SUMMARY REPORT:

REVIEW OF CANDIDATE "BEST PRACTICE" CRITERIA

EXECUTIVE SUMMARY

Reference points and stock status are part of a complete precautionary approach (PA) framework under Canada's PA Policy. Twenty-five respondents provided input on four inter-connected candidate "best practice" criteria for the defensible selection of indicators, limit reference points, and stock status metrics. The refined set of criteria is summarized as follows.

Consistent with an objective to avoid serious harm to the stock

An objective to avoid *serious harm* is central to Canada's PA Policy. It is commonly described in general terms of recruitment overfishing, impaired productivity, loss of resilience or an ability to recover from perturbation, depensation (Allee effects), or very depleted stock states where dynamics become uncertain. Serious harm could also include loss of genetic diversity, structure or distribution, contraction of age/size structure, and extirpation. It could be caused by overfishing or by ecosystem impacts (altered predator-prey dynamics, habitat loss, etc.)

A wide range of stock status indicators could be considered by which to define thresholds to serious harm. Whatever indicators are chosen should usually be representative of the entire "stock," or at least a representative subunit of the stock, and may show proportionality with stock attributes they are intended to measure, but the relative importance of these characteristics in driving the choice of indicator may be context-dependent.

Based on best available information

Best available information, which provides the evidence basis for choice of indicators, reference points or stock status, will vary from stock to stock. Broadly speaking, it can be described as: relevant (appropriate), peer-reviewed, verified and validated, inclusive, objective, timely, transparent, open and accessible, accounts for uncertainties, accurate, consistently gathered, conflicting or alternative information considered, adequate, representative, reproducible (repeatable), clear and complete.

When possible, evidence of serious harm and meta-analysis (basic biological information) may be preferred rationales on which to base choices of limit reference points, over policy guidance (40% B_{MSY}) or common practice. However, these rationales are not exclusive, and choices may based on other stock- or context-specific "best available" information, such as specific stock dynamics or traditional knowledge, and will be ultimately constrained by data poverty.

Operationally Useful

At minimum, indicators, LRPs and therefore status should be feasible to measure or estimate. Other operational reasons to consider indicators, LRPs or status metrics may be: cost-effectiveness (ease of measuring or estimation), communicability (ease of understanding), simplicity, the role that LRPs or stock status may play in HCRs or triggering the need for a rebuilding plan, or the need to estimate stock status to evaluate either trends over time or performance of management measures. These reasons are not mutually exclusive, and the relative priority of reasons that make a choice of indicator, LRP or metric of status "operationally useful" will be context-dependent, both within or across traditional versus procedural paradigms.

Reliably Estimable

Reliable estimation of limit reference points or stock status metrics can mean acceptable consistency, accuracy or precision of estimates (i.e., acceptably low variance or low bias), and robustness to a range of possible uncertainties (assumptions, stock scale, data points and/or model structure). Reliability of reference points or status may also depend on reliable and consistent data collection. Reliability may be evaluated by examining uncertainty in estimates, sensitivity tests, evaluating the reasonableness of assumptions, simulation testing and/or by comparison to other similar stocks.

INTRODUCTION

Four weeks in advance of the workshop, participants were invited to answer 22 questions as a pre-requisite to workshop exercises. The questions were aimed to: a) elicit input from workshop participants regarding four candidate guidance criteria that help to define what makes a “good” or "best practice" LRP/stock status indicator, and b) encourage participants to think about different rationales for choosing LRPs/indicators. Twenty-five respondents completed this activity.

The four candidate guidance criteria are:

1) Consistent with an objective to avoid serious harm to the stock,
2) Based on best available information,
3) Operationally useful,
4) Reliably estimable.

A "stock" can be defined in many ways, but for this purpose, can be considered a semi-discrete group of aquatic animals (fish, invertebrate, marine mammals) with some definable attributes in common that are of interest to managers. An "attribute" is a quality, characteristic or feature of the stock. An "indicator" is some measurement that provides information on the state of the stock, and may include model-based estimates of biomass, fishing mortality or exploitation rate, or suitable proxies for these such as survey indices. A reference point is some value of an indicator that represents a target or threshold that management measures aim to either achieve on average, or alternatively surpass (or avoid breaching); limit reference points are a type of threshold. Lastly, "stock status" is the relationship between some estimate of the indicator (e.g., the most recent or "current" estimate) and a reference point (e.g., the LRP). Stock status is a metric (or statistic). It can be used to evaluate performance (a statistic used to evaluate management measures, and can be calculated from simulated states of nature). It can also be used for monitoring (a statistic estimated and tracked retrospectively to provide information on realized stock or fishery performance over time).

The process of selecting an LRP may involve decisions made by stock assessors at several steps: selecting stock attributes by which to interpret serious harm, evaluating and selecting indicators to represent those attributes, evaluating and selecting thresholds to be LRPs, and estimating and reporting stock status as a metric.

There are various means by which to approach these steps and select an indicator/LRP, including theoretical, historical (both model-based) or empirical versions of fractions of K, B_0, or B_{MSY}, points from stock-recruitment relationship, $B_{recover}$, or other thresholds. "Model-based" indicators are estimates of quantities generated from models, while "empirical" indicators are directly observed (catch, catch-per-unit-effort or CPUE, survey indices). LRPs may be derived from theoretical values such as B_{MSY} (biomass associated with fishing mortality at maximum sustainable yield), B_0 (unfished biomass), K (carrying capacity), or a stock-recruitment relationship. They may also be derived historically (points along a time series; e.g., $B_{recover}$, the lowest biomass from which a stock has demonstrated a secure recovery).

RESULTS

Candidate Criterion 1: Indicators and LRPs should be "consistent with an objective to avoid serious harm to the stock"

This criterion first requires an interpretation of what "serious harm" means as an undesirable stock state to be explained, and then additional considerations to operationalize those definitions in a way appropriate to the stock. "Serious harm" is commonly described in general terms of recruitment overfishing, impaired productivity, loss of resilience or an ability to recover from perturbation, depensation, or very depleted stock states where dynamics become uncertain.

