey designs</td>
<td>Cost constraint</td>
<td>Sampling design</td>
<td>Charter vessels, vessel-days, staff</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>None available</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Advice value submodel</td>
<td>Cost–benefit analysis of management outcomes</td>
<td>Population and/or ecosystem model estimates</td>
<td>Management consequence</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Variance in stock-status estimates</td>
<td>Expected catch, catch variance</td>
<td>Multispecies</td>
<td>None available</td>
<td>Hollowed</td>
<td></td>
</tr>
<tr>
<td>Multivariate spatio-temporal estimation and operating model</td>
<td>Fit to multivariate, spatially referenced data and potentially simulate new data conditional on fits</td>
<td>Sampling design, samples</td>
<td>Survey products</td>
<td>Biomass, Biomass-at-age</td>
<td>Index of abundance</td>
<td>NA</td>
<td>Multispecies</td>
<td>VAST, Ono, INLA</td>
<td>Ono</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Model Information</td>
<td>Modeling Outputs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Univariate spatio-temporal operating model</td>
<td>Fit to univariate spatially referenced data and simulation new data from posterior samples</td>
<td>Sampling design, samples</td>
<td>Index of abundance, index-at-age, age/length data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Samples; Survey products</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td>Single-species VAST, Ono, INLA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kotwicki and Ono 2019</td>
<td>Conner, Laman, von Sza- lay, Kotwicki, Ono, Munro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empirical resampling of survey tows</td>
<td>Resample or subsample survey data to calculate indices or variance of indices</td>
<td>Sampling design</td>
<td>Index of abundance, physical index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Survey products</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td>Single-species None available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>None available</td>
<td>Jorgensen, Lambert, Yeung, Hamel, Rogers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resampling survey data and refitting stock assessments</td>
<td>Resample or subsample survey samples, refit data-processing methods, and refit a stock-assessment model to each set of data products</td>
<td>Population and/or ecosystem model estimates</td>
<td>Stock status, biological reference points, overfishing limit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sampling design</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Population and/or ecosystem model estimates</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stock status, biological reference points, overfishing limit</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stock status, biological reference points, overfishing limit</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expected catch, catch variance</td>
<td>Single-species None available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>None available</td>
<td>Cunningham, Kupschus, Hamel, Spencer, Walker, Richar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manage- ment strategy evaluation</td>
<td>Closed loop simulation of management effects on populations</td>
<td>Survey product</td>
<td>Index of abundance, abundance-at-age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Management procedure</td>
<td>NA</td>
<td>Stock status, overfishing limit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td>Expected catch, catch variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marshall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate-linked operating models for spatial dynamics</td>
<td>Simulation of environmentally linked species distribution and abundance</td>
<td>Sampling design, samples</td>
<td>Sampling design</td>
<td>NA</td>
<td>Biomass, physical</td>
<td>Sampling design</td>
<td>Sampling design</td>
<td>Single-species</td>
<td>VAST</td>
<td>Bryan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>----</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>---------------</td>
<td>------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrapping to predict changes in input sample size</td>
<td>Bootstrap the set of tows and age/length samples within each tow, calculate proportions for each bootstrapped sample, and calculate input-sample size from the variance of bootstrapped proportions</td>
<td>Survey products</td>
<td>Survey products</td>
<td>NA</td>
<td>Bio-mass-at-age</td>
<td>Composi- tion data and input sample size</td>
<td>NA</td>
<td>NA</td>
<td>Single-spe- cies</td>
<td>Being developed for AFSC by SSMA/MESA</td>
<td>Thorson et al. 2020</td>
<td>Thorson, Blaine (multi-species)</td>
</tr>
<tr>
<td>Predict changes in effective sample size resulting from new input sample size using stock-assessment data weights</td>
<td>Survey products</td>
<td>Population and/or ecosystem model estimates</td>
<td>NA</td>
<td>NA</td>
<td>Composi- tion data and input sample size</td>
<td>Weight of composi- tion data</td>
<td>NA</td>
<td>Single spe- cies</td>
<td>Theoretical prediction, tested via simulation</td>
<td>Thorson et al. 2020</td>
<td>Thorson</td>
<td></td>
</tr>
<tr>
<td>Integrated spatial-population operating model</td>
<td>Cost constraint</td>
<td>Management consequence</td>
<td>Sampling design, number of samples</td>
<td>Biomass, biomass-at-age</td>
<td>Index of abundance, compositional data</td>
<td>Stock status, biological reference points, overfishing limits</td>
<td>Harvest control rule</td>
<td>Single species</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>------------------------</td>
<td>-----------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td>-------</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>An operating model that simulates spatio-temporal variation in density at age/length given alternative management procedures, such that it can simulate the consequences of different sampling designs upon stock-assessment variance estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Historically, resampling predominated in the studies that developed simulated survey products, which necessarily could only be compared to existing survey products of unknown quality and frequently did not provide estimates of the uncertainty in the evaluation. Simulations based on distributional operating models were able to assess biases and uncertainty conditional on the quality of the operating model. The latter approach tended to dominate in the studies that evaluated the impact of survey effort changes on the assessment metrics. However, with the modular approach, the data generating (or resampling) process is not a necessary function of the desired level of evaluation. Historically, assessors and modellers tended to follow the simulation approach, while survey implementers tended to use resampling due to interest in the changes to their product. However, this distinction was not borne out in the WKUSER presentations, where some assessors and modellers used resampling methods, while some survey-team members used operating models (e.g. the Ono-Conner spatio-temporal model).

2.4.3 What are the major challenges?

The majority of current analyses presented at this meeting and in the literature have focused on evaluating the impact of changing survey effort and design on individual species. Studies where multiple species were considered, were done without identifying the trade-offs of changing survey strategy relative to one another. An important next step will be to synthesize the results to identify the trade-offs in performance among species, across different sampling strategies. Future analyses should compare the chosen performance metric (e.g. survey CV, bias in biomass estimate, etc.) across species to evaluate the impact of different sampling strategies via their impact on multiple species sampled by a given survey. A qualitative example of this would be to use a multidimensional plot with sampling frequency and sampling intensity on the x-axes, the performance metric on the y-axis, and each species would be represented by a different color. The next challenge would be to determine the relative importance (e.g. fishery value, revenue, ecosystem importance, etc.) of each species and use this as a weighting factor to identify the best sampling strategy for a multispecies survey.

Another potential method for identifying single-species performance trade-offs is MSE (see Marshall presentation, Muridian et al. 2019, Hutniczak et al. 2019). In the special case of mixed species bottom-trawl surveys, the information needs for informed decision-making requires consideration across species and fisheries. Several talks provided case studies showing the merits of examining trade-offs in changes to survey frequency or design across species. Figure 2.4.1 illustrates a potential analytical framework under development at the AFSC to address this issue. This framework builds on output derived from one or more of the simulation tools described in Table 2.4.2. Completion of this framework requires a partnership between the survey analytical team and stock assessment scientists (similar to that proposed in Thorson et al. 2020).

Within the proposed framework the analytical team could use one or more of the tools identified in Table 2.4.2 to identify viable survey scenarios that could be entertained given existing constraints (funding, ships etc.). To fully assess trade-offs associated with changing survey design or frequency, we recommend that status-quo scenarios are always carried forward through the framework for comparison purposes.

The group recognized that simulating the short- and long term implications of changes in survey design or frequency for all species monitored by the survey would be a daunting task. Therefore,
under the proposed framework, analysts would select representative suites of species life history characteristics/distributions for analysis. Selection of suites of representative species should include at least one group that are relatively rare and/or patchily distributed species that typically exhibit relatively high interannual variability and high within year coefficients of variation.

Figure 2.4.1. Potential analytical framework at the Alaska Fisheries Science Centre to address survey effort reduction.

Suites of simulated data streams would be examined to determine expected consequences of changing survey design, effort, and products on expected management performance. This can be done using management strategy evaluation to define an operating model for survey data, resulting survey products, stock-assessment performance, and resulting management consequences. Alternatively, expected consequences can be identified by linking together a sequence of targeted analyses that pursue each of these analyses individually. In either case, it requires some output involving management consequences (see Table 2.4.2). Based on these statistical properties, an operating model would be developed to simulate fish and shellfish populations, surveys, catch-at-age using standard approaches in closed-loop management strategy evaluations (Punt et al. 2014, also see presentation by Marshall). We recognize that the simulated data streams are approximations of those currently used for stock assessment however, these likely represent the best available scientific information for assessing these issues.

Several case studies identified cases with abrupt changes in production, distribution or species overlap in response to changing environmental conditions. Scenarios for future surveys could include projected environmental change derived from downscaled global ocean model projections of future climate change. Spatio-temporal operating models are well suited for this purpose.

Projected outcomes of proposed survey changes could be simulated over a suitably long period under status-quo, and proposed sampling designs. Suites of core performance metrics could be
estimated to inform decision-makers (e.g. Fulton et. al. 2019). Candidates include (but are not limited to) metrics of projected risk for loss of yield, unintentional overfishing, future estimates of biological reference points (e.g. BMSY, FMSY, MSST, MFMT). In cases of environmental change, secondary analyses can be included to assess the implications of changing survey design or frequency in a changing climate.

Extensions of this effort to fully address a region’s effort to implement an ecosystem approach to fisheries management could be informed by outputs from this framework. In this circumstance, analysts would work with fisheries managers (ICES ACOM/FMCs) to develop suites of core socio-economic and fisheries management indicators that US and European agencies are developing. These secondary performance metrics could include valuation of the fishery, community access, sectoral allocation and others.

While examples of how such methods have been implemented numerically were presented at the workshop (Table 2.4.1), application of the same methodology over different survey situations to identify the risks and benefits of different methods under specific conditions is still missing. Before such methods are used to inform on survey strategy, it is vital that the magnitude of multiplicative propagation of the process and observation errors in such linked simulations is evaluated against the scale of the contrast that is likely to occur in response to survey changes. The group acknowledges that there are unknown risks arising from misspecification of modelling assumptions that are used to inform changes in survey operations. In the end, the appropriate tool will likely be dependent on the quality and availability of data, magnitude and sources of uncertainty, overall objectives of the study, and the time frame and resources available for analysis.

2.4.4 Research recommendations. Collaborative research and tool development projects

The description of methods used to evaluate future survey options in this report has tried to identify methodologies based on management objectives. This has led to an aggregation of methodologies based on the data required and the output provided by each method, with the aim of showing how different methods can be combined sequentially to propagate the change in uncertainty associated with different sampling options. The “method grouping approach” taken at the workshop (Table 2.4.1, Figure 2.4.2) emphasizes the perspective of the data user, i.e. what type of data is available. Specifically, it classifies examples based upon the available data. An alternative approach may emphasize the data collector perspective to evaluate the impact of survey changes on management. Future work needs to consider data input types (i.e. surveys) from an inference perspective, particularly sampling design or data characteristics. Not all methods that produce SSB estimates for example can be equally well applied to different survey designs or data collection (e.g. acoustic vs trawl). Some of these elements have been commented in Table 2.4.1 under the “risks and benefits columns”, but cannot be readily aggregated across different survey types to identify trade-offs between different survey design changes. Future effort is needed to better define and evaluate the suitability of different methods under combinations of the existing data, and the scale of change in ecosystem condition or species distribution.

