Fisheries and Oceans
Science

Pêches et Océans
Sciences

C S A S

Canadian Science Advisory Secretariat

S C C S

Secrétariat canadien de consultation scientifique

Proceedings Series 2001/08

Série des comptes rendus 2001/08

Proceedings of the Fisheries Management Studies Working Group

8-11 January 2001

**Bedford Institute of Oceanography
Dartmouth, Nova Scotia**

**R. G. Halliday and R. Mohn, Chairmen
Marine Fish Division
Bedford Institute of Oceanography
Fisheries and Oceans
P.O. Box 1006, Dartmouth
Nova Scotia, Canada
B2Y 4A2**

March 2001

Proceedings of the Fisheries Management Studies Working Group

8-11 January 2001

Bedford Institute of Oceanography
Dartmouth, Nova Scotia

R. G. Halliday and R. Mohn, Chairmen
Marine Fish Division
Bedford Institute of Oceanography
Fisheries and Oceans
P.O. Box 1006, Dartmouth
Nova Scotia, Canada
B2Y 4A2

March 2001

Foreword

These Proceedings are a record of the Fisheries Management Studies Working Group meeting which was held 8-11 January 2001. The report records as faithfully as possible the contributions and discussion that transpired at the meeting. However, the individual interpretations and opinions expressed at the meeting are not necessarily or in all cases scientifically sustainable or supported by other participants. The discussion summaries document the deliberations, which led to the tabled proposals. No statements are to be taken as reflecting the consensus of the meeting unless they are clearly identified as such. Moreover, additional information and further review may result in a change of decision where tentative agreement has been reached.

Avant-propos

Le présent compte rendu relate les travaux de la réunion du Groupe de travail et d'étude sur la gestion des pêches tenue du 8 au 11 janvier 2001. Il reflète aussi fidèlement que possible les contributions et discussions des participants à la réunion. Toutefois, les opinions et interprétations individuelles qui y sont présentées ne sont pas nécessairement ou systématiquement soutenables sur le plan scientifique ou appuyées par les autres participants. Le résumé des discussions documente les délibérations ayant abouti aux propositions déposées. Aucune déclaration ne doit être considérée comme une expression du consensus des participants, sauf s’il est clairement indiqué qu’elle l’est effectivement. En outre, des renseignements supplémentaires et un plus ample examen peuvent avoir pour effet de modifier une décision qui avait fait l'objet d'un accord préliminaire.

TABLE OF CONTENTS

ABSTRACT

The Working Group met on 8-11 January 2001 to discuss the use of the Traffic Light method for application of the Precautionary Approach to fishery management planning. Specifically, the intentions were:

- to learn from the experience of a pilot application of the Traffic Light method at the November 2000 groundfish RAP and assess client reactions,
- evaluate possible solutions to outstanding issues and assess Traffic Light performance through examination of test cases,
- establish a plan of action for production of a citable version of the Methods Workbook produced by an ad hoc study group on the Traffic Light method in July 2000, and
- make recommendations for further research and on future applications of the method.

These aspirations were met.

RÉSUMÉ

Le Groupe de travail s'est réuni du 8 au 11 janvier 2001 pour discuter de l'utilisation de la méthode des feux de circulation dans l'application de l'approche de précaution à la planification de la gestion des pêches. La réunion visait les buts suivants :

- tirer des leçons de l'application expérimentale de la méthode des feux de circulation au PCR sur le poisson de fond de novembre 2000 et évaluer les réactions des clients;
- évaluer les solutions possibles aux questions restant en suspens et évaluer les résultats de la méthode des feux de circulation par l'examen d'essais-types;
- établir un plan d'action pour la production d'une version-source du guide de méthodes produit par un groupe d'étude spécial sur la méthode des feux de circulation en juillet 2000;
- formuler des recommandations d'autres recherches et d'applications futures de la méthode.

Ces buts ont été atteints.

FISHERIES MANAGEMENT STUDIES WORKING GROUP

REPORT OF MEETING of 8-11 January 2001

The Fisheries Management Studies Working Group (FMSWG) met on 8-11 January 2001 in the 6th Fl. Polaris Boardroom, Bedford Institute of Oceanography, Dartmouth, to address the agenda in Annex 1. Participants were as listed in Annex 2.