Question 1. How might you define and diagnose "recruitment overfishing", "loss of resilience", "impaired productivity" or "depensation"? What data do you need? (pick one of the terms to define and explain)

Respondents explored and described serious harm via its common interpretations, particularly recruitment overfishing, as shown in Figure 1. Serious harm in the context of recruitment overfishing was defined by participants as:

- Insufficient biomass, abundance or number of spawners to support "normal" recruitment levels, the ability to produce offspring, or the ability for the stock to maintain or replace itself
- Stock levels associated with the steep parts of the stock-recruitment curve, or those to the left of the inflection or break-point
- Associated with an increased risk of stock collapse if prolonged or combined with poor environmental conditions
- A result of overfishing or high exploitation rates

Impaired productivity was defined by participants as:

- Persistent periods of low productivity (in relation to some historical period)
- Non-responsiveness in terms of increasing biomass to changes such as reduction in fishing pressure

Loss of resilience was defined as the inability to adapt to environmental perturbations.

Depensation (also called predator pits or Allee effects) was defined by participants as:

- Inverse density-dependence, or declines in the per capita rate of growth of a population as stock density or size declines
- Caused by loss of reproductive potential through the loss of large fecund individuals or through density-dependent processes (e.g., fertilization success)

Respondents suggested a wide range data may be needed for a diagnosis of serious harm, including:

- Time series of declining or persistent low recruitment, spawning stock biomass, survey indices of biomass, and/or contracting size/age structure

- Time series of catches, landings or fishing mortality
- Time series showing response of stock to decreased fishing pressure
- Analyses of surplus production (and data required to generate these estimates)
- Information on fecundity, size/age structure, growth and natural mortality
- Understanding of the relationship between reproductive biology and density
- Ability to estimate the stock-recruitment relationship (longer the time series the better) and a threshold from that relationship
- Ability to apply alternative stock-recruitment models to test for inflection points due to depensatory effects
- Negative trends in the residuals of the stock-recruitment relationship
- Estimates of recruits per spawner < 1 (i.e., below the replacement line)

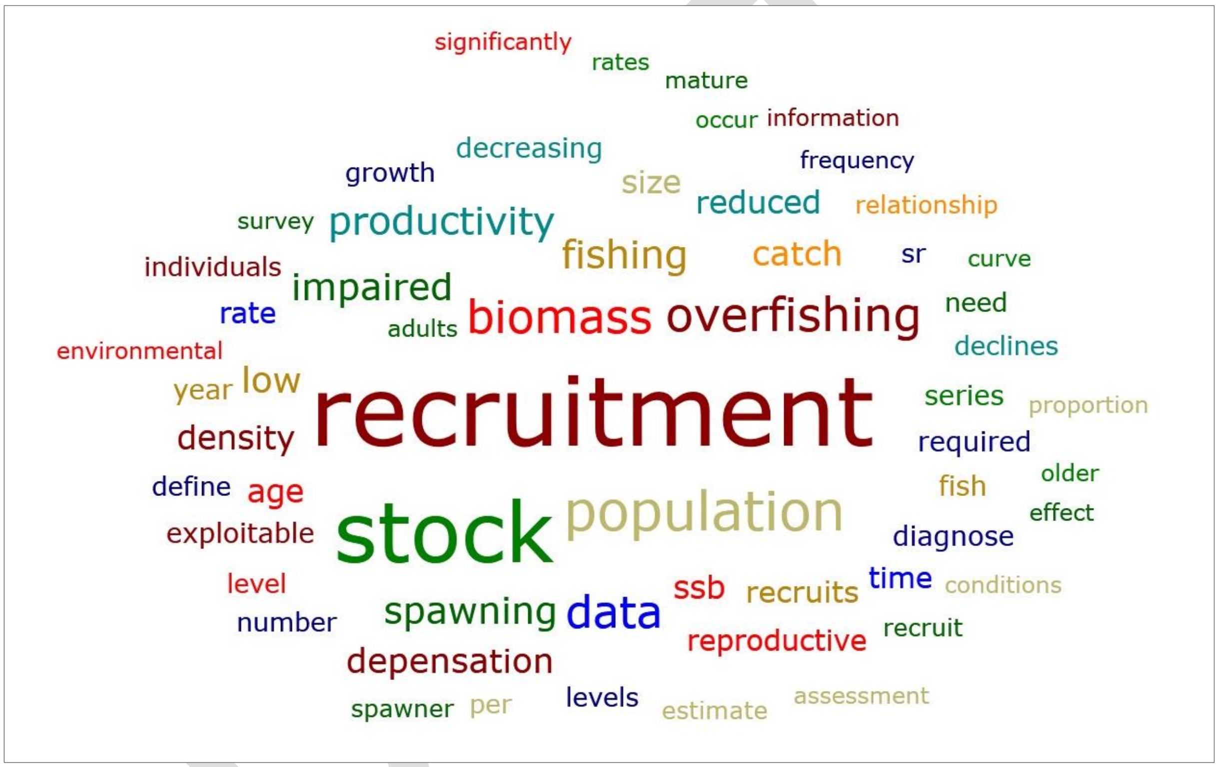

Figure 1: A word cloud of all terms used by respondents at least four times in their descriptions of various traditional approaches to serious harm, where font size is proportional to the number of times it was mentioned.

Question 2. In your opinion, are there other undesirable stock states or outcomes that represent "serious harm" and that are not mentioned above?

Respondents suggested several other outcomes that could constitute serious harm (Figure 2), including:

- Loss of genetic diversity (six participants)
- Changes in stock structure or behavior (e.g., spawning and movement; range shifts; serial depletion) (five participants)

- Taking into account ecosystem considerations into thresholds of serious harm, or considering ecosystem thresholds (as a result of considering predator-prey interactions, habitat loss, or prevailing conditions) (three participants)
- Smaller size or age structure (two participants)
- Extirpation (two participants)
- Poorer condition, evolutionary shifts toward smaller body size and earlier age of maturity, growth overfishing and ecosystem overfishing (one participant each)