Table 2.4.2 describes some methodologies based on output requirements to appropriately link the evaluation chain. For example the “operational sampling model” is needed to translate
changes in survey design to changes in cost, which is in turn necessary to compare different survey options based on a precision/cost trajectory. Currently, no automated methods exist to provide such cost information and should be developed to allow automated simulation testing. Similarly, further development of methods for quantify cost–benefit trade-offs in socio-economic terms associated with a change in precision of the management recommendations is also necessary.

Particularly important for the long term sustainability of surveys is the development of methods that allow managers to evaluate the “total value” of surveys (including for specific assessments, process studies, political constituencies, economic enterprises, etc.) across different resource or management objectives (e.g. across species, qualitative ecosystem metrics and assessment of physical conditions). This requires suitable prioritization for such metrics, including the value of surveys in research and development, the value of understanding ecosystem responses to management actions, and addressing as yet unforeseen ecosystem threats. Better methods to combine results across multiple methods and species are needed in the short-term to evaluate survey changes while quantitative methods are being developed for situations where time and budget constraints do not allow for a full evaluation, which is in itself a lengthy process even if those methods were available.

Most methods presented and reviewed by WKUSER predominantly evaluated impact of survey changes on commercially exploited fisheries resources through data application: ‘Stock assessment’ and ‘multispecies stock assessment’. Similar methods that can be used to assess more complex management metrics, although we have not evaluated those options.

Tools to identify uncertainty thresholds will be required for all ecosystem metrics and potentially different uncertainty responses under different ecosystem states. Particularly the CV may be much smaller during low abundance periods compared to high abundance periods. Useful tools may consider how various plans to modify survey effort and sampling designs result in changes in the coefficient of variation (CV) of survey biomass estimates (or other survey outputs), and the implications for important assessment and management related qualities such as variability of survey biomass estimates (i.e. input to stock assessments) and variability in estimated biomass (i.e. output from stock assessments). For example, simulation modelling presented during the workshop by Spencer et al. revealed that when low sampling rates are applied to highly skewed delta-lognormal distributions of biomass (a typical condition for marine stocks), the estimated variances of survey biomass estimates may underestimate the variance of deviations between survey estimates and true biomass (i.e. estimation error). Conversely, at high sampling rates the estimated variances of survey biomass estimates are an accurate representation of the sampling variance of estimation error. Analyses of this type can provide a threshold for a level of survey CV, as underestimation of the standard deviation of estimation error should be avoided (particularly in assessment models in which the estimated variances of survey estimates are used to weight individual survey biomass estimates).

The evaluation of uncertainty thresholds of survey data product is also needed beyond the measures such as relative SSB estimates but also to other management metrics such as, for example, fishing mortality to more clearly describe the impacts of survey effort modification at various outputs from the assessment models. The criteria described above focused on the variances of the input data to assessments, and could be compared across different protocols for altering survey effort. A second stage of analysis would describe the associated uncertainty with important
assessment output such as estimated scaled and trend of biomass. A third stage of development would be to relate the assessment output to a common currency of value (considering the economic, ecological, and social value), as this would help evaluate trade-offs of altering survey effort across multiple stocks.

2.4.5 References


Figure 2.4.2. Schematic decision tree for identifying applicable methods for analysis of changes in survey design or effort allocation.
3 Synthesis of Workshop Results

3.1 Summary of each TOR group

Presentations and discussions over the 1st two days of the WKUSER focused on the types of data produced by different types of surveys, the importance of the data to stock assessment, modelling approaches to addressing data gaps, and examples of the application of different simulation and modelling methodologies to best inform approaches to unavoidable survey reduction. There were 29 presentations given that focused on various aspects of surveys primarily in the North Atlantic and north Pacific Oceans. Presentations varied widely with empirical and simulation studies, with multiple survey designs addressing single-species and multispecies sampling. In addition, applications to survey collection of non-fisheries data, such as zooplankton and oceanographic variables were also addressed. In summation, the initial presentations spoke to a wide range of ideas on methods that can be used to minimise information loss, facilitate better contingency planning and convey the consequences of survey reductions to scientists and policy-makers.

During days 3 and 4, the WKUSER has worked in subgroups on different terms of reference. In all subgroups, the discussions and evaluation of presentations led to at least one decision tree for the respective TOR topic. In this summary, the relationship among the main outcomes is described, as well as the gaps following from it.

TOR 1 discussed current methods that have been used for survey reductions and potential considerations that should be included when making decisions on survey effort reductions. The subgroup divided these into two categories, long- (Figure 2.1.1) and short-term (Figure 2.1.2) survey reductions, which have similar decision points, but potentially different considerations and outcomes. Figure 2.1.1 focuses on the possibilities to implement a change in survey effort to a survey (see also section 2.1.5) on the longer term, in a stepwise approach. Figure 2.1.2 presents an additional decision tree for short-time effort reduction, for unexpected and unavoidable events while at sea. The TOR 1 subgroup recommended careful consideration of priorities and trade-offs in multispecies surveys and also recommended conducting process-oriented studies (addressing both spatially explicit and multispecies concerns) that are useful for making decisions on survey reductions (see also section 2.1).

The TOR 2 subgroup addressed total survey uncertainty in the face of survey reductions. Figure 2.2.1 schematizes the steps towards the estimation of total survey uncertainty, which is defined as a comprehensive representation of the accuracy and precision of a survey data product. Estimates of total survey uncertainty are necessary to assess the quality of survey data products and also can be useful in the in stock assessments to weight survey data inputs. The survey uncertainty contains three main pillars: uncertainty due to population processes, observation processes during the survey, and analytical processes on the data. Due to the complexity of the processes affecting survey uncertainty variance of the survey data product estimated from samples alone often does not reflect total survey uncertainty. Therefore, it is necessary to perform separate studies to estimate total survey uncertainty. The TOR 2 subgroup recommended that research be conducted to partition uncertainty into its different components (e.g. gear related, habitat related). The group also recommended that realistic simulations be conducted to addressing
where uncertainty is propagated through models and decisions and finally, the group recommended that survey reductions be balanced with research to quantify uncertainty where possible. More information on survey uncertainty is available in section 2.2.

The TOR 3 subgroup addressed questions of when and how surveys can be calibrated. Much of the discussion centred on when models could be used for inter-survey calibration and at which point experimental calibrations were best used. Figure 2.3.1 describes the need for physical or modelled inter-survey calibration (see also section 2.3). Inter-survey calibrations can be used to obtain either an index developed depending on multiple vessels and countries or to deal with changes in survey design. Effort reduction may lead to changes in survey design and inter-calibration may be needed in order to use the data from the survey over time. The TOR 3 subgroup recommended that proxy variables can be identified and used for calibration within modelling approaches and that when possible, both design-based and model based approaches to calibration be used and compared.

The TOR 4 subgroup was tasked with identifying decision-making tools for survey reductions including defining how reductions might impact the quality of survey deliverables. Figure 2.4 provides an overview of applicable methods for stock status assessment after reduction of survey effort. The TOR 4 subgroup also assembled a decision Table 3.4.2 that describes in detail methods, trade-offs and implications of survey reductions. The group recommended pursuing research to automate cost/benefit analysis of survey reduction using appropriate measurement metrics for surveys with multiple objectives, species and ecosystems. The group indicated that these tools should define thresholds for survey reduction impacts (such as where reductions caused increases in uncertainty that rendered the survey useless for stock assessment) using a common currency. For further background, see section 2.4.

### 3.2 Relation between the decision trees

The outputs of the TOR subgroups were a group of decision trees and tables that formed step-by-step instructions that could be followed during periods of survey reduction. The trees were interlinked by common concepts and pathways for decision-making (Figure 3.1). For example, the pathway to consider long term reductions in survey effort led directly into questions of whether a calibration would be needed in order to provide continuity of the time-series for stock assessment, or whether the types of changes incurred by the survey reduction could be addressed more easily through application of a method such as spatial-temporal modelling that could join the two time-series. Similarly, the ‘Applicable methods’ in Figure 3.1 informs the first step in Figure 2.4.1. In that same step, evaluation of total survey uncertainty (Figure 2.2.1) should be taken into account. The total survey uncertainty estimate furthermore provides information on the impacts of the short-term reduction in survey effort (Figure 2.1.2). Together, the decision trees can form a pathway from a proposed survey reduction to a series of questions and tools that can be used to evaluate the quality of the impacts on the resulting data and decision-making.
Figure 3.1. The relationship among decision trees.

3.3 Interpretation of quantitative outputs

The methods identified by the outputs of the WKUSER workshop are widely applicable across various survey reduction, change, or expansion scenarios. The methods can address the complete loss of data or the necessity of making do with less data. Methods also address when a survey can be continued with only a minor adjustment versus the requirement of a total recalibration and redesign of a survey.

Despite the fact that methods for evaluation of survey reduction provide quantitative output it is difficult to give predefined acceptable levels of survey data product uncertainty. The acceptable survey uncertainty may be different for a situation where data from multiple sources are available than for data where only the survey provides information on a specific stock. Likewise, a shift from one survey design to another may occur even when the inter-calibration ability is not optimal for specific cohorts or species while the results of inter-calibration for other cohorts or species are good.

Along the same lines it is not possible to a priori name the ‘best’ survey design, as the combination of e.g. available ship-time, funding, vessel size (and so: numbers of staff on board), objectives, target species or areas, may in the end lead to an achievable survey design. There are trade-offs to each of the approaches (see TOR 4 section 2.4 for details) and these trade-offs are occurring along multiple axes.

One of the key points from this workshop with regards to interpretation of outputs from survey reduction exercises was the need to define the evaluation process relative to survey objectives. If the objective of a survey is to produce a single species index for stock assessment, the objective of any evaluation of reduction of the survey will likely be simpler than for the multi species or ecosystem survey. In this case, a simpler analysis stopping only at assessing the impact of survey reduction on uncertainty for the single species biomass estimate might be adequate. However, if
the survey data are collected to inform ecosystem indicators that then feedback into stock assessment through recruitment processes or ecosystem productivity estimates, a more complex objective and analysis may be needed to fully realize the impacts of survey reductions on the data. In other words, clarity in purpose for the survey and how the resulting data are used is important when evaluating the impacts of survey reductions.