1. Introduction

History of Events

Halliday presented the following perspective: The WG developed a discussion paper on the Precautionary Approach in 1999 (published as Annex 3 to CSAS Proc. Ser. 2000/02). Dissatisfaction with current stock assessment methods, in particular their failure to incorporate the full scope of biological knowledge on populations and the effects of the environment on resource productivity, was expressed by a number of participants in these discussions. Also, there was no consensus in the WG that the direction being taken by ICES and the NAFO Scientific Council on definition of precautionary reference points had an adequate scientific foundation. Meanwhile, FRCC discussion papers on reopening closed fisheries, the development of methods for defining precautionary reference points for data-poor stocks (Caddy, 1999,NAFO Sci. Coun. Studies, 32: 55-68), and application of the Traffic Light method in shrimp assessments, attracted the WG's attention. The Traffic Light method was seen as an alternative stock assessment methodology that potentially could accommodate all data relevant to stock status, for stocks both rich and poor in data, while also providing a basis for defining precautionary reference points.

Initial trials of the method for groundfish were conducted by the Marine Fish Division (MFD), Science Branch, at the November 1999 RAP (CSAS Proc. Ser. 99/39). Summary tables of indicator results were included in Stock Status Reports (SSRs). This did not amount to a full application, but the WG found the results encouraging (CSAS Proc. Ser. 2000/02). An *ad hoc* group centred in the Bedford Institute of Oceanography took it upon itself to produce a draft Methods Workbook, which it completed in July 2000, and organised a short workshop to get more views on the Workbook proposals in September 2000 in Moncton. The minutes of that *ad hoc* meeting are attached as Annex 3. Most recently, MFD decided to do a full scale pilot of the method at the November 2000 groundfish RAP (CSAS Proc. Ser. 2000/33).

Pilot applications of the method to date have been in a purely stock assessment context. However, the function of FMSWG is to provide technical advice on all aspects of fishery management, i.e. on economic, statistical, compliance and regulatory issues as well as science issues. Thus the topic of the present meeting was set as "Use of the Traffic Light Method for Application of the Precautionary Approach to Fishery Management Planning", not "to Stock Assessment".

The Departmental Objective-Based Fisheries Management initiative (OBFM) is relevant to present discussions. Through this initiative, the Fisheries Management side intends to implement revisions to the Integrated Fisheries Management Planning (IFMP) process that emphasise the setting of quantifiable objectives, monitoring of performance in relation to these objectives, post-season review and risk management. The Traffic Light methodology could potentially serve as the technical basis for this initiative and this option needs to be evaluated.

Current Status of the Methodology

Mohn reviewed the rapid evolution of the Traffic Light methodology applied to groundfish assessment and its current status, as the starting point for present discussions. A glossary of terms being used and an illustrative Traffic Light table are available in the report of the September 2000 meeting (Annex 3).

Purposes of the Meeting

It was agreed that the general objectives of the present meeting would be:
- Learn from experience of the pilot application of the Traffic Light method at November 2000 groundfish RAP and assess client reactions
- Evaluate possible solutions to outstanding issues and assess Traffic Light performance through examination of test cases
- Establish a plan of action for production of a citable version of the Methods Workbook
- Make recommendations for further research and on future applications

Agenda and Schedule

The agenda was approved as in Annex 1. The meeting initially used case study applications (agenda item 8) of the Traffic Light method to address the technical issues in items 5-6 (boundary determination and integration of indicators), and some aspects of 4 (indicators), in aggregate. It then addressed the remaining aspects of 4, the questions of how to construct the interface between model outputs and fishery management decision making (item 7), and finished with a discussion of next steps in method development and application (item 9). The report follows the chronology of the discussions.

2. Traffic Light Case Studies

Case studies for winter skate, cod in 4X, haddock in 4X and sandlance were used to explore in detail the methodological issues associated with the Traffic Light method. It was noted that the Traffic Light method was being considered in a context where the precautionary approach was an essential consideration. The results are summarised, on the basis of issues.

<u>Integration of Indicators</u>

Integration of indicators involves two aspects, scaling the indicators to make them comparable and applying an operation to summarise the results from many indicators. The scaling of indicators to a common currency is mandatory before summarising. It was suggested that, where possible, scaling should be done with the indicators in their continuous, not in their discrete, state. For example, if the indicators at hand are catch rate 'like' it is possible to pursue calibration in the continuous state using techniques like GLM. It is anticipated, however, that there will be cases where some indicators are not continuous, e.g. some fishermen's indicators, and not all of the summarizations that analysts wish to make are of 'like' indicators.