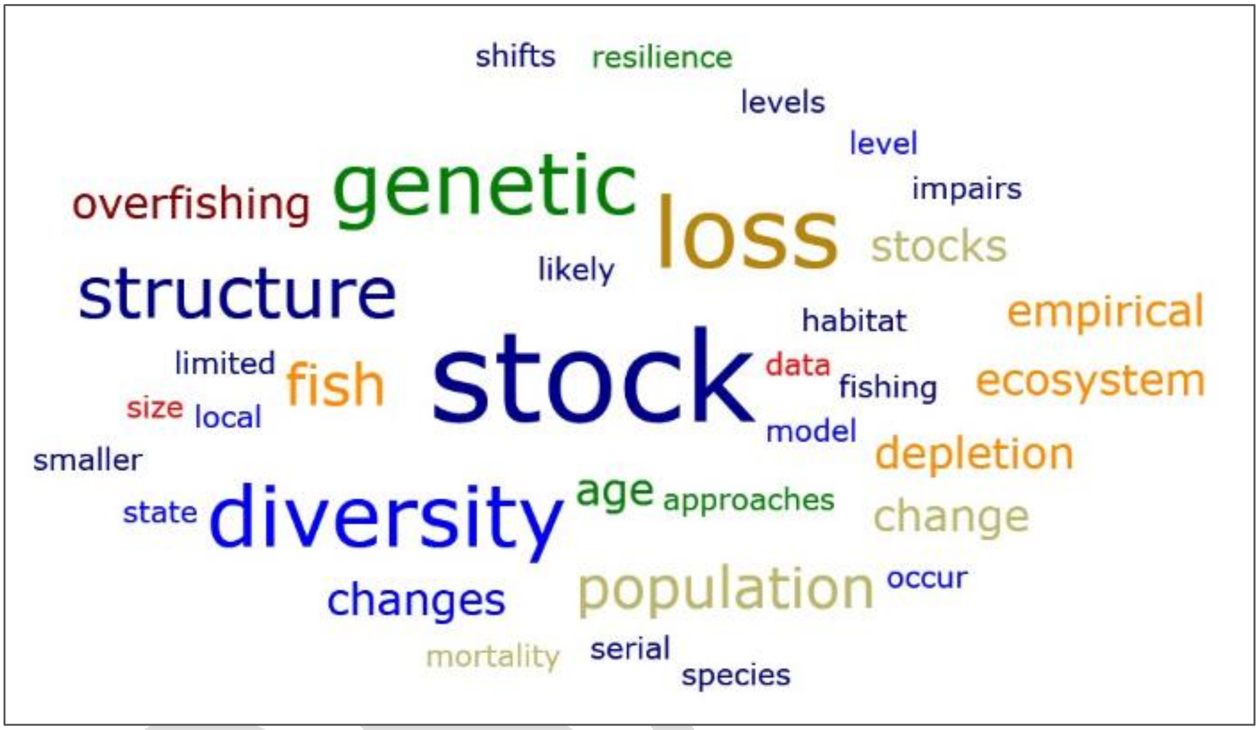

Figure 2: A word cloud of all terms used by respondents at twice when identifying other states or outcomes that could be associated with serious harm, where font size is proportional to the number of times it was mentioned.

Question 3. What are some of the possible difficulties you could face when trying to identify a threshold to states of serious harm using any given indicator, consistent with the undesirable stock states discussed or identified above?

Respondents identified a number of potential difficulties in selecting a LRP consistent with serious harm, including:

- No or poor-quality indicators or data (short time series, fishery-dependent indicators only, or highly uncertain) (ten respondents)
- Missing or uncertain life history information (six respondents)
- Picking one threshold from multiple options (because of uncertainty in estimates, that serious harm is a continuum or hasn't been observed, and subjectivity in choice) (six respondents)
- A lack of a stock-recruitment relationship (or lack of one with a break-point) (five respondents)
- Non-stationary environmental processes (five respondents)
- Picking only one indicator for serious harm when several may be pertinent (three respondents)

- Complex stock structure or time and model poverty (one participant each)

Question 4. Common stock attributes that could have thresholds associated with serious harm include the size of the stock (measured with model-based or empirical indicators of abundance, biomass, or subsets such as spawning stock biomass), egg production, or surplus production. Can you suggest any alternative stock attributes that could have thresholds associated with serious harm? What indicators could be used to estimate (or approximate) those attributes?

Respondents suggested some additional stock attributes that might have thresholds that could be associated with serious harm, such as genetic diversity, sex ratios or the abundance of males (i.e., resulting in sperm limitation), environmental tolerance, and distribution.

Possible indicators of genetic diversity could be the number of genetic polymorphisms or rate of polymorphisms, measures of heterozygosity or genetic flow, the number of alleles, allelic richness and variants among alleles, and the average number of alleles per locus, and the proportion of polymorphic loci.

Other possible indicators that could be used to estimate or approximate important stock attributes (and from which a threshold to serious harm could be selected) were proposed. These indicators included the number of (or ratio of) males to females, age and size composition of the catch, length-at-age, body condition, and the proportion of mature fish. CPUE could serve as an indicator of stock abundance under certain conditions, and stock distribution could be measured in terms of area occupied or the distribution of spawning. A variety of indirect indicators of stock status were suggested as well, including removal rates (or fishing mortality), the extent of habitat alteration, the occurrence of extreme heat events, the distribution or rate of spread of aquatic invasive species, concentrations of pollutants, predator or prey densities, water temperature and oxygen concentration.

Question 5. Indicators may be model-based estimates (e.g., spawning stock biomass) or empirical (e.g., fishery-dependent or independent indices). How important is it to consider whether the chosen stock status indicator is...

a. **Representative of the state of the entire "stock"**
b. **Proportional (exhibits a linear relationship with) the stock attribute(s) selected to represent serious harm**

A majority of respondents felt that it was always important for indicators to be representative of the state of the entire "stock," and with the exception of two respondents, most also felt that the indicator should be proportional to the stock attribute it was intended to represent (Figure 3). Respondents who felt that these characteristics were more important in some situations than others were asked to expand upon that choice in question 6.