3.4 Scope for further development

During the workshop, most presentations and discussions focused on the effect of reduction on one or a limited number of species in a survey. In general, these were the most commercially important species. To broadly apply the methods to (fisheries-independent) surveys, development of evaluation tools for multiple aspects (e.g. species, age groups, ecosystem information) of a survey has to become available. In addition, the evaluation tools need to have as outputs a common currency, for example a currency that is expressed in terms of increased uncertainty or potential loss of fishery yield. Answers to these questions were not directly addressed during the workshop, however, there was broad agreement that future research and development should examine these topics. Suggestions for future research and action include several actions on the short-, mid-, and long term:

**Short term**-

- Decision trees for survey managers in areas and for survey types which were not covered in subgroup meetings
- Development of universal survey analysis tools to process survey data. Assure that tools can be used across different survey types and databases (e.g. R packages, augment existing code from researchers and make it universal for the processing survey data and/or estimating reductions impacts. Make analytical studies easier. Continue to streamline the process of QC of the survey data.
- Importance criterions for survey locations. Use information from the past, e.g. densities, variance, spatio-temporal covariance to identify consistent patterns to allow for lower sampling density. Use of the information from the past surveys to allocate effort relatively to the expected information contents. Adaptive sampling based on previews of expected distribution.
- Developing methods for dealing with elimination of entire survey areas (e.g. due to area closures to trawling, wind farms, high commercial fishing densities)
- Use of technological developments to improve ability to collect more data with higher accuracy (e.g. automated age reading)
- Increase ecosystem data collection during survey for use in model-based estimation and process studies
- Use simulation studies to obtain approximate estimates of total survey variance and explore causes of additional variance in survey data inputs estimated within assessment models.
- Perform easy calibration exercises during existing surveys to test different designs using the same gear.
- Continue exploration of model-based estimation in challenging situations.
Mid term-

- Multispecies/multi-objective optimization, trade-offs, approaches. How to agree on a common currency for what you are optimizing (what is the composite metric)? How do you get stakeholders to agree on weights for different species? Need of increased/appropriate feedback from stakeholders (e.g. year-to-year). Look at the multivariate matrix to weigh between environmental and biological objectives. How do we balance survey objectives (e.g. focus particularly on less abundant species, environmental information).
- We want to be prepared for the ecosystem change. How to embed more ecosystem process studies to inform catchability and spatial dynamics studies. How to implement the use of novel technology in surveys. Develop methods for coming up with designs that allow flexibility in options and strategies in survey effort allocation between years areas to maximise information from surveys.
- Identify decision-makers to get feedback from to clarify objectives and priorities in conjunction with development of methods to evaluate survey changes across multiple objectives. Augment a structure of advisory process to assure feedback to survey groups (e.g. SSC feedback provided to AFSC survey group in 2018) on priorities with respect to importance of the areas and species to direct survey effort to accordingly to the needs.
- Test model-based flexibility to different sample allocation strategies. Flexible designs should provide the same expected values but can differ in variance estimates.
- Look for correlations between environmental covariates and densities (occurrence) or catchability. Strive to increase your data collections, e.g. collect easy to collect environmental data such us for example pictures of the bottom. Develop tools to process on the spot automatically.
- Consider smarter stratification and sample allocation schemas to achieve better precision from a given effort. For example: change sampling densities within strata of existing surveys in response to expected changes in distributions.
- In the effort to address distribution shifts consider research on expanding surveys into the new areas where species are moving to by for example stretching existing sampling effort or doing something different.

Long term-

- Develop technological methods to obtain absolute estimates of biomass or abundance to calibrate existing surveys and obtain estimates of survey catchability and variation in catchability
- Develop methods to incorporate open codend tows with cameras for some stations. Consider technologies that could reduce the number of tows to be processed by hand.
- Test Artificial Intelligence or other multivariate approaches to discovering predictable spatial relations to aid estimation for advisory products.
In addition, there are major logistical challenges associated with limited resources, administration, governments, and boundaries. Some issues include:

- Limited number of staff to do the work. This can affect survey group’s ability to complete full surveys.
- Organizational divide between data collections and assessment programs. Organizations should create mechanisms for the survey and stock assessment scientist to have productive communication and collaboration. The example of such collaboration in the work of the WGISDAA and WKUSER.
- Limited number of research boats and boats available for charter. Boats bidding for charters are old and in the future it is expected that their number may be reduced.
- There are serious logistical challenges with expanding surveys to new areas to follow fish populations. Developing new surveys, especially in the far north, is challenging because there is very limited information on the distribution of species in these previously under-sampled areas and habitats.
- Problems with fish moving across survey and country boundaries creates logistical challenge because not all surveys can be easily modified due to jurisdictional issues.

3.5 What’s next?

- WKUSER participants agreed that there is a need for a second WKUSER. The issues that were identified during this first meeting are of very complex nature and there are no easy solutions that can be applied across the board to all surveys. The research coordination and cooperation is needed to assure progress in the critical research areas as defined by WKUSER TORs or by refined TORs in the future.
- WGISDAA will discuss the necessity for the next WKUSER and the group meeting in October 2020. WGISDAA will also discuss prioritizing a list of challenges remaining and consider creation of the new group within ICES or asking existing groups for help to address challenges the lie ahead for fisheries independent surveys.
4 Abstracts from Presentations

Blaine, Jen  Washington Department of Fish and Wildlife

WDFW Puget Sound scientific bottom-trawling: sampling design changes and consequences

Jen Blaine, Robert Pacunski, and Dayv Lowry

Since 1987, the Washington Department of Fish and Wildlife (WDFW) has been conducting bottom-trawl surveys throughout Washington waters of the Salish Sea (i.e. greater Puget Sound) to estimate the relative abundance, species composition, and biological characteristics of benthic fauna. Occasional surveys have also occurred in Canadian waters (1997, 2000, 2001, and 2003). While all surveys have incorporated the same depth-stratification, and generally occurred in spring, the overall survey design has varied, with surveys conducted at irregular intervals and at different spatial scales. Surveys in 1987, 1989, and 1991 sampled all of Puget Sound except the San Juan Islands sub-basin. From 1994-1997 and 2000-2007, surveys were conducted annually and focused on 1-4 of nine demarcated Puget Sound sub-basins. Starting in 2008, the WDFW implemented a synoptic survey design, with tows occurring annually at fixed index sites throughout eight of the Puget Sound sub-basins in order to get annual, Sound-wide spatial coverage. From 2008 to 2013, two trawl tows were conducted at each station; however, based on the similarity of catches in these paired tows at most stations, and in the interest of reducing survey costs and minimising bottomfish mortality associated with the survey, the protocol was altered in 2014. After the first tow is completed, the processed catch is compared to the average catch at that station since 2008, and a second tow is conducted only if the catch was substantially different than expected. This presentation will discuss the qualitative and quantitative costs and/or benefits of the major survey design changes.

Bryan, Meaghan D, and James T. Thorson  Alaska Fisheries Science Center (NMFS)

The impact of survey frequency and intensity on detecting environmental anomalies and shifts in abundance

Meaghan D. Bryan and James T. Thorson

Climate-driven shifts in abundance are becoming increasingly familiar in many systems. At the forefront of our ability to detect these shifts are fisheries-independent surveys that provide vital information about the scale of abundance and trends of exploited stocks. The ability to accurately detect changes in abundance is greatly influenced by the survey’s spatial extent, which is commonly constrained by management jurisdiction and logistical and budgetary considerations. In the US North Pacific, we have a consistent, annual survey in the eastern Bering Sea, a focal area of fishing and fisheries management, and a less consistent survey in the northern Bering Sea. Information from both is vital to understanding the changing dynamics of the system. It is of
great importance to determine how these separate pieces of information can be combined to provide a comprehensive view of stock-specific changes in the Bering Sea and for use in stock assessments. Another overarching concern of survey programs is the changing budgetary environment within which they operate. Budgets directly control survey sampling frequency and intensity and understanding their impact on our ability to estimate environmental variation and abundance is crucial. We completed simulation experiments to address these issues and to determine how well a spatio-temporal model estimated 1) the local impact of a regional environmental index given annual, full sampling; annual, reduced sampling; and biennial, full sampling over a range of correlated spatial scales and 2) the proportion of biomass in an inconsistently sampled area given the range of sampling intensity and frequency from simulation 1. We show that annual surveys, even with reduced sampling intensity, better estimate the local environmental effect on abundance. More importantly this is the first simulation exercise to demonstrate that a spatio-temporal model can accurately estimate the proportion of abundance in an area not consistently sampled with relatively low error from historical data. Indices of abundance for Bering Sea walleye pollock and Pacific cod are currently being developed using this same spatio-temporal model; hence, this is a meaningful and practical result for North Pacific fisheries management.

Conner, Jason  Alaska Fisheries Science Center (NMFS)

Impact of reducing sample density on the accuracy and precision of design-based estimators of an abundance index for a bottom-trawl survey in the eastern Bering Sea

Jason Conner, Kotaro Ono, Stan Kotwicki, Lewis Barnett

Fisheries-independent resource surveys are an indispensable component of fisheries stock management programs, providing estimates of abundance indices (AI) that are dependent only on the assumptions of sampling design and the accuracy and precision of the estimator. NOAA has conducted the Eastern Bering Sea Continental Shelf Bottom-trawl Survey (BTS) of Groundfish and Invertebrate Resources annually since 1979, and has conducted sampling with a standardized bottom-trawl since 1982. BTS trawl samples are collected systematically from a 20 nmi grid, covering 492,898 km2. As we consider potential changes to BTS effort, it is prudent to identify the optimal design-based estimator of AI under varying sample densities for the present geographical extent of the BTS. We simulate the spatial distributions of 4 species (Gadus chalcogrammus, Gadus macrocephalus, Atheresthes stomias, Limanda aspera) from 1982-2016 with known mean CPUE. We then simulate the BTS with 3 sampling strategies (simple random, stratified random, and systematic) over 4 sampling densities, and evaluate the performance of estimates of mean CPUE for each design. Additionally, it is known that there is no unbiased estimator for the variance of an estimate from systematic sampling, so we utilize the simulation framework to provide a comparison of a post-stratified local variance estimator with the naïve variance estimator of mean CPUE. We demonstrate that 1) the bias of systematic sampling for mean CPUE depends on the generalized distribution of a given species, 2) given the estimation error for some species, random sampling with smaller sample sizes may be desirable, and 3) the precision of the variance estimate currently used for mean CPUE may present challenges in interpreting the point estimate of a BTS.
Cunningham, Curry  University of Alaska Fairbanks

Implications of changes in bottom-trawl survey effort on the quality of stock assessment results

Curry Cunningham

Biomass indices and age composition data from NOAA bottom-trawl surveys are a key source of information for assessment of Alaska’s commercially valuable fish stocks. Unavoidable reductions in survey effort will require examination of the trade-offs among alternatives for reducing survey effort across space and time. However, only quantifying the changes in design-based or model-based uncertainty in derived biomass indices is unlikely to be a sufficient basis for evaluating these trade-offs. More importantly, it is necessary to examine how these trade-offs translate to differences in stock assessment results. To quantify the implications of changes in survey effort, we explore the impact of removing information from the NOAA Alaska Fisheries Science Center Eastern Bering Sea Shelf Survey on the assessed status of the Walleye pollock (Gadus chalcogrammus) and Pacific cod (Gadus macrocephalus) stocks. By removing survey catch per unit effort and age composition observations from specific survey strata-year combinations and using these data to inform a range of simplified stock assessment models, we identify the sensitivity of assessment model estimates and uncertainty to reductions in survey effort. First, we provide a general description of the survey effort reduction alternatives explored and the range of assessment models used to estimate trends in stock status. Next, we provide a comparison of design-based and Vector Autoregressive Spatio-Temporal (VAST) model-based biomass indices produced under a given level of survey effort. Finally, we compare assessed trends and uncertainty in stock status and management reference points when biomass indices and age composition data available under these survey effort alternatives are provided to the range of simplified assessment models.