Three methods of integrating indicators were presented to the WG other than the method used in Traffic Light applications to date, which was called the 'Strict Traffic Light' scheme. In this strict scheme, two cut points are defined which break an indicator into three regions designated by colours. This is shown schematically in Figure 1. The conversion to colours is the mechanism for scaling that fulfils the need for indicators to be in a common currency before integration. Once all relevant indicators have been converted to colours they may be averaged. No operator other than averaging has yet been considered for summarising several indicators. The centre portion of this figure shows the converter, which is a series of three steps. As only three colours are used, the conversion loses a great deal of detail and has sharp transitions from one light to another. The three other methods of integration that were presented give more resolution. These were labelled the 'Fuzzy Logic', the 'Continuous Conversion', and the 'Ramp-Step Conversion', schemes. The principal proponents of these schemes provided the following summaries of their arguments. In the case of the fuzzy logic scheme, a simplified working model for managing by TAC was demonstrated.

The Fuzzy Logic scheme – **Fanning** and **Silvert** (see Figure 2 for schematic presentation): The Traffic Light method, as developed so far for application in fisheries management, requires three levels of decision-making. In the first instance, there is a need to determine the boundaries between red and yellow and yellow and green. The second level of decision-making determines how different indicators are integrated into characteristics or even a single result. The final level of decision-making is the determination of the appropriate fisheries management measures as a result of the stock status. Most of the development on the Traffic Light approach so far has been directed to the first and second levels of decision-making. In spite of this, there remains considerable controversy about appropriate techniques for both boundary determination and integration. The Traffic Light Method is an elegant way of presenting complex data in a graphical and easily understood form, but it has some limitations that we believe can be avoided by the use of fuzzy sets.

The Traffic Light method is actually based on set theory. Conversion of indicators to traffic lights is equivalent to assigning each indicator to one of three sets, based on its value. There are obvious advantages to doing this, along with some disadvantages. Among the advantages is the ability to deal with non-linear indicators, such as temperature. Temperatures close to the optimum can be represented by a green light, the two ranges further out can be yellow, and the extremes red. It is possible to do this with non-linear functions of course, but then one gets into difficult and messy mathematics. The advantages of the Traffic Light Method are even greater when one deals with variables that are not well understood and are harder to calibrate in the laboratory, such as upwelling indices.

On the other hand, the discontinuous nature of the Traffic Light method is a serious disadvantage. As the Methods Workbook and the report of the September workshop (Annex 3) show, setting the threshold values is difficult and can be very controversial, and there is some reluctance to convert continuous variables into discrete categories. This has the potential to become a very serious problem if, for example, a number of indicators fall just below the threshold value and the result is a yellow light, when changing an arbitrary parameter value from 60% to 58% might give the industry a green light. Another disadvantage discussed in the September report is the difficulty of representing uncertainty, i.e. distinguishing between a yellow light obtained by integrating a lot of yellow lights and one that represents equal numbers of red, green and yellow lights.

These disadvantages can be easily avoided by treating the traffic lights as fuzzy sets. This simply means that instead of having only one light on at a time, signifying that the relevant indicator falls into one of the three categories, it is possible for more than one light to be on at a time, showing that the indicator can be associated with more than one of the categories. If we imagine looking at the lights while an indicator increases from very poor to very good values, the lights in the standard Traffic Light method would change suddenly from red to yellow and then from yellow to green, while with fuzzy lights the original bright red would fade and the yellow light would brighten, and then the yellow light would fade while the green light appeared. Although we still have only three lights, the changes are much more continuous and provide a better representation of the condition of the system. With normal traffic lights it is impossible to tell whether yellow means just a bit better than red or just a bit worse than green, but with fuzzy lights the former situation would correspond to a mixture of yellow and red and the latter to a mixture of yellow and green, and a bright yellow light by itself would show that we were in the middle of the yellow region.

It should not be difficult to generalize the concepts used in the standard Traffic Light method to fuzzy sets. For example, the definition of the red-yellow boundary as a limit reference point could be changed by saying that the red component must be less than 50%, i.e. the red light must not be more than half illuminated.

The use of fuzzy sets also makes the representation of uncertainty easy. If all indicators are yellow, then the integrated value will also be yellow, just as in the standard Traffic Light method. If, however, the indicators are a mixture of red, yellow and green lights, then the integrated value will show all three lights illuminated. This shows clearly that there is conflict between the different indicators and that the result is not conclusive.

There are several procedures that can be used to obtain the best parameter values for the fuzzy approach. One is to set them on the basis of the statistical properties of the time series, as was suggested for the boundaries of the normal Traffic Light method. However, we have found that in such complex cases where a great deal of scientific expertise is available, that relying on the judgement of researchers and other knowledgeable individuals, a sort of Delphi approach, is very good and gives reproducible results. Another approach that has proved promising is to use neural networks or genetic algorithms to build a fuzzy decision support system. No matter what approach is used, if it does not engender confidence in the people who use the results, it is of little value.