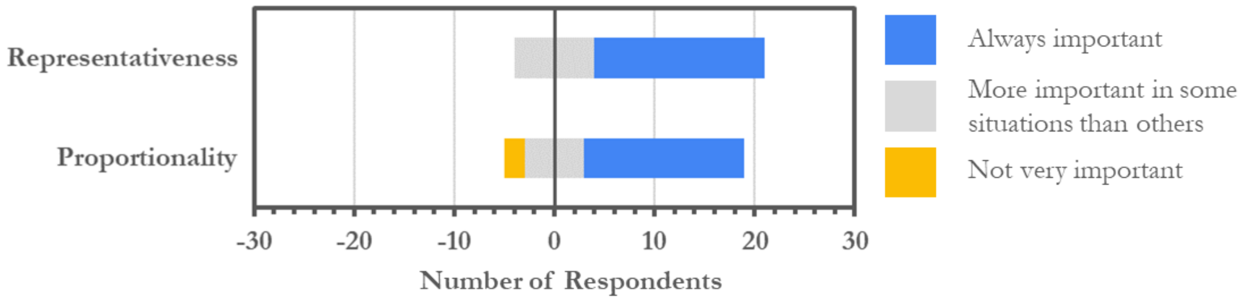

Figure 3: Diverging stacked bar chart showing the number of respondents ranking the relative importance of two characteristics in selecting an indicator to represent a stock attribute relevant to serious harm. Responses have been standardized; positive responses are graphed to the right of the 0 on the X axis, in opposition to negative responses. Note that neutral responses straddle positive and negative.

Question 6. If you selected "More important for some situations than others" in the question above for either proportionality or representativeness, can you explain what those situations are?

For representativeness, respondents noted that:

- It depends on the definition of a "stock"
- For stocks that are meta-populations, indicators could be used only for certain subunits, or for all subunits if a model is used that explicitly accounts for spatial structure
- In salmon management units (SMUs), which consist of multiple conservation units (CUs) some indicators may be used or be appropriate for some CUs but not others and this may have important implications for which management measures may apply to which CUs (and which threats each are subject to)
- Indicators used in multi-indicator frameworks may reflect only one component of a stock, where that component then serves as an "indicator" for the entire stock
- Empirical fishery-dependent indicators already do not inform on the state of the entire "stock" because they inform on only that component available to the fishery
- There may be smaller subcomponents of the stock that are not represented by a given indicator (and may not be assessed or frequently targeted by management measures), but the conservation risk to the stock may be low so long as they are small or not under intense fishing pressure. If these subcomponents were subject to fisheries or other sources of mortality, then representing them would become more important.
- Representativeness may be more challenging for large stocks versus small stocks (geographically speaking), and consideration should be given for larger stocks where distribution would not be consistent throughout the entire area

For proportionality, respondents noted that:

- Linear relationships between indicators and attributes are often assumed in models
- Linear relationships may not always exist (Type II or III responses)
- Some LRPs can be derived in the absence of some assumptions of linearity ($B_{recover}$)
- Empirical indicators often exhibit linear relationships with stock attributes

- Proportionality may not be essential for use of indicators in all cases as long as the indicator can be assigned a threshold to serious harm
- Non-linear relationships between indicators and attributes that are unknown or not accounted for can be problematic
- Indicators used in multi-indicator frameworks may have complex relationships to stock attributes

Candidate Criterion 2: Indicators and LRPs should be "based on the best available information" for the stock.

"Best available information" will vary from stock to stock and might produce different reasons for selecting an indicator/LRP. Some jurisdictions, like the US, provide criteria for evaluating scientific information as "best" that include such factors as relevance, inclusiveness, objectivity, transparency and openness, timeliness, verification and validation, and peer review, as appropriate. (https://www.ecfr.gov/current/title-50/chapter-VI/part-600/subpart-D/section-600.315)

Question 7. What criteria would you use to define what makes "best available information"?

Respondents suggested a range of terms that could be used to define "best available information". Several respondents found this a difficult question and potentially subjective. Some felt that "best available" does not mean perfect, but it also doesn't mean that any available information qualifies (i.e., while having some information is better than none, it might not be sufficient to support advice; 'information' is not the same as 'data', and risks associated with the quality of the information should be considered). Several noted Traditional Knowledge in the context of inclusiveness. Others cautioned that it may not be possible to maximize all criteria at the same time, particularly timeliness.

Criteria	No. of Respondents
Relevant (US criterion) [i.e., recent and pertaining to the stock/area]	9
Peer Reviewed (US criterion)	9
Verified and Validated (US criterion)	7
Inclusive (US criterion)	6
Objective (US criterion)	6
Timely (US criterion)	6
Transparent and Open (US criterion)	4
Acknowledges limitations and uncertainties	4
Accurate	3
Available and Accessible	3
Consistently gathered	2
Reliable	2
Alternative or conflicting information considered	2
Appropriate	2
Adequate	2

Criteria	No. of Respondents
Representative	2
Defensible	1
Reproducible	1
Clear	1
Complete	1

Question 8. What priority would you assign different reasons to select a given threshold as LRP? Rank in order from more to less preferred (and assume all information is available). Please use each rank only once (no ties).

a. **Consistent with policy choice or guidance [40% BMSY]**
b. **Common practice (similar to LRPs elsewhere)**
c. **Direct evidence (demonstrated serious harm in this stock)**
d. **Meta-analysis (using basic biological information, e.g., "low productivity", or analogies from other stocks)**

Respondents strongly preferred to base a choice of LRP on evidence of serious harm for the stock (assuming such information was available), or failing that, based on meta-analyses or basic biological information over a choice consistent with policy guidance, or to an LRP selected because it might be common practice (Figure 4).