Hamel, Owen  Northwest Fisheries Science Center (NMFS)

The effect of trawl survey frequency and intensity on U.S. West Coast groundfish stock assessments


Fisheries management systems rely on stock assessments, which, ideally, are informed by well-designed and comprehensive surveys, providing the data necessary to estimate scale and trends in fish populations. Given limited budgets and the financial demands of conducting surveys and the concomitant laboratory and analytical requirements, it becomes necessary to consider trade-offs in survey design, as well as evaluate the impact of alternative approaches to reducing survey effort and cost. We conducted a retrospective analysis of the impact of 50% sampling reduction in intensity or frequency of the U.S. West Coast Groundfish Bottom-trawl survey. The influence of the survey reductions on assessment outputs and catch limits for 11 species were evaluated.
and found to depend upon species life history, frequency of occurrence in the survey, and the data-richness of each assessment. All approaches to reducing survey sampling led to increased uncertainty in stock assessment results, especially for more rarely encountered species, for species that have not been heavily exploited, and for a set of data-moderate assessments, which rely more heavily on survey indices.

Hollowed, Anne  Alaska Fisheries Science Center (NMFS)

SSC perspective on trade-offs among trawl survey schemes in federal waters off Alaska under varying funding scenarios

Anne Hollowed (SSC co-chair)1, Gordon Kruse (SSC co-chair)2, George Hunt3, Alison Whitman4, Dayv Lowry5, and Dana Hanselman6

1Alaska Fisheries Science Center, Seattle, WA
2Emeritis, College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, Juneau, AK
3 University of Washington, Seattle, WA
4 Oregon Department of Fish and Wildlife, Newport OR
5 Washington Department of Fish and Wildlife, Olympia, WA
6 Alaska Fisheries Science Center, Juneau, AK

In the United States, Federal Fishery Management Councils are required to review and approve research priorities every 5 years. The North Pacific Fishery Management Council (NPFMC) receives advice and recommendations on research priorities from the Scientific and Statistical Committee (SSC). These priorities are maintained in a public database https://research.psmfc.org/ to provide public access to the ongoing and evolving research needs relevant to federally managed fisheries for fish and shellfish in waters off the coast of Alaska. Given the plethora of potential research needs for such a large region, the SSC recommended partitioning research needs into four categories: Critical ongoing monitoring, Urgent, Important, and Strategic. In 2019 the Council adopted a 3 year cycle for review of research priorities, and recommended that development of a “top ten” list to highlight issues of particular relevance to Council needs.

As key input to the research prioritization process, the SSC annually receives updates from the National Marine Fisheries Service, Alaska Fisheries Science Center’s Science Director on agency research priorities including survey planning. This presentation typically includes consideration of the balance of ecosystem and stock assessment surveys, current rationale for priority surveys, and emerging science questions. In June 2018, AFSC Leadership requested a subcommittee of the SSC be established to meet with AFSC staff to develop recommendations for survey prioritization in FY2019 and beyond. In September 2018 SSC subcommittee met with AFSC survey team to discuss 2019 survey prioritization under reduced funding scenarios. In October 2018, SSC subcommittee delivered a white paper to the NPFMC documenting their review of AFSC’s research priorities. This review exercise serves as a case study for how decisions regarding unexpected reductions in surveys were reconciled for federally managed fisheries off the coast of Alaska.
Key findings drawn from this case study serve as a potential approach for science informed decision-making for responses to unexpected survey reductions that have implications for federally managed, public resource. The process started with a full day workshop with a subcommittee of survey experts and stock assessment scientists to allow the SSC to review the best available scientific information regarding the pros and cons of different survey prioritization options. Clear terms of reference were provided to the SSC in advance to focus the group on key issues to focus the discussion.

The AFSC typically receives funding for 5 boats for its trawl surveys. Allocation of this funding includes an annual trawl survey of the southeastern Bering Sea Shelf; biennial surveys of the northern Bering Sea, Aleutian Islands and EBS slope in even years; and biennial surveys of the Gulf of Alaska in odd years. In this case study, scientists were asked to consider how to allocate ship time under a reduction from a 5 boat to a 4 boat survey. The timing of this request coincided with a period when highly valuable fish stocks were experiencing unexpected range shifts in response to a marine heat wave (Stabeno and Bell, 2019; Stevenson and Lauth, 2019; Thorson, 2019). The rapid changes in environmental conditions and fish abundance and distribution in response to Marine Heat Waves accentuated the need to continue, or increase, trawl survey frequency and sampling density in the northern Bering Sea. These abrupt changes in the distribution and productivity of key marine species raised several pressing management questions including: emerging transboundary management of stocks, concerns regarding stock structure (Spies et al.), non-stationary patterns of productivity (Duffy-Anderson et al., 2017), and shifting species interactions (encounter rates of non-target species in catch, and prey availability for top trophic level consumers). Thus, decisions regarding how to address reductions in survey effort was complex and multifaceted.

The SSC recognized that the trawl survey enterprise, creates and maintains indispensable data that substantially contribute to scientific understanding and management of fish and crab populations, fisheries, and communities dependent upon those fisheries. Discontinuation or diminishment of the research that provides these datasets would leave a significant gap in the science needed to support sustainable and successful fisheries management in the North Pacific. The SSC determined a priority ranking for AFSC’s surveys based on considerations of the implications for stocks in number of very valuable fisheries monitored by the survey, the areal coverage of the survey, the dependence of coastal communities on the fisheries, importance of monitoring stock structure, and data to assess and manage populations, and current stock status.

The decision process was informed by the quantitative investigations into the implications of survey reduction (many of which were presented at this workshop). The decision process also benefitted from the ongoing research collaborations between stock assessment modelers and survey experts. It was noted that there are inherent trade-offs in balancing field and laboratory activities designed to understand processes underlying ecosystem change and core monitoring and assessment for tactical fisheries management. It is anticipated that these trade-offs will be considered in planned strategic modelling exercises as part of the Fisheries Ecosystem Plan (FEP) Climate Module.
References:

Jorgensen, Elaina  Alaska Fisheries Science Center (NMFS)

Systematic reduction in survey effort and the effect on variance of fish abundance

Elaina Jorgensen

The Eastern Bering Sea Survey (EBSS) has been conducted annually for over 30 years. Sampling is based on a standard 20 x 20 nm grid. From this real dataset, a mean and variance of abundance by species and grid cell was calculated. A simulated dataset, based on the spatial extent of the EBSS but with a 2 x 2 km grid, was used to calculate an annual mean and variance by species and EBSS grid cell. Grid cells with a simulated mean and variance less than the lowest 20% of the real data were removed incrementally (by 1% (approximately 4 stations)), and a new mean and variance was calculated for each of four commercially important species (pollock, cod, yfs, and turbot).

Kotwicki, Stan  Alaska Fisheries Science Center

The effect of variable sampling efficiency on the reliability of observation error as a measure of uncertainty in abundance indices from scientific surveys

Kotwicki, Stan and Kotaro Ono

The major goal of fisheries surveys is to obtain abundance indices (AIs) of fish populations. The AIs from surveys are important because they provide necessary information for research on fish populations and for fisheries management. While AIs provide information on population abundance and trends over time, AI variance (AIV) provides information on reliability or quality of the AI. AIV is an important output from surveys and is commonly used in formal assessment of survey quality, in survey comparison studies, and in stock assessments. However, uncertainty in the AIV estimates is poorly understood and studies on the precision and bias in survey AIV
estimates are lacking. Typically, they are “design-based” and derived from sampling theory under some aspect of randomized samples. Inference on population density in these cases can be confounded by unaccounted process errors such as those due to variable sampling efficiency (q). We show that the bias and precision of AIV depends on the mean q and variance in q. We conclude that to fully evaluate the reliability of AI, both observation error and variability in q must be accounted for when estimating AIV. A decrease in mean q and increase in the variance in q results in increased bias and decreased precision in survey AIV estimates. These effects are likely small in surveys with mean q ≥ 1. However, for surveys where q ≤ 0.5 these effects can be large. Regardless of the survey type, AIV estimates can be improved with knowledge of q and variance in q.

Kupschus, Sven  Centre for Environment, Fisheries, and Aquaculture Science

An empirical approach to predicting the effects of fisheries independent survey effort reductions on biases and precision characteristics of stock dynamic metrics used in the management of fisheries. An example for three gadoid species in the North Sea

Sven Kupschus

Fisheries surveys are increasingly used to assess multiple species with different distributions and stock dynamics assessed by an increasing number of assessment methodologies. Given the increased need for fisheries independent information under budgetary constraints a better quantitative understanding of the effects of survey effort is required to make decisions in the long term and at sea decision-making on effort deployment.

This study applied an empirical approach to predicting and evaluating the impact of potential future changes to survey design and effort changes on management metrics. The survey data index calculation and stock assessment (implemented with SAM) processes for three North Sea gadoid species were automated in a MSE framework implemented as a closed loop process to evaluate two different methods of reducing survey effort. The first option randomly selected 50% of the stations based on a post stratification design. The second option used all stations, but randomly selected individual fish from the catch using a binomial error distribution with mean probability = 0.5. Both the ICES index calculation method and the GAM index method currently used for gadoid stocks were implemented for each of the three species and the simulated assessment results were compared.

1. Generally, the loss of precision of assessment metrics (SSB, F and Recruitment) was minimal given the 50% reduction in abundance information
2. The delta-gamma based GAM indices (model-based) tended to produce increasingly biased assessments with decreasing sampling effort unlike ICES design-based indices which seemed to remain largely unbiased at least at 50% sample reductions.
3. Empirical standard error estimates for two different survey effort reduction strategies for the two strategies did not suggest a substantial degree of autocorrelation in the length and age information within stratum despite the clustered biological sampling protocol.
4. Comparison with current assessments for the various combination of index methods and effort reduction strategy indicated a lack of a consistent criteria for bias and precision-loss suggesting a case specific examination is necessary when considering changes to multi-purpose surveys.

The results highlight the importance of both quantitative methods for examining changes to surveys and a clear understanding of the prioritization of survey objectives when making decisions on survey effort reductions.