The approach described here is very similar to that of fuzzy control theory, and many of the results and methods of fuzzy control theory can be used in this problem. This gives the methodology a degree of standardisation that may make it more palatable to managers. However, the mathematical formalism is not trivial, and the integrating procedure has to be flexible enough to reflect biological reality. Clearly some indicators, such as stock size, are more reliable than others, and should be weighted more heavily in determining the summary values.

The Continuous Conversion scheme – **Evans** (see Figure 3 for schematic presentation):
There is a widespread and strong feeling that the Traffic Light approach should be reluctant to relinquish the advantages of continuous variables, either in its computations or in its presentations. Note that the computation and the coloured presentation are separate issues. Each value, or distribution of plausible values, for an indicator can be thought of as a probability distribution. It is then easy to form a (possibly weighted) average of distributions for different indicators to obtain the distribution for a characteristic. This method automatically handles any combination of discrete and continuous distributions. It must be recognized, however, that averaging is not always the appropriate operation - especially when we seek the distribution of likely values for the true quantity, which will be narrower than the distribution of likely values of indicators of the quantity. (This, incidentally, is the basic reason why Bob Mohn's moving cut points were inappropriate [when setting boundaries for combined indicators, see below]. When you combine several estimators, you do indeed end up with more certainty (less colour range) in the combined estimate.) There will be times when two indicators blatantly contradict each other, in which case there is NO good mechanical rule for combining them.

Once we have a probability distribution for a characteristic, however it was obtained, we can present it as a continuous band of colour as follows. Construct the cumulative distribution function. Assign a colour from the spectrum to each point on the abscissa (ranging from red (rotten) through amber (apprehensive) and green (good) to blue (brilliant)). Assign the colour of the appropriate abscissa to each point on the cumulative distribution function, and project these colours onto the ordinate. The ordinate will then be a coloured bar going from 0 to 1, with a dominant colour coming from the region where the distribution is steepest. Rainbows instead of traffic lights.

The claim may be made that averaging probability distribution functions can be thought of in terms of fuzzy sets. But even if it can be, this is not in itself an argument that it is useful to do so. It often happens in mathematics that there are many different sets of primitive concepts for a subject, such that each set can be derived from every other set. So showing that set B can be derived from set A is in no way an argument that set A is more fundamental, let alone more useful, than set B. If the commonplace of averaging will give us what we want, why fuss with the arcana of fuzziness?

The Ramp-Step Conversion scheme – **Mohn** (see Figure 4 for schematic presentation): Ramp-step conversion is a hybrid of strict traffic lights and continuous analysis. The ramp-step conversion is like Strict Traffic Light analysis in that it has three defined states, but the transition between them is now continuous and is shown as ramps. This approach removes the sudden jumps of the strict method while being less complex than fuzzy logic. As with the fuzzy logic example, the two-colour (pie chart) lights could be rounded to the larger slice if only three lights are wanted at the final state. An example application of this approach is shown in Figure 5. The bar chart on top of this figure is for the abundance characteristic and contains two pairs of horizontal lines. The top pair marks the width of the ramp for the yellow-green boundary and the lower two are from the yellow red boundary. Also note that some traffic lights are split into two colours, which reflects their position on the ramp between the two adjacent steps (colours).

Conclusion: It was proposed that all four of these (strict traffic light, fuzzy logic, continuous and ramp-step conversion) methods be evaluated by an *ad hoc* group on behalf of FMSWG. Common data sets should be used for test cases. The objective would be to describe fully the issues involved in application of these methods and the extent to which the methodology changes the aggregated indices and hence the interpretation of the data.

Setting Boundaries for Indicators

In applications of the Traffic Light method so far, the boundaries for indicators have been determined on a series by series basis, and the default is to define them from averages based on history. These have the advantage of simplicity and serve to scale indicators so that they may be summarised. It is not immediately obvious, however, that a common rule applied to several indicators automatically scales them 'appropriately' in a precautionary approach sense. The probability mass attributed to each state, red, yellow or green depends on how the arbitrary rule deals with the statistical distribution of each indicator. This issue was reflected in the discussion of survey data where applying the rule to the raw observations gives different boundaries to applying the rule to log transformed data. Even if the statistical distributions of indicators were made similar through transformation, a 20% reduction in biomass, for example, may not mean the same degree of 'worsening' as a 20% reduction in condition factor. It would be preferable to have an exogenous way to identify boundaries that reflected the biology of the resource.