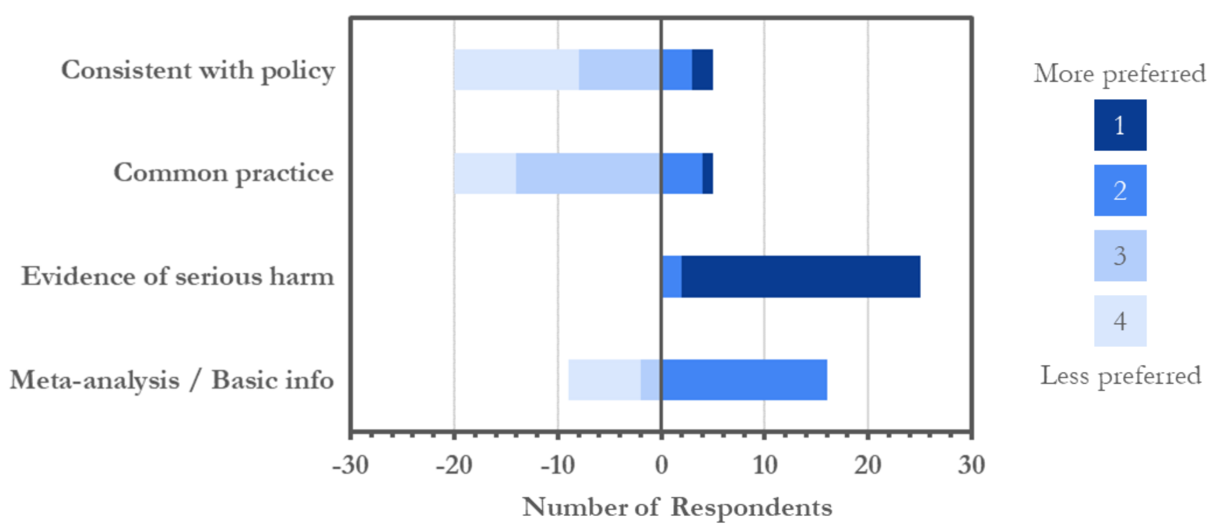

Figure 4: Diverging stacked bar chart showing the number of respondents ranking preferred rationales on which to base the selection of an LRP. Responses have been standardized; positive responses are graphed to the right of the 0 on the X axis, in opposition to negative responses.

Question 9. What reasons might produce LRPs that are more likely to be challenged? Rank in order from more to less likely to be challenged (and assume all information is available). Please use each rank only once (no ties).

a. **Departing from policy choice or guidance [i.e., not 40% BMSY]**
b. **Less common practice (dissimilar to LRPs elsewhere)**
c. **Lack of direct evidence (demonstrated serious harm in this stock)**
d. **No meta-analysis or basic biological information (e.g., “low productivity”, or analogy to other stocks)**

Respondents felt that LRPs selected in the absence of evidence of serious harm, or that deviated from common practices, to be the most likely to be challenged (Figure 5). LRP choices made in the absence of meta-analyses or basic biological information, on the other hand, were considered to be the least likely to be challenged.

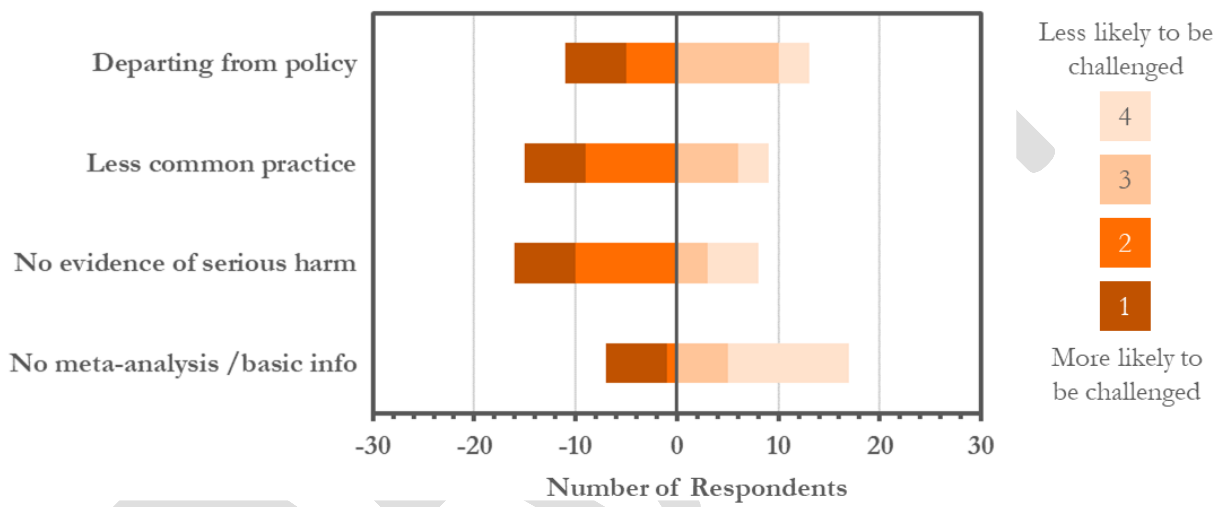

Figure 5: Diverging stacked bar chart showing the number of respondents ranking the likelihood that a choice of LRP that deviates from, or does not use, particular rationales will be challenged. Responses have been standardized; positive responses are graphed to the right of the 0 on the X axis, in opposition to negative responses.

Question 10. Are there any reasons to select an LRP that are not included in the list of four above?

Several respondents suggested additional reasons to select LRPs:

- LRPs may be considered appropriate for stocks based on knowledge of its dynamics in a way that is unrelated to other stocks
- LRPs may be selected due to unknown dynamics below certain points ($B_{recover}$)
- LRPs may be selected because they are easy to understand, or easy to calculate/estimate reliably
- LRPs may be selected in the context of the management measures they may be linked to
- LRPs may be selected after consultations with fishery interests to determine agreed-upon undesirable states
- Data poverty may drive LRPs selected

- LRPs pre-dating policy guidance may also be considered appropriate for a stock in some circumstances, e.g., where there is large support for the choice

Question 11. What evidence would you use as a rationale to depart from the provisional default LRP of 40% B_{MSY} identified in the PA Policy? (For this question, assume you can at least estimate B_{MSY}.)

In general, respondents felt that there were a variety of reasons related to estimability, simulation testing and a variety of biological rationales that could be used to depart from a provisional default of 40% B_{MSY}. However, 40% B_{MSY} was noted as a convenience and may be a reasonable starting point in many cases.

- Where estimates of B_{MSY} are very uncertain or unreliable, and especially if B_{MSY}:B_0 is low
    - E.g., due to uncertainties in model parameters pertaining to selectivity, steepness or natural mortality
- Where estimates of B_{MSY} may be impacted by non-stationarity (changing productivity regimes)
- Where there is evidence of serious harm (in a given stock, or similar stocks) at levels higher than 40% B_{MSY} and thus 40% B_{MSY} is not expected to be sufficiently precautionary
    - E.g., Stocks with late maturity and/or slow growth
- Where other LRPs have stronger rationales and support given the historical data for the stock (or similar stocks), including:
    - Allee thresholds,
    - Periods of depressed productivity,
    - Demonstrated recoveries in the past below levels corresponding to 40% B_{MSY}
- Where other LRPs show better performance in closed-loop simulation
    - E.g., S_{gen} (the number of spawners that result in recovery to S_{MSY} within one generation in the absence of fishing under equilibrium conditions) for salmon
- Where the stock is not a finfish or has unusual life histories
- Where the stock is extremely productive
- Where the role that the LRP plays as an operational control point for management measures in relation to the USR may be a concern (e.g., narrower Cautious zones)

Candidate Criterion 3: Indicators and LRPs should be "operationally useful."