**Laman, Ned  Alaska Fisheries Science Center (NMFS)**

**Effects of sampling density changes on biomass estimates from stratified random bottom-trawl surveys in the Gulf of Alaska**

*Ned Laman*

Sampling density has been reduced on 4 of 10 biennial summer bottom-trawl surveys conducted in the Gulf of Alaska (GOA) between 1996 and 2015. The objectives of our project were to assess how biomass and uncertainty estimates change with changing sampling density and to compare the performance of different estimation methods under those conditions. Using a known population, we simulated standard GOA stratified random surveys for the years 1996-2015 at four different sampling densities and examined estimates of biomass and uncertainty for four federally managed species (arrowtooth flounder, Pacific cod, Pacific Ocean perch, and sablefish) in this large marine ecosystem. Reducing sampling density produced less accurate and more uncertain estimates of biomass. Performance of estimation method and selection of the best model varied among species.

**Lambert, Gwladys  Centre for Environment, Fisheries, and Aquaculture Science**

**Reducing effort in a stratified fixed station survey – impact on survey indices and assessment of a data rich and data-limited stock**

*Gwladys Lambert*

The Irish Sea and Bristol Channel Beam Trawl survey, BTS-UK(E&W)-Q3, is an annual survey that samples ICES Divisions 7.a and 7.f-g. One of the main target species is plaice. This survey provides the only age-based index for the plaice 7.a stock assessment, which is a category 1 (data rich) stock by ICES definition and uses an age-based analytical approach (SAM model). Two other survey indices contribute to this assessment. The BTS-UK(E&W)-Q3 survey provides a biomass index for the plaice 7.f-g assessment. The survey predominantly covers the 7.f part of the stock, while another survey provides an index for 7.g. This is a category 3 (data-limited) stock and the assessment used is a stochastic surplus production model in continuous time (SPiCT).
BTS-UK(E&W)-Q3 is a stratified fixed station survey where 108 fixed stations are spread across depth bands and within survey sectors (geographical regions). We simulated a reduction in survey effort by either systematically or randomly removing 10%, 20% and 50% of the stations per strata, thereby keeping all strata sampled, but reducing the overall effort. We re-calculated the survey indices the same way that they are currently calculated for the stock assessments and then re-ran the assessments to assess the impact of reduced survey effort on the stock perception and on the management indicators for these two different stocks; i.e. comparing the impact of survey reduction on a category 1 versus a category 3 stock assessment for a single species.

Marshall, Kristin  Northwest Fisheries Science Center

Understanding trade-offs with survey frequency using Management Strategy Evaluation (MSE): a Pacific Hake case study

Marshall, Kristin, Nis Jacobsen, Ian Taylor, and Aaron Berger

The consequences of changes in survey effort have the potential to propagate through fisheries management systems. Potential effects include changes to indices of abundance, stock assessments, and perceived stock status, which could influence management decisions like setting catch limits. Assessment and management of Pacific Hake, the largest groundfish fishery on the U.S. West Coast, depends on an acoustic survey currently conducted biannually. The frequency of this survey has been the subject of debate for nearly a decade. Here, we use a Management
Strategy Evaluation (MSE) framework to better understand consequences and trade-offs associated with operating the survey every 1, 2, or 3 years. We discuss these trade-offs in the context of previous analyses, and planned future work to investigate other survey methods, including a recruitment index and the use of saildrones. We also highlight pros and cons of a MSE approach to better understand the consequences of changes in survey effort.

Martin, Michael  Alaska Fisheries Science Center (NMFS)
Keynote Address--US Perspective

Munro, P  Alaska Fisheries Science Center (NMFS)

Efficiency of four estimators of change in mean trawl survey catch per unit effort (CPUE), evaluated using empirical Mean Squared Errors (MSE)

Trawl surveys produce standardized measures catch rates, the catch per unit effort (CPUE), that may be taken as proxies for fish density. Change in mean CPUE from survey to survey can then be interpreted as a measure of change in abundance. Using simulated fish populations, four estimators of change in mean CPUE were evaluated using empirical calculations of the Mean Squared Error of the estimate (MSE). All four estimators produced estimates of change in mean CPUE that were slightly biased but with no variation in bias as a function of the magnitude of the change in abundance. Differences in empirical MSE appeared to be due to the variance of the estimate, with the relative efficiency of the estimators remaining constant across sample sizes.

Methot, Richard  Office of Science and Technology (NMFS)

Challenges for Next Generation Surveys

Richard Methot and Cisco Werner

Our marine resource survey enterprise is challenged to support expanding needs. Climate changes are moving fish into new locations requiring us to expand survey coverage and to conduct more ecosystem data collection to understand causes of these shifts. Ships, staff, and budget to support these needs have not expanded to meet the requirements. Our response is to work smarter as guided by recent planning documents, including the NMFS Updated Stock Assessment Improvement Plan (2018). The SAIP guides us to make improvements on three time and technology scales. First is to optimize the allocation of existing survey effort. We can use spatial models to optimize the information obtained from existing survey observations. With simulation tools such as Management Strategy Evaluation we can evaluate the impact on fishery management of changes in survey frequency and coverage. The work being conducted by the WkUSER group exemplifies such work. Second, we can use technological improvements to magnify the data we can collect from each day at sea. Advancements such as automated image analysis allow
rapid translation of raw image data into fish counts and sizes. Robotic platforms such as sail-drones can be force multipliers to expand the coverage of some pelagic fish surveys. Other technological improvements and dedicated field studies are allowing us to achieve absolute, not just relative, calibration of some surveys to provide more direct information on fish and shellfish abundance. Third, we continue to track radical new technologies and conduct applied research to fine-tune them to our fish assessment needs. Currently we are working on eDNA as a tool to detect fish remotely and on genetic mark-recapture approaches that hold promise as a new tool to measure the abundance of some fish stocks.

Ono, Kotaro  Norwegian Institute of Marine Science

A spatio-temporal operating model for simulation testing Alaskan bottom-trawl survey effort and design

Ono, Kotaro, Lewis Barnett, and Stan Kotwicki

When faced with reduced resources to conduct fishery-independent surveys, monitoring agencies require objective methods for determining how best to adapt or redesign surveys. We present a statistical modelling framework that serves as the basis for several ongoing studies estimating the likely effects of reduced Alaskan bottom-trawl survey effort on groundfish abundance indices and how these consequences may be mitigated using different estimators, survey designs and effort allocations. The objective was to generate predictions of groundfish density at all locations within a survey domain for each period of interest, to represent true species distributions from which simulated surveys can be conducted. To generate such predictions from survey data, where many 0 values are often present, we developed a dynamic species distribution model in the form of a hurdle model, where two generalized linear mixed models (GLMMs) are used to separately estimate the probability of occurrence and the magnitude of density when present. The model accounts for spatial and temporal dependence by estimating a spatial random field, which evolves through time following a first-order autoregressive process. To produce predictions that replicate the extreme patchiness often observed in survey data, we include an option to predict by sampling from the posterior predictive distribution of parameter estimates rather than predicting from the mean values of each parameter. We fitted the model to all years of bottom-trawl surveys available through the NOAA Fisheries Alaska Fisheries Science Center, creating density predictions for each of a suite of commercially and ecologically important species in the eastern Bering Sea and Gulf of Alaska, selected to represent a range of life histories and distributional properties. We also present guidance on diagnostics to assess model fit and representation of realistic survey catch distributions.
Richar, Jon  Alaska Fisheries Science Center (NMFS)

Considering changes in sampling density and survey frequency, and their effects on eastern Bering Sea crab population time-series

Jon Richard

Design of stock assessment surveys is a critical aspect in maintaining data integrity and confidence in management decisions, and changes in the design(s) of established surveys can have dramatic effects on population estimates and time-series. In addition to temporally variable trends, crab are contagiously distributed, thus survey design can have dramatic effects on estimates of population size structure. We examined two alternative designs for the eastern Bering Sea trawl survey; 1.) Reconstructing the 44 year eastern Bering Sea trawl survey to calculate area swept estimates for even and odd years, and 2.) Reconsidering high density vs. low density sampling efforts within the regions of the Pribilof and St. Matthew Islands. These alternative survey designs were evaluated based on how they affected both absolute population estimates, and statistical properties of estimates, for regional stocks of southern Tanner crab, Chionoecetes bairdi, opilio crab, Chionoecetes opilio, blue king crab, Paralithodes platypus, and red king crab, Paralithodes camtschaticus. Autocorrelation within population time-series drove statistically significant correlations between even and odd year estimates used to evaluate the alternating year approach for all stocks, though significant variability was still observed in the relationships. Due to patchiness in spatial distributions, eliminating corner stations increased CVs for Pribilof stocks of red and blue king crab, and the St. Matthew stock of blue king crab, in all maturity and size categories evaluated. Abundance and biomass estimates were also generally reduced for all stocks, with magnitudes of up to -100% observed. Comparatively reduced effects were observed in opilio and bairdi stocks. These results suggest that serious consideration should be given to the underlying temporal and spatial variance of these crab stocks before changes to the survey are considered.

Rideout, Rick Department of Fisheries and Oceans, Canada

An Overview of Fisheries and Oceans Canada’s Multispecies Bottom-trawl Surveys in the Newfoundland and Labrador Region: Survey Coverage Issues and Implications for the Provision of Science Advice

Rick Rideout

The Newfoundland and Labrador Region of Fisheries and Oceans Canada conducts two annual research vessel bottom-trawl surveys. The current survey designs cover an expansive spatial area, spanning six NAFO Divisions (2HJ3KLNO) and 515,000 km2 in autumn and four NAFO Divisions (3LNOP) and 324,000 km2 in spring. Surveys are currently conducted using two ageing research vessels (>30 years of operation) that are also shared with other Fisheries and Oceans Canada regions. Extensive mechanical delays in recent years have resulted in reduced survey coverage, interchange of research vessels outside of their normal area coverage pattern, and have
extended the time required to complete surveys of the individual divisions. There is currently no formalized mechanism in place to deal with unavoidable reductions in survey effort but typically attempts are made to reallocate remaining survey time to cover high priority areas (i.e. areas considered important for high priority stock assessments) and/or to cover areas that will allow the largest possible number of stock assessments to be completed. Survey time constraints have often resulted in some areas (e.g. deep strata in Divs. 2H and 3L in autumn survey) being eliminated from the survey area in order to prioritize completion of other areas. Also, areas typically covered toward the end of the surveys (e.g. Div. 3L in spring survey) are often subject to coverage issues. Deficiencies in these surveys have had direct impacts on the advice produced for many groundfish and invertebrate assessments, which may be particularly concerning for those stocks on the Grand Bank where ecosystem changes appear to be occurring and many stocks are in decline.