When boundaries are based on history, inclusion of fewer or more years could change the categorization of indicators quite a bit. For example, the 4X haddock stock was more productive in the 1950s and 1960s. However, this was mainly as a result of recruitment of an exceptionally large 1963 year class, and 1960s events could be considered anomalous. Limits or targets need to be realistic in the current context. Shrimp stocks in the Northwest Atlantic provide a contrasting example. The most recent period is the most productive and the current means may be a poor representation of medium to long term expectations. Barndoor skate indicators, on the other hand, should perhaps be red in recent decades as observations are all coming from a period of poor abundance. There is no guarantee that a long time series will cover the full dynamic range but such a series is more likely to have covered a wide range than is a shorter one. Expert knowledge is necessary to "design" boundaries (maybe using Delphi techniques) that take into account the full dynamic range of an indicator, even if such range is not revealed by the data.

Setting Boundaries for an Integrated Group of Indicators

In the Traffic Light program that was used for the recent RAP, boundaries were set for the composite indices (characteristics, summaries) that assigned an equal probability mass to each colour, taking into account the number of indicators that were combined.

The Rationale – **Mohn** explained as follows: Table 1 shows the results of 30000 random draws of two indicators that can be red (1), yellow (2) or green (3). In this example, about 11% of the time the answer was pure red (1) or pure green (3). If the combined decision points were set at 1.51 and 2.49, the random draws would be split into equal thirds.

Table 1. Combinations of 2 indicators from 30000 random draws

Colour	# of occurrences	Proportion
1.00	3339	0.111
1.50	6750	0.225
2.00	9968	0.332
2.50	6621	0.221
3.00	3322	0.111

However if 6 indictors were averaged and the decision points were still at 1.51 and 2.49 the results would be quite different. In Table 2 the results for 30000 random such draws shows that less than 11% would be green and less than 11% would be red and the remaining 78% would be yellow. In order to have equal probability of red, yellow or green from random inputs the decision points would have to be moved nearer the yellow (to approximately 1.8 and 2.2).

Table 2. Combinations of 6 indicators from 30000 random draws

Colour	# of occurrences	Proportion
1.00	36	0.001
1.16	253	0.008
1.33	912	0.030
1.50	2013	0.067
1.66	3708	0.124
1.83	5239	0.175
2.00	5796	0.193
2.16	5095	0.170
2.33	3783	0.126
2.50	2073	0.069
2.66	838	0.028
2.83	226	0.008
3.00	28	0.001

Comments: Several participants expressed concern about the moving of boundaries in response to an increased number of indicators. When you combine several estimators you do indeed end up with more certainty (less colour range) in the combined estimate and moving cut points was thought to be inappropriate. Boundaries for summary or characteristic indices will be the input to management decision rules, and thus decisions on boundary locations should likely not be a purely statistical issue. More work is required on this.

Short Time Series

The Indicators for the stocks reviewed included some from industry surveys that have only a few years of observations (about five years) as well as research vessel series extending for 30 years. In the November 2000 assessments, short time series were given low weight (0.1 or 10%) when integrating them into summary series or stock characteristics. It was suggested that weights are the wrong tool for dealing with series of different duration. At issue is the fact that it is less likely that short series have covered the full dynamic range of observations and that, in such a case, boundaries established from historical averages would most likely be meaningless. On the other hand, some were concerned that assigning the same colour to all points in the short series would equate to such series not adding any information to the integration process. Of course, if they are assigned only a 10% weighting, they do not influence the integration much anyway. It was pointed out that, if the uncertainty is in where the boundaries are, not in the series themselves, ways to keep such series while making them relevant to the overall integration need to be explored.

Techniques for integrating indicators in the continuous scale, like GLM mentioned above, can accommodate series of different duration. It was suggested that, particularly in the case of time series of different duration, scaling should be done with the indicators in their continuous, not in their discrete, state. The scaled indicators could use the scaled boundaries, thereby removing the need for defining separate boundaries for each indicator. It is conceivable that a short series may end up with points all of the same colour using that technique; this is a desirable property as short series are less likely to span the full dynamic range of a stock characteristic.

Some suggested that short representations of the indicators (using the same number of observations for averaging all series) be produced as a means to address the issues related with short series. Such an approach may not provide boundaries in the precautionary approach sense but may provide a suitable scaling and an improved graphical representation of the recent trends in indicators and characteristics.

Weighting

Weighting of indicators should be reserved for specific roles. For example, some indicators may be more important than others and this could be accounted for in the integration process by varying the weight given to each. Similarly, some indicators may be less precise than others and weights based on the inverse of their variance could be used to reflect this. Some weighting factors may best be determined at stages of the decision process that includes user groups.