Like "best available information," what is considered operationally useful may vary widely from stock to stock. Here, "operationally useful" can be considered to mean the choice of indicator and LRP are "ready to use" in advice and management of a particular stock.

Question 12. What could be operational reasons to select a given indicator or LRP (the combination of which yields stock status)? Evaluate each in terms of their relative importance. Please use each rank only once (no ties) *HCR = harvest control rule

a. Feasible (possible) to estimate status

b. **Cost effective to estimate status frequently (i.e., considering costs of data collection and analysis, including modelling)**
c. **Status is easy to communicate and understand**
d. **Status is, or could be, an input (or LRP an operational control point) to a HCR**
e. **Status needed as monitoring metric: Can be applied to existing data/time series to estimate past/current status and trends**
f. **Status needed as performance metric: Enables provision of forward-looking advice (i.e., forecasts or simulations)**

Of all the operational reasons to pick an indicator or LRP, feasibility was ranked the most important by a strong majority of respondents (Figure 6). Cost-effectiveness and the role that indicators or LRPs might play in HCRs were generally considered less important. The role of status as a metric for monitoring status and trends was considered slightly more important than as a role as a performance metric in forward-looking advice.

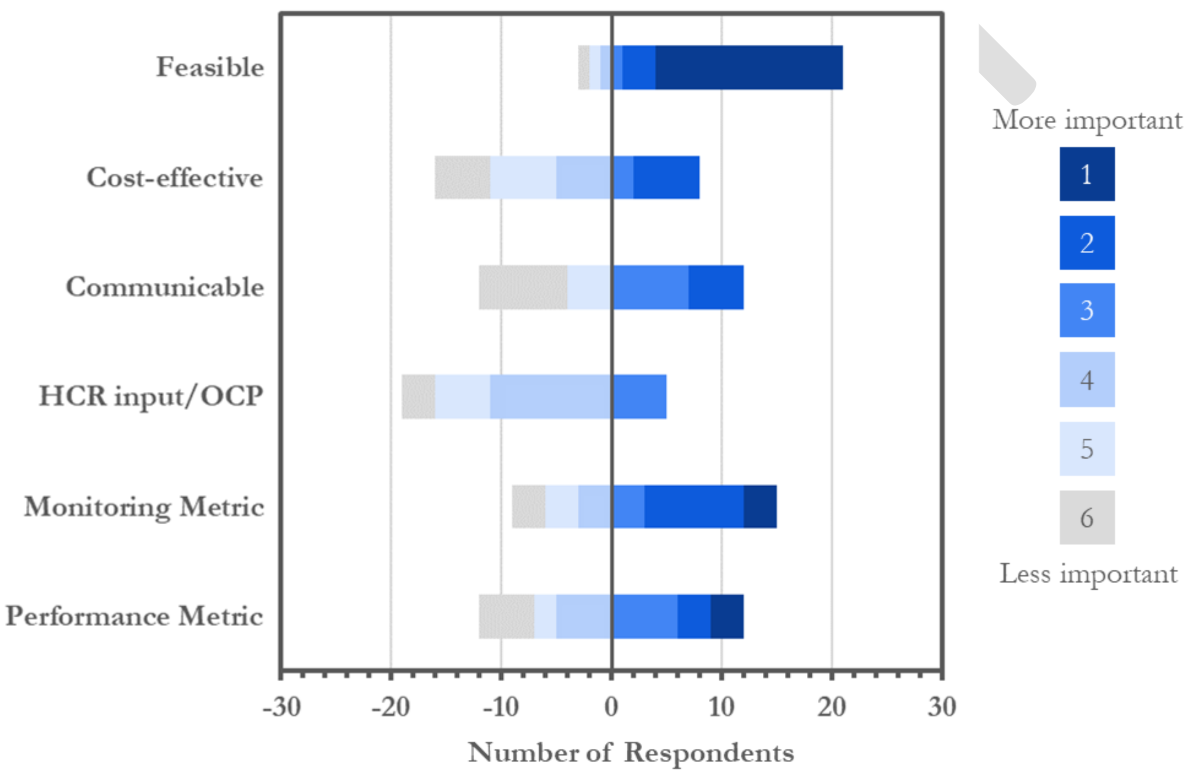

Figure 6: Diverging stacked bar chart showing the number of respondents ranking the importance of different operational reasons to select an indicator, LRP or method of estimating stock status. Responses have been standardized; positive responses are graphed to the right of the 0 on the X axis, in opposition to negative responses.

Question 13. Can you suggest any other considerations or reasons (apart from the six given above) that would make stock indicators or LRPs operationally useful?

Other operational considerations to select indicators or LRPs that were suggested by respondents included:

- Ease of measuring
- Utility in achieving management objectives such as avoiding collapse
- Supported by Traditional Knowledge
- Minimizing complexity – e.g., if two options perform equally effectively, it might be desirable to keep it simple and use only one
- Whether status is estimated with enough certainty to be used by managers to trigger management actions

Some respondents noted that ranking operational reasons to select indicators or LRPs in Question 12 was challenging:

- The options are not mutually exclusive and can be considered simultaneously
- Feasibility is mandatory; if something is not feasible or cost-effective, then none of the other may considerations matter
- Relative importance may also depend on the stock in question and that it was important to be flexible with operational considerations that may be pertinent to implementation in a variety of contexts

Question 14. Would the relative importance of different operational reasons to choose a particular indicator/LRP (question 12) change with different paradigms? If yes, please explain.