Rogers, Lauren  Alaska Fisheries Science Center (NMFS)

Evaluation of a survey with an adaptive sampling domain to capture climate-driven shifts in larval fish distributions

Lauren Rogers and Kathy Mier, EcoFOCI, AFSC

The EcoFOCI program conducts biennial spring surveys in the Bering Sea to study the early life stages of fish, in particular Walleye Pollock, in order to better understand the ecosystem processes underlying fisheries recruitment and productivity. Studies have shown that the spatial distribution of Walleye Pollock larvae differs in cold and warm years, and that the existing survey design was failing to capture the extent of their distribution. In 2015-16 we reassessed the survey design with three main objectives: 1) to improve its coverage of the major known spawning areas of Walleye Pollock on the southeastern Bering Sea shelf, 2) to ensure it captures the variable spatial extent of larvae, and 3) to ensure it can be used to create an index of larval abundance for ongoing studies of early life stage survival and recruitment. The new design has a coarser sampling grid and two types of stations: core sampling stations that are sampled every survey, and adaptive sampling stations on the edges of the core that can be added dynamically during a survey if the edge of the distribution of larvae has not yet been reached. The new survey design was implemented in May 2016 and 2018. We are now conducting a simulation study to assess the performance of the new survey design relative to a fixed grid station design when the underlying patches of larvae are shifting in space. We are also evaluating both model- and design-based estimators of abundance. As more species shift their distributions under climate change, it is becoming necessary to consider and carefully evaluate new survey designs, especially when resources for ship time are limited.
Rooper, Chris Department of Fisheries and Oceans, Canada

Accounting for habitat variables to improve abundance indices in Alaska trawl surveys with an emphasis on results from averaging multiple modelling methodologies

Chris Rooper

Stock assessments for Alaska fish species rely on time-series of abundance data from fishery-independent surveys, primarily bottom-trawl surveys conducted on the continental shelf and upper slope of the Gulf of Alaska (GOA), eastern Bering Sea and Aleutian Islands (AI). For some species, such as Pacific ocean perch (POP), yelloweye rockfish, and northern rockfish in the GOA and Atka mackerel, POP, and northern rockfish in the AI, the variability in abundance estimates can be quite high (coefficients of variation > 0.50) and in some years the change in estimated survey abundance can be greater than appears biologically feasible. This high variability is a critical problem with existing bottom-trawl time-series that increases uncertainty in stock assessments. In part, this variation can be attributed to the random allocation of trawl survey stations within strata without respect to habitat type. For example, the allocation of stations in the GOA at a relatively small area (Snakehead Bank) of known high abundance of rockfish has ranged from 5 to 19 in the random allocation for surveys from 1996-2013. The coefficient of variation of the biomass of northern rockfish for the larger strata (Kodiak Outer Shelf stratum and Chirikof Outer Shelf stratum) that contains the Snakehead Bank decreases linearly with an increase in the number of tows at the Snakehead Bank. The goal of this project is to reduce uncertainty in the stock assessments for selected rockfishes and Atka mackerel in the GOA and AI by developing and testing models that explicitly account for habitat-related trends in abundance. Five approaches for producing biomass estimates using habitat modelling were explored with respect to each time-series biomass estimates, variances, and deviance explained. Finally, we created time-series that combined methodologies into a single estimate through model-averaging or ensembling.

Spencer, Paul Alaska Fisheries Science Center (NMFS)

Variance propagation from fishery-independent surveys to stock assessment outputs

Paul Spencer

Fisheries stock assessments typically rely on abundance indices (AIs) from resource surveys to scale the estimated population size and inform population trends, with the estimated abundance index variances (AIVs) quantifying the uncertainty in the AIs. Stock assessment models commonly weight individual AI estimates in inverse proportion to their estimated uncertainty. Similar to AIs, the estimates of AIV are a function of sampling data and are subject to estimation error and/or bias. However, stock assessments do not often consider the uncertainty in the AIV estimates, leading to the potential errors in assessment outcomes due to erroneous (i.e. uncertain
or biased) AIV estimates. In this study, we simulated spatial and temporal patterns in population abundance for three species (walleye pollock, Pacific cod, Pacific ocean perch) based on the Alaska Fisheries Science Center Gulf of Alaska bottom-trawl survey data, taking these to represent the true abundance. We estimated replicate AIs and AIVs from these data by simulating surveys using simple random sampling at a range of sampling intensities. We then used the simulated survey estimates in a non-age structured population assessment model to evaluate how variance in survey biomass estimates influenced stock assessment results (estimated scale and trends in abundance over time). Sampling intensity affected both the survey biomass estimates and its variances, resulting in time-series of estimated survey biomass that can differ substantially from the true biomass. The ability to detect trends is a function of the sampling variance of the biomass estimate relative to the interannual variation of true biomass, and the strength of the underlying trend. This modelling framework can provide guidance for the required sampling intensity necessary to achieve a desired level of precision and accuracy of assessment biomass estimates.

Thorson, Jim Alaska Fisheries Science Center (NMFS)

Measuring the impact of increased ageing effort: theory and case-study demonstration

Thorson, James T., Meaghan Bryan, Pete Hulson, Haikun Xu, André E. Punt

Ocean management involves a sequence of interconnected research activities, typically involving planned monitoring, development and application of biological models, and resulting stock-status estimates that are used to inform policy changes. However, few science organizations have a recurring, quantitative process to optimize the effort allocated to research efforts across different assessments, perhaps due to complexities when comparing the impact of changing survey information within different assessment models. For example, effort spent processing samples to measure fish age can be easily reallocated among species, but methods to conduct “power analysis” for age-composition data have been limited to large simulation studies that are expensive to update, difficult to condition upon specifics for different assessment models, and where responsibility for applying these methods are impractical to divide among different research teams. To address this difficulty, we decompose a power analysis for age-reading effort into three independent steps: (Step 1) the sensitivity of input-sample size to the number of raw age measurements; (Step 2) the sensitivity of effective sample size to changing input-sample size; (Step 3) the sensitivity of stock-status variance to changing effective sample size; and (Step 4) the sensitivity of management performance to stock-status variance. We propose a bootstrap estimator to conduct Step 1, and derive a novel analytic estimator for Step 2 when reweighting age-composition data using a Dirichlet-multinomial likelihood. We then provide two simulation studies to evaluate our two proposed estimators, and show that the bootstrap estimator for Step 1 underestimates the likely benefit of increased age reads while the analytic estimator for Step 2 is unbiased given a plausible mechanism for overdispersion. We conclude by recommending additional research regarding Step 3 (the sensitivity of stock-status estimates to changing effective sample size) and outline how these results could be used within a formal process to update survey efforts for stock assessment.
Von Szalay, Paul Alaska Fisheries Science Center (NMFS)

A Comparison of Bottom-trawl Sampling Strategies in the Gulf of Alaska: Design vs. Model-Based Approaches

Paul Von Szalay

Using 2x2 km resolution fish density maps derived from 1996-2015 survey data and superimposed on the Gulf of Alaska survey area, we simulated full-scale bottom-trawl surveys of 820 survey stations. Comparisons of mean catch-per-unit effort and variance were made between a design-based and a model-based survey design. Results from both approaches were compared with the true mean and variance derived directly from the density maps. Four major and representative species were considered: Pacific cod (Gadus macrocephalus), Pacific Ocean Perch (Sebastes alutus), arrowtooth flounder (Atheresthes stomia), and sablefish (Anoplopoma fimbria). A simple random sampling design represented the design-based approach and a spatio-temporal model (VAST), which was applied to simulated surveys derived using three different sampling strategies (uniform and simple random distributions and an information criterion-based distribution), represented the model-based approach. We expect the model-based approach with the information criterion sampling strategy to perform the best because it incorporates both temporal and historical variability information into the sampling strategy.

Walker, Nicola CEFAS

Is the North Sea IBTS oversampled – computer-based study of the effects of reduced sampling on stock assessments?

Nicola Walker

The North Sea International Bottom Trawl Survey (IBTS) occurs twice each year with participation by eight countries, typically with two different countries sampling each ICES statistical rectangle. Here we investigate the effects of removing each of the Q1 (Quarter 1) and Q3 (Quarter 3) survey indices from assessments of North Sea stocks to evaluate whether both time-series are needed. We then investigate whether excluding data collected by individual countries reduces the precision of indices for North Sea cod and the effect this has on the assessment of the stock.

Martin, Michael Alaska Fisheries Science Center (NMFS)

Keynote Address--US Perspective
Williams, Kresimir  Alaska Fisheries Science Center (NMFS)

Cameras vs Catch: potential effects of implementing open codend tows for acoustic midwater fish surveys

Kresimir Williams

The Midwater Assessment and Conservation Engineering (MACE) program conducts acoustic-trawl (AT) surveys of pollock in the north Pacific using midwater trawls to characterize acoustically sampled fish aggregations. The MACE program developed a trawl-mounted stereo-camera device (CamTrawl) to estimate fish lengths as they pass through the trawl. This allows for non-retention (open-codend) trawl hauls to be conducted to provide image-based fish size and species composition estimates. However, open-codend tows do not provide other biological data such as fish weight, sex, maturity, and age structure samples. The effect of reducing the number of these sample data was explored using simulations from using historic MACE survey data from the eastern Bering Sea and Shelikof Strait (Gulf of Alaska). Estimates of biomass at length, sex ratio and length at 50% maturity were recomputed under two scenarios: 1) reducing the number of tows from which biological data (excluding length) were taken, and 2) proportionally reducing the number of samples (excluding length) taken from each tow. The effects of reducing the biological sampling by 50% in total number of tows, or for a proportional reduction of sampling across all tows were similar for the EBS and showed small effects on the reference biomass with 90% of simulations showing < 2% change across all surveys. Effects on biomass for Shelikof Strait were slightly higher, with 90% of simulations showing < 3.5% change across all surveys. The effects on length at maturity estimates in Shelikof Strait exhibited the largest changes when haul effort was reduced by 50%, with 90% of simulations showing < 5% change for all survey except one survey, where 90% of simulations were within 12%. Potential application of these results in the design of future surveys, as well as implications for the additional uncertainty from camera-derived length compositions will be discussed.