Overall Summaries

The necessity of defining an Overall Summary (direct or indirect) for application of decision rules was questioned. It was suggested that stock characteristics were sufficient for that purpose and accounted for important "axes" of stock status. The main questions to address during an assessment are:

- What is the size of the resource?
- How fast is it renewing?
- How are human actions affecting the stock characteristics?
- How might these things change (due to environment, ecological regime shifts, intrinsic variability e.g. in recruitment, etc.)?

We should answer these questions separately. Attempting to address such questions related to conservation through a single 'roll-up' indicator does not help clarify the conservation issues.

It was noted that the stock characteristics used to date, i.e. abundance, productivity and fishing mortality cover the first three of these questions. Abundance and fishing mortality are similar to the biomass and exploitation axes used in the precautionary approach frameworks used by ICES and the NAFO Scientific Council. These are the elements that support decision rules based on the advice of these organizations.

It is important to keep the number of characteristics manageable for the decision making process. Abundance (biomass), production and fishing mortality are a good starting point.

With respect to the 'production' characteristic, it was suggested that we need to construct a model linking recruitment and growth to production. If a model is available, it should be used to obtain the characteristic and its boundaries. An arbitrary integration of recruitment and growth indicators should be used only if a model is not feasible. Such a model might also be useful for prediction.

Relevance of Indicators

The biological mechanisms linking a given indicator to stock status are unclear for many indicators. Nevertheless, the utilization of such indicators deserves to be considered. Indicators of longevity were raised as examples, e.g. how many fish are within 80% of $L_{infinity}$ or how many are within 80% of the oldest age? It is hoped that it will be possible to bring in these types of indicators by using the Traffic Light method. Some expressed concern that introducing these indicators without interpretation in relation to stock status would not help the decision making process and could hinder it if interpreted in disparate ways. For example, reduced growth may be seen as either positive because it signals increased competition for food due to increased abundance or negative because it indicates deteriorating environmental conditions. However, ancillary information may be available, in this case on abundance and environmental conditions, to distinguish which interpretation is appropriate.

Correlated Indicators

The integration of correlated indicators was discussed. For instance, indicators originating from a model that has used a given index should not be added again in the summaries containing such index. It was argued that there is an advantage of having redundant information, as the raw index could show different trends or variability in more recent years. Others argued that, if output from a model such as VPA is acceptable (e.g. no retrospective pattern, relatively good precision), it should be used as the abundance characteristic but if problems with the VPA are impossible to resolve, it is preferable to use the index as the indicator. Concern was also expressed that not using a VPA when one is available results in the cohort information in the data being lost. Cohort effects are not easy to obtain from other Traffic Light indicators. On the other hand, the Traffic Light method allows the inclusion of indicators that are not typically considered in age-structured calibration models. In summary, an attempt should be made to reduce the amount of redundant information or correlated indicators so that those integrated in the Traffic Light table are independent. In the end, however, the abundance characteristic should be the best aggregate of the abundance information. This has to be judged on a case by case basis.

Biomass 'Quality' Indicators

It was recognized that abundance has several dimensions (biomass, distribution, size and age structure). We need to measure the 'quality' of biomass as well as the biomass level itself. Scientists recognized that the distribution of cod was changing before they measured the stock collapses based on biomass measures. New indicators based on distribution in relation to depth, for instance, may assist in identifying significant changes in abundance. Can we identify average distribution patterns and relate yearly changes to them? Are there distribution patterns that are representative of low, medium and high abundance? One alternative is to look at maps, but these may not fit in the Traffic Light framework. The definition of indicators of distribution should be pursued, and indicators of age (size) structure need to be developed.

These new measures of the distributional dimension of biomass are important but the interpretation of the measures we have is unclear. Area Occupied (% of tows where target species was seen) has received the most extensive use but was not considered to be a truly independent measure of abundance or its 'quality'. In most cases it is not telling us more than the biomass index. If we want to measure resource contraction, we have to measure this differently. Contiguous area occupied would be a better measure of habitat use. Another measure that has been used is Resource Concentration (the minimum proportion of sets to reach x% of survey biomass estimate). Questions were raised as to interpretation - is concentrated abundance good, or is it bad? Is concentration during the short time period of the research vessel survey an adequate indicator of inter-annual changes?

Ecosystem Indicators (with special reference to Prey Species)

Prey species represent special problems in resource assessment. Firstly, they are usually less well sampled by research surveys and rarely are seen in commercial sampling. They are also less likely to be aged or to have detailed biological sampling. The second source of difficulty is that they may be more important to fishery yields of other species through ecosystem linkages than through directly supporting a fishery. For example, concern is raised about the Antarctic krill resource not because it has direct economic significance but because it may affect the status of whale stocks. Thus, a traffic light synopsis for krill might well include a whale indicator.