Most respondents (16) felt that yes, the relative priority of operational reasons to select a given indicator or LRP would change depending on the advice or management paradigm for the stock in question. Others felt that the operational reasons either might not or would rarely change (4), or weren't certain (3). The reasons for a change in priority were given as follows:

- The need to move to simulation for the purpose of evaluating management procedures
  - Monitoring metrics are more important for traditional stock assessment and the *Sustainability Survey*, while performance metrics are more important for MSE
- The need to consider uncertainty differently in advice given for management measures
- The need to accommodate different levels of data poverty or in the event of the rejection of models
- Different paradigms may result in advice being given at different frequencies (impacting cost-effectiveness)
- Several respondents noted that ease of communication may become more important in MSEs or consultative processes where HCRs are developed
- Cost-effectiveness of frequent updates may be more important for traditional stock assessment

Some noted that the choice of paradigm itself may be driven by operational reasons, with procedural approaches being increasingly preferable as data poverty increases. Furthermore, if the stock is in a very depleted state and is a candidate for rebuilding, a role of the LRP as HCR operational control point may be less prioritized over other operational considerations.

Candidate Criterion 4: LRPs and indicators should be "reliably estimable."

Reliable estimation can mean acceptable consistency or precision of estimates (i.e., acceptably low variance, low bias), robustness to uncertainty in assumptions, and/or robustness to uncertainty in stock scale.

Question 15. Are there other ways that you would define "reliable estimation"?

Seven respondents suggested that "reliable" LRPs and indicators should be supported by "reliable" data (consistent, feasible and sufficiently frequent data collection; i.e., not sporadic). Another noted that "reliability" could be defined via peer-review.

Several respondents highlighted robustness to uncertainty and expanded upon it as follows:

- Estimates should not be influenced by individual data points
- Estimates should be robust to model structure, e.g., via cross-validation or retrospective analysis

Question 16 How might you evaluate whether a LRP is reliably estimated?

Respondents provided a range of suggestions for evaluating reliable estimation, including:

- Examining the accuracy and precision of indicator and LRP estimates (i.e., confidence bounds)
- Evaluating the sensitivity of LRP estimates to missing data, assumptions or priors
- Simulation testing of the estimation procedure under different scenarios
- Considering the reasonableness of the assumptions
- Examining model diagnostics to evaluate the fit (and therefore of derived reference points)
    - Convergence, examination of established statistical criteria, magnitude of the standard error of the estimate, in Bayesian applications whether the posterior probability updated from the prior (is the result data-driven or prior-driven)
- Retrospective analysis
- Cross-validation using statistical techniques
- Comparison of estimates with different model-based approaches
- Comparison with other stocks of the same species or to similar species
- Over time, with periodic review of stock status and associated monitoring with consistent methods
- Comparison with time series to determine association with evidence of serious harm
- Ensuring estimates are reproducible
- More generally, that this must be done via peer review

Question 17. What characterizes a "reliable" stock status indicator?

Respondents characterized stock status indicators as reliable when they were associated with:

- Low uncertainty; accuracy and precision; low bias
- Robust to changing conditions
- Ability to provide information on changes in stock status quickly
- Ease of measurement
- A basis of trustworthy information that is readily and consistently available to support monitoring (and which is expected to continue to be so in the future)
- Data-richness, in general
- Consistency with other metrics of stock health
- Evidence that they are thought to be or demonstrated to be consistent with the state and dynamics of the stock and its attributes of "productivity" over time, even if not linearly related to those stock attributes
- Proportional to "true" abundance in a relative time series (assumption of constant catchability)
- Derivation from robust and consistent analytical methods
- Monitoring that subsequently confirms prior simulation testing, if simulation is used
- Being reliant on minimal or reasonable assumptions
- Being well-documented
- Providing information needed to sustainability manage the fishery and achieve desired goals

Question 18. Please rank the proposed minimum "best practice" criteria for LRPs and indicators that were highlighted above, from more to less important. Please use each rank only once (no ties).

A majority of respondents ranked a criterion of "consistent with an objective to avoid serious harm to the stock" as the most important of the four candidate criteria explored in this study, and tended to rank "operationally useful" as least important (Figure 7).

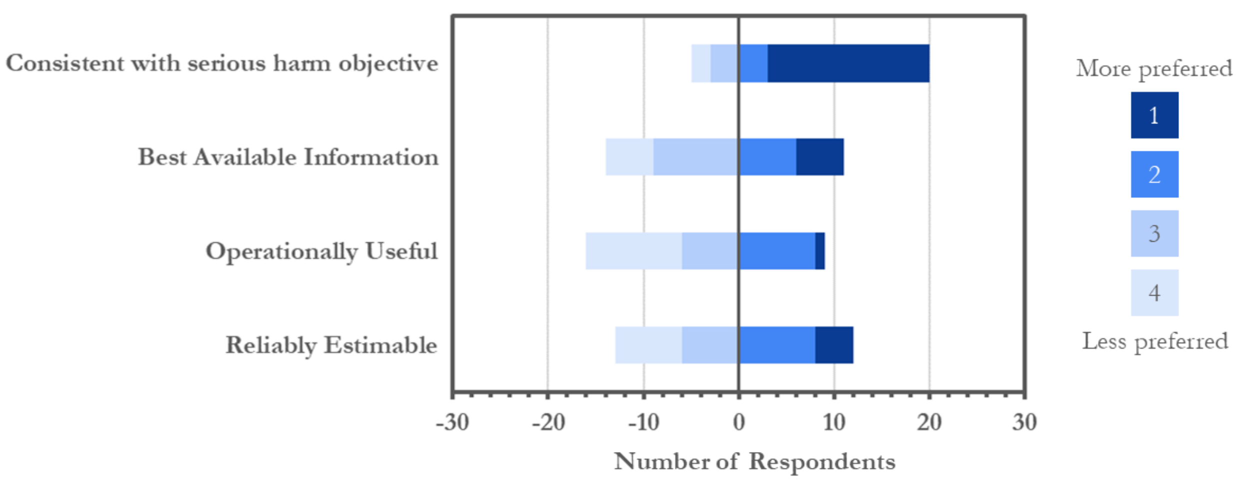

Figure 7: Diverging stacked bar chart showing the number of respondents ranking the importance of different candidate "best practice" criteria for LRPs and indicators. Responses have been standardized; positive responses are graphed to the right of the 0 on the X axis, in opposition to negative responses.

Question 19. Are any of these four proposed criteria listed above not essential (i.e., should not be part of a set of minimum "best practice" criteria for LRPs and indicators to meet)? If yes, please explain.