Yeung, Cynthia  Alaska Fisheries Science Center (NMFS)

Survey Effort Reduction Impacts on the Assessment of the Thermal State of the Bering Sea Ecosystem

Cynthia Yeung

Bottom temperature is an abiotic variable routinely measured by the Eastern Bering Sea, Aleutian Islands, and Gulf of Alaska bottom-trawl surveys conducted by the Alaska Fisheries Science Center. It was measured using expendable bathythermographs in the early years, and using a depth and temperature data logger mounted on the headrope of the trawl in the recent years. It is of paramount importance as an indicator of the thermal states of these marine ecosystems. The continuous warming of the high-latitude oceans in the recent years is associated with significant changes in the distribution and abundance of marine species, including groundfish species of
high economic value. Outlook on the thermal environment is incorporated into stock assessments to inform management decisions. In the Eastern Bering Sea, key derivatives of bottom temperature include the area and the geographic extent of the "cold pool" - a pool of ≤2°C bottom-water resulting from spring ice melt that is a characteristic feature detected in summer in the middle shelf (50- to 100-m depth). The cold pool presents a thermal barrier to cold-intolerant groundfish. Under the current warm ecosystem state of the Eastern Bering Sea, sea ice formation has been curtailed and the cold pool has shrunken and retreated northwards. Correspondingly, there have been major changes in the groundfish community. A reduction of bottom-trawl survey effort temporally and/or spatially is likely to affect the accuracy of assessing the present and projecting the near-future thermal states, and may subsequently impact the stock assessment and management of groundfish.
5 Examples of Surveys Contributed at the Workshop

5.1 Washington Department of Fish and Wildlife Southern Salish Sea Annual Bottom Trawl Survey

Jen Blaine, Washington Department of Fish and Wildlife
See workshop abstract.
5.2 West Coast Bottom Trawl Survey

Amiee Keller, Northwest Fisheries Science Center

Title: The Northwest Fisheries Science Center’s U.S. West Coast Groundfish Bottom Trawl Survey

Abstract: The Northwest Fisheries Science Center (NWFSC) conducts an annual bottom-trawl survey along the upper continental slope and shelf of the U.S. west coast. Our mission is to provide scientific information for management of west coast groundfish and their ecosystem. Since 1998, the survey has occurred in conjunction with the commercial fishing industry aboard chartered west coast bottom-trawlers. The survey targets groundfish resources in trawlable habitats at depths of 55–1280 m from U.S.-Canada to US-Mexico and is the primary source of fisheries-independent data used in the majority of stock assessments for commercially important west coast species. We focus on abundance, distribution and biological characteristics of the 90+ demersal fish in our management plan. We additionally collect environmental (salinity, temperature, dissolved oxygen, chlorophyll, turbidity and bottom depth) and habitat data. The NWFSC assumed responsibility for the survey in 1998, extending two pre-existing west coast surveys conducted by the AFSC since 1977: a triennial shelf survey and an annual slope survey.
5.3 Bering Sea Bottom Trawl Surveys

Elaina Jorgensen, Nancy Roberson, Alaska Fisheries Science Center

**Bering Sea slope**
- Larger stratified random
- Sampled in volumes 2 million m³
- Property of AFSC
- License: 2019
- Timing: May to late August
- Spatially stratified
- Survey area: 1,500 m depth
- Total collection of groundfish
- Minimum 300 kg
- Mean length: 1,200 kg
- Biological species

**Eastern Bering Sea shelf**
- Larger stratified random
- Sampled in volumes 2 million m³
- Property of AFSC
- License: 2019
- Timing: May to late August
- Spatially stratified
- Survey area: 1,500 m depth
- Total collection of groundfish
- Minimum 300 kg
- Mean length: 1,200 kg
- Biological species

**Northern Bering Sea shelf: Ecosystem Survey**
- Larger stratified random
- Sampled in volumes 2 million m³
- Property of AFSC
- License: 2019
- Timing: May to late August
- Spatially stratified
- Survey area: 1,500 m depth
- Total collection of groundfish
- Minimum 300 kg
- Mean length: 1,200 kg
- Biological species

**Chukchi Sea: Ecosystem Survey**
- Larger stratified random
- Sampled in volumes 2 million m³
- Property of AFSC
- License: 2019
- Timing: May to late August
- Spatially stratified
- Survey area: 1,500 m depth
- Total collection of groundfish
- Minimum 300 kg
- Mean length: 1,200 kg
- Biological species

The AFSC conducts comprehensive bottom-trawl surveys in the Gulf of Alaska (GOA) designed principally to monitor trends in abundance and distribution of groundfish populations. The survey area includes the continental shelf and upper continental slope (out to 1,000 m depth) in the Gulf of Alaska and extends from the Islands of Four Mountains 2,300 km east to Dixon Entrance. The survey began in 1984 and was conducted triennially and then biennially since 1999. During a typical survey, three commercial trawlers are chartered for 75 days each, during late May - early August, sampling the standard 320,000 km² survey area with approximately 550 to 820 survey stations. The catch data result in observations of catch per unit area which are averaged and expanded by survey area to estimate the relative abundance of important groundfish species. These estimates are provided to stock assessment scientists who use the estimates and biological information from the surveys to determine Allowable Biological Catch (ABC) and Total Allowable Catch (TAC) for the North Pacific Fishery Management Council. Other information collected during the survey is used to improve understanding of life history of the fish and invertebrate species and the ecological and physical factors affecting their distribution and abundance.
The GOA Bottom Trawl Survey is a multispecies survey based upon a stratified-random design and the area-swept method of estimating abundance. The GOA shelf and upper slope to 1000 m are divided into strata based upon management area, sub-region, habitat, and depth and overlain with a grid consisting of 5 km by 5 km cells. A station is a successfully sampled trawl deployment or tow within a cell. Stations are selected at random within a stratum, except that stations that are too rocky for a successful tow are excluded. Stations are allocated among strata based upon an optimal allocation scheme using past estimates of abundance and variance, strata areas, and the economic importance of key groundfish species. The Alaska Fisheries Science Center contracts with two or three commercial fishers and vessels to conduct the survey. The skipper is given up to 2 hours to search the station for a suitable, smooth position to deploy the trawl. If the station is untrawlable, then the vessel moves on to the next adjacent station within the stratum and repeats the process. The net is deployed from the vessel and allowed to sink to the seabed while the vessel steams ahead at 3 knots. Acoustic net sensors provide real time depth and opening width measurements from the net. Once on the seabed, the net is towed for 15 minutes after which the net is retrieved, all the time maintaining the target speed of 3 knots. The catch is then processed by the scientific crew who identifies all living organisms, weighs and counts them, and takes biological samples from key groundfish species or other species of interest. CPUE is calculated by dividing the catch weights by the area swept by the net (distance fished by the net times the average net width). The CPUEs are averaged among stations in the stratum and then multiplied by the stratum area to estimate biomass for each stratum. Standard errors are estimated as the simple standard deviation divided by the square root of the sample size within the stratum. Population biomasses are compared among years to establish relative trends or are used directly with other parameters to set ABCs.
5.5 Aleutian Islands Bottom Trawl Survey

Wayne Palsson, Alaska Fisheries Science Center

Aleutian Islands Biennial Bottom Trawl Survey

The AFSC conducts comprehensive bottom-trawl surveys in the Aleutian Islands (AI) designed principally to monitor trends in abundance and distribution of groundfish populations. The survey area covers the continental shelf and upper continental slope (out to 500 m depth) in the Aleutian Islands from Islands of Four Mountains west to Stalemate Bank. The survey has been conducted since 1980, first at approximately three year intervals and then biennially since 2000. NOAA charters two commercial fishing vessels for 70 days each, during June - August, sampling the standard 64,400 km² survey area with approximately 420 survey stations. The catch data result in observations of catch per unit area which are averaged and expanded by survey area to estimate the relative abundance of important groundfish species. These estimates are provided to stock assessment scientists who use the estimates and biological information from the surveys to determine Allowable Biological Catch (ABC) and Total Allowable Catch (TAC) for the North Pacific Fishery Management Council. Other information collected during the survey is used to improve understanding of life history of the fish and invertebrate species and the ecological and physical factors affecting their distribution and abundance.

The AI Bottom-trawl Survey is a multispecies survey based upon a stratified-random design of previously successful stations. The AI shelf and upper slope are divided into strata based upon management area, sub-region, and depth zones (0-100 m, 101-200 m, 201-300 m and 301-500 m). Stations are allocated among strata based upon an optimal allocation scheme using past estimates of abundance and variance, strata areas, and the economic importance of key groundfish species. Additionally, a grid of 5 km x 5 km is overlain over the survey area and stations are targeted to fit within the preselected grid of the station. The net is deployed from the vessel and allowed to sink to the seabed while the vessel steams ahead at 3 knots. Acoustic net sensors provide real time depth and opening width measurements from the net. Once on the seabed, the net is towed for 15 minutes after which the net is retrieved, all the time maintaining the target speed of 3 knots. The catch is then processed by the scientific crew who identifies all living organisms, weighs and counts them, and takes biological samples from key groundfish species or other species of interest. CPUE is calculated by dividing the catch weights by the area swept by the net (distance fished by the net times the average net width). The CPUEs are averaged among stations in the stratum and then multiplied by the stratum area to estimate biomass for each stratum. Standard errors are estimated as the simple standard deviation divided by the square root of the sample size within the stratum. Population biomasses are compared among years to establish relative trends or are used directly with other parameters to set allowable catches.
5.6 Fisheries and Ocean Canada Newfoundland and Labrador Annual Trawl Surveys

Rick Rideout, Canadian Department of Fisheries and Oceans

See workshop abstracts.
### 5.7 Irish Groundfish Survey

David Stokes, Marine Institute Ireland

**Survey Specification**
- **Design:** Stratified Random
- **Frequency:** Annual
- **Time Series:** 2003 - Current
- **Area:** 103,521 km²
- **Stations:** 170
- **Avg. Sweep Area:** 5.30m²
- **Net:** GNV 36x74 ft pants high headline trawl, Type D 15'
- **Timing:** October - December, North - South
- **Sharks:** 15 trawl following Multi-Région Tree analysis based on target species, management regions, depth, gear type required.