The sandlance case study focussed attention on a data poor species as well as the concept of an ecosystem characteristic, the appropriateness of particular ecosystem indicators, and the role of these in individual species stock assessments. Can a cod biomass indicator be used as an indicator of sandlance abundance? Could cod condition factor be taken as an indicator of sandlance mortality? Is predation mortality rate on sandlance what we should be looking at?

Need to Show Indicators

It was noted in the sandlance case that the number of indicators and characteristics were not much different, making the indicator information largely redundant. This raised the general question of the need for the indicator part of the Traffic Light matrix. It was argued that there is value in retaining the transparency provided by showing all the inputs.

Need for Lights

Some suggested that it might not be necessary to show the lights for the indicators. For instance, this would have the advantage of removing the need to scale the indicators when integration is done in their continuous state. Others pointed out that it was important to show the indicators in the Traffic Light fashion for greater transparency. The hope is for a method that produces a snapshot of stock status using a consistent format for all information.

Key Conclusions from Case Studies

1. Important to document the basis for indicators and their boundaries.
2. Scaling the indicators to a common currency is mandatory before integration.
3. Boundaries of characteristics should attempt to capture stock dynamics features and be set so as to be meaningful in the precautionary approach context (e.g. red means below acceptable limit). For instance, boundaries could be model based, so as to capture stock dynamics and biological characteristics.
4. Important to avoid redundancy when integrating indicators into characteristics.
5. Summaries are not essential to the use of the Traffic Light method in stock assessment. It was suggested that stock characteristics were sufficient to display all important "axes" of stock status. Decision rules can be formulated directly from the stock characteristics.
6. It is advisable to work with the continuous variables (indicators) as long as possible within the Traffic Light method. Analyses done with continuous variables can serve to integrate continuous indicators, thereby providing a means to deal with time series of different lengths and reducing the need to invoke arbitrary weighting of indicators. Make the output discrete only when analyses on continuous indicators have been completed.
7. Some techniques, e.g. fuzzy logic, models based on continuous variables, or models based on probability distribution functions of indicators, are available to address issues related with the abruptness of the boundaries. Current Traffic Light software does not implement fuzzy logic techniques or continuous techniques and more research/development is needed in this area.

3. Multi-Criteria Decision Making

Dr. Karasakal, Faculty of Administration, University of Ottawa provided a review (Annex 4) of the Traffic Light Workbook draft of July 2000 but unfortunately was unable to attend this meeting. Many of the points she raised were addressed during this session. An important outstanding issue is the need to consider the multi-criteria decision making literature. The suggestion that time series analysis could be considered is also of value. It is one of the tools available for linking Traffic Light output to decision rules or decision making through projections. It was agreed that Dr. Karasakal would be invited to attend the next meeting related to the development of the Traffic Light method.

4. The Broader Context

Three presentations were made on developments in ecosystem and socio-economic status assessment methods and comparisons made to the Traffic Light method for stock status assessment.

Ecosystem Assessment - **O'Boyle**: Canada, along with a number of other countries (e.g. USA and Australia), is endeavouring to implement ecosystem-based management, of which fisheries assessment would be a component. The Traffic Light method would fall under the Impacted Species aspect of Productivity that is part of meeting the conservation objective, which itself is one of several objectives. The Traffic Light method has common elements to the Index of Biotic Integrity approach that is currently being used to assess the state of freshwater ecosystems. Indeed, the concept of Marine Environmental Quality is based on a similar approach. Use of a common assessment approach will facilitate incorporation of single species stock assessment into the broader ecosystem framework. This is one of the topics of a DFO workshop on ecosystem objectives, indicators and reference points planned for the end of February 2001. The discussion on so-called data-based assessment methods, e.g. the Traffic Light method, versus model-based assessment methods, is topical internationally.

Index of Biotic Integrity (IBI)– **Halliday**: Two examples of IBI implementation chosen from the literature show the similarities to the Traffic Light method. In particular, indicators were classed in three states, each of which had a numerical value (1,3 or 5). The results were integrated by summing to provide overall scores. The summary indicator was then classed into five states. Features of particular interest were that the definitions of classes of the summary indicator were discontinuous and scores falling between classes were assigned to adjacent classes based on the analyst's judgement, and the lack of connection between individual and summary indicator boundaries.