Eighteen respondents indicated that all four criteria were important, with several noting that the criteria were inter-related and that it was difficult to rank them.

Three respondents made recommendations or provided additional considerations. One participant suggested that “consistency with an objective to avoid serious harm” has not been consistently applied nor interpreted by managers. Another participant indicated that “operationally useful” may in fact be self-evident and not helpful when selecting an LRP, while a third clarified that “operationally useful” should be defined at minimum as an LRP and indicator that can function as a trigger for a rebuilding plan, if not an HCR.

Question 20. Can you suggest any other desirable 'minimum' best practice criteria that LRPs and stock status indicators should meet?

Several respondents suggested that the set of minimum criteria should be enhanced to mention:

- Less corruptible (less sensitive to changes in data available, such as fishery-dependent data or survey indices, or in situations with, for example, dome-shaped selectivity and cryptic biomass)
- Evidence-based
- Defensible
- Repeatable
- Enforceable
- Consistency with PA Policy
- Consistency with depensatory or Allee effects
- Peer-reviewed
- Not based on assumptions of stationarity
- Supported with timely data collection
- Refined as more knowledge is gained

Question 21. Would you like to share anything about your own experiences with setting LRPs? What challenges did you face? Please include the type of stock (e.g., large pelagic, etc.).

Respondents identified the following challenges based on their own experiences:

Stock Type	Challenges
Groundfish	• How to set LRPs when natural mortality due to predation is high • Whether to use a dynamic LRP • How to select LRPs and estimate stock status in MSE • Difference between LRP as a concept (e.g., Brecover, 40% Bmsy) and particular estimates of a LRP (in tonnes) that may change over time, e.g., in relation to MSEs and/or meeting Department reporting requirements • Mis-alignment between assessment and management units • Challenging stock-recruitment relationships (e.g., appearing linear, no evidence of density dependence in available time series) • Ability to estimate Bmsy

	• Estimating Bmsy from surplus production models in more data-limited contexts
Pacific Salmon	• Disconnect in scale when LRPs are set at level of salmon management unit (SMU) but conservation units (CUs) to be preserved under Wild Salmon Policy • Controlling harvest rates is at level of SMU, but additional management measures beyond controlling harvest (habitat and hatchery) that operate at the level of the CU • It is unclear what a potential SMU status metric (with LRPs set to ensure 100% of CUs are above "red status") will be for rebuilding plans at the SMU level
Invertebrates	• Data poverty • Reliance on fishery-dependent data • For broadcast spawners, there is a disconnect between stock size and recruitment in the same location due to dispersal during pelagic larval period – traditional stock-recruitment relationships are not well-suited to these species
Small pelagics	• Using LRPs for one stock based on other stocks may not be appropriate, due to different productivities among stocks

Some respondents provided more general comments on challenges experienced in setting LRPs:

- Setting LRPs is less likely to be challenged when stocks are not depleted
- Setting LRPs is impacted by trade-offs between short-term economic objectives and long-term sustainability objectives

Question 22. What challenges would the "one stock, one LRP" requirement of the Fish Stocks provisions provide for you, and how might that challenge be resolved? Please include the type of stock (e.g., large pelagic, etc.).

Respondents suggested the following challenges posed by "one stock, one LRP" requirements:

- **MSE** (due to multiple operating models in contrast to traditional stock assessment paradigms)
- **Stock complexes** (salmon, redfish)
- **Stocks under co-management**: partners may not agree with LRP, scale of LRP, or choice of management measures; a solution could be a co-management agreement among member nations around decisions such as choice of management measures
- **Data poverty** (in general)
- **Management units that cover only part of biological units and/or have different data sources and methods for each** (Georges Bank haddock, shellfish)
- **Multiple assessment or biological units in one management unit**
    - For invertebrates – LRP at bed level or stock level?
    - **Resolution of mismatch between Salmon SMUs and CUs**: LRPs define on basis of one or more CUs falling into "red zone" of WSP

- **Sessile invertebrates** where life history parameters and environmental variables vary over fine spatial scales and which may be naturally more abundant in some areas than others; understanding of what a stock is may depend on "best available information"
- **Coast-wide stocks where management and assessment are very limited in scale** (swimming scallops, green sea urchins, some clams)
- **Incorporating ecosystem approaches (EAFM) into reference points** (i.e., non-stationarity versus intermittently static estimates of LRPs, the need for ecosystem-level LRPs)
- **Stocks depleted because of poor prevailing productivity conditions**
- **Identifying "stocks" and their scale/structure**
- **Definition of "stock"** (i.e., it should be biologically defined given that LRPs are supposed to be biologically based on concept of serious harm)

Recommended References

Several respondents suggested references for selection of indicators and LRPs:

- Forrest, R.E., Rutherford, K.L, Lacko, L., Kronlund, A.R., Starr, P.J., and McClelland, E.K. 2015. Assessment of Pacific Cod (Gadus macrocephalus) for Hecate Strait (5CD) and Queen Charlotte Sound (5AB) in 2013. DFO Can. Sci. Advis. Sec. Res. Doc. 2015/052. xii + 197 p.
- Forrest, R. E., Holt, K. R., & Kronlund, A. R. (2018). Performance of alternative harvest control rules for two Pacific groundfish stocks with uncertain natural mortality: bias, robustness and trade-offs. *Fisheries Research*, *206*, 259-286.
- Kronlund, A.R., Forrest, R.E., Cleary, J.S., and Grinnell, M.H. 2017. The Selection and Role of Limit Reference Points for Pacific Herring (Clupea pallasii) in British Columbia, Canada. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/009. ix +125 p.
- Mace, P., & Sissenwine, M. (1993). How much spawning per recruit is necessary. *Risk evaluation and biological reference points for fisheries management*, *120*.
- Peacock, S.J., and Holt, C.A. 2010. A review of metrics of distribution with application to Conservation Units under Canada's Wild Salmon Policy. Can. Tech. Rep. Fish. Aquat. Sci. 2888: xii + 36 p.
- Peacock S. J. and Holt, C. A. 2012. Metrics and sampling designs for detecting trends in the distribution of spawning Pacific salmon (*Oncorhynchus* spp.). *Canadian Journal of Fisheries and Aquatic Sciences*. 69(4): 681-694.