---

**Irish Groundfish Survey 2019**

[Map showing survey areas]
Figure 1. 2020 WKUSER Participants.
Annex 1: List of participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Institute</th>
<th>Country (of institute)</th>
<th>E mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lewis Barnett</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:lewis.barnett@noaa.gov">lewis.barnett@noaa.gov</a></td>
</tr>
<tr>
<td>Jennifer Blaine</td>
<td>Wash. Dept. of Fish &amp; Wildlife</td>
<td>USA, Mill Creek, WA</td>
<td><a href="mailto:jennifer.blaine@dfw.wa.gov">jennifer.blaine@dfw.wa.gov</a></td>
</tr>
<tr>
<td>Patrik Börjesson</td>
<td>Dept. of Aquatic Resources, SLU</td>
<td>Sweden</td>
<td><a href="mailto:Patrik.Borjesson@slu.se">Patrik.Borjesson@slu.se</a></td>
</tr>
<tr>
<td>Lyle Britt</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:lyle.britt@noaa.gov">lyle.britt@noaa.gov</a></td>
</tr>
<tr>
<td>Meaghan Bryan</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:meaghan.bryan@noaa.gov">meaghan.bryan@noaa.gov</a></td>
</tr>
<tr>
<td>Jason Conner</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:jason.conner@noaa.gov">jason.conner@noaa.gov</a></td>
</tr>
<tr>
<td>Curry Cunningham</td>
<td>University of Alaska, Fairbanks</td>
<td>USA, Fairbanks, AK</td>
<td><a href="mailto:cjcunningham@alaska.edu">cjcunningham@alaska.edu</a></td>
</tr>
<tr>
<td>Ingeborg de Boois</td>
<td>Wageningen Marine Research</td>
<td>The Netherlands</td>
<td><a href="mailto:ingeborg.deboois@wur.nl">ingeborg.deboois@wur.nl</a></td>
</tr>
<tr>
<td>John Field</td>
<td>SWFSC, NOAA Fisheries</td>
<td>USA, Santa Cruz, CA</td>
<td><a href="mailto:john.field@noaa.gov">john.field@noaa.gov</a></td>
</tr>
<tr>
<td>John Gauvin</td>
<td>Alaska Seafood Cooperative</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:gauvin@seanet.com">gauvin@seanet.com</a></td>
</tr>
<tr>
<td>Owen Hamel</td>
<td>NWFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:owen.hamel@noaa.gov">owen.hamel@noaa.gov</a></td>
</tr>
<tr>
<td>Lisa Hillier</td>
<td>Wash. Dept. of Fish &amp; Wildlife</td>
<td>USA, Olympia, WA</td>
<td><a href="mailto:lisa.hillier@dfw.wa.gov">lisa.hillier@dfw.wa.gov</a></td>
</tr>
<tr>
<td>Anne Hollowed</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:anne.hollowed@noaa.gov">anne.hollowed@noaa.gov</a></td>
</tr>
<tr>
<td>Jim Ianelli</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:jim.ianelli@noaa.gov">jim.ianelli@noaa.gov</a></td>
</tr>
<tr>
<td>Walter Ingram</td>
<td>Mississippi Laboratories, NOAA Fisheries</td>
<td>USA, Pascagoula, MS</td>
<td><a href="mailto:walter.ingram@noaa.gov">walter.ingram@noaa.gov</a></td>
</tr>
<tr>
<td>Elaina Jorgensen</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:elaina.jorgensen@noaa.gov">elaina.jorgensen@noaa.gov</a></td>
</tr>
<tr>
<td>Bill Karp</td>
<td>NOAA Fisheries, Retired/ICES</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:billkarp1950@gmail.com">billkarp1950@gmail.com</a></td>
</tr>
<tr>
<td>Aimee Keller</td>
<td>NWFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:amiee.keller@noaa.gov">amiee.keller@noaa.gov</a></td>
</tr>
<tr>
<td>Stan Kotwicki</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:stan.kotwicki@noaa.gov">stan.kotwicki@noaa.gov</a></td>
</tr>
<tr>
<td>Sven Kupschus</td>
<td>Center for Ecosystem, Fisheries, &amp; Aquaculture Sciences</td>
<td>United Kingdom</td>
<td><a href="mailto:sven.kupschus@cefas.co.uk">sven.kupschus@cefas.co.uk</a></td>
</tr>
<tr>
<td>Ned Laman</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:ned.laman@noaa.gov">ned.laman@noaa.gov</a></td>
</tr>
<tr>
<td>Gwladys Lambert</td>
<td>Center for Ecosystem, Fisheries, &amp; Aquaculture Sciences</td>
<td>United Kingdom</td>
<td><a href="mailto:gwladys.lambert@cefas.co.uk">gwladys.lambert@cefas.co.uk</a></td>
</tr>
<tr>
<td>Name</td>
<td>Affiliation</td>
<td>Location</td>
<td>Email</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------------------------------</td>
<td>-------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Kristin Marshall</td>
<td>NWFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:kristin.marshall@noaa.gov">kristin.marshall@noaa.gov</a></td>
</tr>
<tr>
<td>Michael Martin</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:michael.martin@noaa.gov">michael.martin@noaa.gov</a></td>
</tr>
<tr>
<td>Richard Methot</td>
<td>Office of Science and Technology, NOAA Fisheries</td>
<td>USA, Silver Spring, DC</td>
<td><a href="mailto:richard.methot@noaa.gov">richard.methot@noaa.gov</a></td>
</tr>
<tr>
<td>Cole Monnahan</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:cole.monnahan@noaa.gov">cole.monnahan@noaa.gov</a></td>
</tr>
<tr>
<td>Peter Munro</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:peter.munro@noaa.gov">peter.munro@noaa.gov</a></td>
</tr>
<tr>
<td>Kotaro Ono</td>
<td>Institute of Marine Research</td>
<td>Norway</td>
<td><a href="mailto:kotaro.ono@hi.no">kotaro.ono@hi.no</a></td>
</tr>
<tr>
<td>Zack Oyafuso</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:poyafuso@gmail.com">poyafuso@gmail.com</a></td>
</tr>
<tr>
<td>Robert Pacunski</td>
<td>Wash. Dept. of Fish &amp; Wildlife</td>
<td>USA, Mill Creek, WA</td>
<td><a href="mailto:robert.pacunski@dfw.wa.gov">robert.pacunski@dfw.wa.gov</a></td>
</tr>
<tr>
<td>Wayne Palsson</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:wayne.palsson@noaa.gov">wayne.palsson@noaa.gov</a></td>
</tr>
<tr>
<td>Jon Richar</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Kodiak, AK</td>
<td><a href="mailto:jon.richar@noaa.gov">jon.richar@noaa.gov</a></td>
</tr>
<tr>
<td>Rick Rideout</td>
<td>NW Atlantic Fisheries Centre, DFO</td>
<td>Canada, St. John’s</td>
<td><a href="mailto:rick.rideout@dfo-mpo.gc.ca">rick.rideout@dfo-mpo.gc.ca</a></td>
</tr>
<tr>
<td>Lauren Rogers</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:lauren.rogers@noaa.gov">lauren.rogers@noaa.gov</a></td>
</tr>
<tr>
<td>Chris Rooper</td>
<td>Pacific Biological Station, DFO</td>
<td>Canada, Nanaimo, BC</td>
<td><a href="mailto:Chris.Rooper@dfo-mpo.gc.ca">Chris.Rooper@dfo-mpo.gc.ca</a></td>
</tr>
<tr>
<td>Paul Spencer</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:paul.spencer@noaa.gov">paul.spencer@noaa.gov</a></td>
</tr>
<tr>
<td>David Stokes</td>
<td>Marine Institute, Ireland</td>
<td>Ireland, Rinville</td>
<td><a href="mailto:david.stokes@marine.ie">david.stokes@marine.ie</a></td>
</tr>
<tr>
<td>Ian Taylor</td>
<td>NWFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:Ian.Taylor@noaa.gov">Ian.Taylor@noaa.gov</a></td>
</tr>
<tr>
<td>James Thorson</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:james.thorson@noaa.gov">james.thorson@noaa.gov</a></td>
</tr>
<tr>
<td>Tien-Shui Tsou</td>
<td>Wash. Dept. of Fish &amp; Wildlife</td>
<td>USA, Olympia, WA</td>
<td><a href="mailto:tien-Shui.Tsou@dfw.wa.gov">tien-Shui.Tsou@dfw.wa.gov</a></td>
</tr>
<tr>
<td>Paul von Szalay</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:paul.von.szalay@noaa.gov">paul.von.szalay@noaa.gov</a></td>
</tr>
<tr>
<td>Eric Ward</td>
<td>NWFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:eric.ward@noaa.gov">eric.ward@noaa.gov</a></td>
</tr>
<tr>
<td>Nicola Walker</td>
<td>Center for Ecosystem, Fisheries, &amp; Aquaculture Sciences</td>
<td>United Kingdom</td>
<td><a href="mailto:nicola.walker@cefas.co.uk">nicola.walker@cefas.co.uk</a></td>
</tr>
<tr>
<td>Kresimir Williams</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:kresimir.williams@noaa.gov">kresimir.williams@noaa.gov</a></td>
</tr>
<tr>
<td>Cynthia Yeung</td>
<td>AFSC, NOAA Fisheries</td>
<td>USA, Seattle, WA</td>
<td><a href="mailto:cynthia.yeung@noaa.gov">cynthia.yeung@noaa.gov</a></td>
</tr>
</tbody>
</table>
Annex 2: Resolution

WKUSER – Workshop on unavoidable survey effort reduction


a) The workshop will reflect on the current processes used in dealing with unavoidable reductions in survey efforts and examine the existing coping strategies (e.g. spatial coverage, survey frequency, or sampling density) and their qualitative consequences (Science plan codes 3.2);

b) Develop key quality metrics that can be used to describe “total survey uncertainty” for survey derived indices of abundance for common survey designs (Science plan codes 3.2, 3.3);

c) Define “changes to survey designs” that require inter-survey calibration and what changes can be resolved by a model-based approach to index generation (Science plan codes 3.2, 3.3);

d) Consider the development of methods that aim to provide quantitative decision-making tools that describe the effects on the quality of the survey deliverables and ultimately advisory products (Science plan codes 3.3).

WKUSER will report by 15 February 2020 for the attention of the ACOM and SCICOM.

Supporting information

| Priority | Marine surveys are expensive and under recent budgetary and political pressures a number of decisions on survey implementation have had to be made at very short notice and with little opportunity to evaluate different options for effort reductions the effects of which will only become apparent in the next few years. Such changes are likely to be a recurring theme, and it is in the interest of national governments making the decisions and ICES using such information for their advice to have a better understanding of their effects on stock assessment advice and a clearer understanding of the mitigation measures that can be implemented to minimise the impact of such events. |
| Scientific justification | Most survey programs are at one time or another asked to make substantial short term savings. Usually these requests leave little time for planning let alone evaluation so there is a real need to develop methods that provide a better understanding of the risks of different implementation options, an investigation of methods that can help to compensate for some of the information loss, and lastly under which survey design and survey objectives these methods are most appropriate. Often survey scientist / managers are having to make decisions on the fly, the consequences of which are poorly understood. Having a framework or a set of methods that can be applied to the specific problem is highly valuable together with summarizations of findings for general cases, which allow survey scientist to |
make decisions in the absence of data or the opportunity to evaluate options statistically.

<table>
<thead>
<tr>
<th>Resource requirements</th>
<th>Many different approaches to evaluate effects and survey options have been developed independently at different times in response to specific cases. A large part of this work is to research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resource required to undertake additional activities in the framework of this group is negligible.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>Unknown at present but likely between 10 and 20 participants</td>
</tr>
<tr>
<td>Secretariat facilities</td>
<td>None.</td>
</tr>
<tr>
<td>Financial</td>
<td>No financial implications.</td>
</tr>
<tr>
<td>Linkages to advisory committees</td>
<td>There is a direct link with the advisory committee as they require knowledge on the sensitivity of the advice to changes in surveys in order to provide precautionary advice when survey information is compromised.</td>
</tr>
<tr>
<td>Linkages to other committees or groups</td>
<td>The workshop should link closely back to WGISDAA which will maintain the tools / methods and broaden the approach over time. Work with stock assessment WG is thought to be essential.</td>
</tr>
<tr>
<td>Linkages to other organizations</td>
<td>The work of this group is closely aligned with similar work in FAO and in the Census of Marine Life Programme.</td>
</tr>
</tbody>
</table>