Genuine Progress Indicator (GPI) **– Liew**: This socio-economic index also integrates multiple indicators. The currently used Gross Domestic Product (GDP) has limitations as a measure of economic health, as it considers only monetary transactions. The GPI includes a wide range of indicators, covering both economic and societal processes. Whereas colour is used to scale indicators in the Traffic Light method, the dollar is used to scale the indicators in the GPI.

Discussion: With regard to the IBI, it was commented that the choice of the class scores as 1,3 or 5 implies classes 2 and 4 and is thus a usage of a fuzzy logic approach without labelling it as such. Some felt that, with the data sets illustrated, a fuzzy logic approach would not make a difference in the final result. In discussion of the GPI, it was pointed out that considerable work at the indicator level is implied to ensure appropriate measurement and scaling. Valuing social processes is not easy. However, after proper scaling, the summation of the indicators is relatively straight forward. There are clear similarities among Traffic Light, IBI and GPI concepts, although various important differences in application. It is evident that these ecosystem and socio-economic initiatives will proceed and the issue of fitting stock assessment into this broader context will be faced.

5. Indicators of the Status of Attributes

The importance of developing comprehensive and authoritative indicator accounts as the basic building blocks of the method was reinforced by the case studies and in other discussions. It was agreed that the first draft template, that served as the basis for accounts produced to date, should be enhanced to include greater emphasis on validation of the relationship of the indicators to stock status and on the establishment of biologically meaningful rather than arbitrary boundaries. A section should be added also on calculation options and the circumstances under which these are appropriate. The description should relate the attribute (of which the indicator provides a measure) to the characteristic or characteristics it provides information about and under what circumstances. Fanning agreed to provide a second draft of the template (see Annex 5).

It is the intention that these indicator descriptions provide a benchmark for building Traffic Light tables and thus the descriptions should be peer reviewed. They should evolve by incorporating new insights. There is no suggestion that they should limit the originality of analysts to develop new indicators or improve established ones. They will, however, require the analyst to establish the basis for an indicator according to the guidelines.

It was proposed that descriptions of characteristics be developed also. This would aid in deciding what attributes should be associated with them. More importantly, it would establish the basis for interpreting characteristics in the context of decision rules.

Stock Status Indicators

The July 2000 draft Methods Workbook has, as appendices, a number of indicator accounts. However, these have not been peer-reviewed and do not include all the indicators used to date in stock assessments.

Biomass 'quality' indicators, such as Geographic Distribution (draft already in Workbook) and Size/Age Distribution are discussed above under Case Studies. Several other indicators were addressed during general discussion. First drafts of Fishery Distribution (Branton) and Relative F (Fowler) indicators were tabled, as was an addendum on calculation methods (Fowler) to the previously produced Condition indicator. Production of a Size/Age at Sexual Maturity indicator was offered based on silver hake as an example (Showell). A verbal presentation of a Catch Curve Z indicator (Mohn) illustrated validation methods. There was insufficient time during the meeting to address these indicators and they were deferred for later consideration (see below under Next Steps).

Fishery Economic Status Indicators

Economic/social indicators are relevant at the fishery (IFMP) level, not the stock level, at least in multi-stock fisheries. Even then, there are interactions between fisheries due to multiple licensing, e.g. groundfish – lobster, and it may only be meaningful to track these indicators at the fleet level. Decisions are required on what is useful. It would be preferable for these indicators to be developed in the context of introducing OBFM features, such as establishment of quantifiable objectives, to IFMP planning. However, in the absence of clear policy objectives, some plausible objectives could be assumed as a basis for developing indicators. More consideration is required and this WG could play a supporting role.

Ideally, indicators could include amount of revenue, number of licences, number of crew, earnings per crew, earnings versus those in other industries, and indicators of revenue concentration/distribution. Fish price information is now difficult to obtain and cost/earnings surveys have not gone well in recent years. It was noted that conditions of licence could now be tracked on an individual vessel basis, which could improve the quality of some participation data.

Reference was made also to historical studies, both by DFO and outside sources, which could be useful.

Regulatory Compliance Indicators

No progress has been made in developing indicators of compliance. Developing Fishermen's Indicators of compliance (see below) was thought to have some potential. So too was the idea of structured surveys of Fishery Officers. Other options identified were the use of the Observer Program and focussing interviews on recently retired fishermen.

It was concluded that the FMSWG forum was not attracting the appropriate people to advance this discussion. More work needs to be done in more precisely identifying the questions that need to be addressed. As discussed above for economic indicators, compliance indicators will be required when introducing OBFM features, such as risk analysis, to IFMP planning. Again, this WG could play a supporting role.

Fishermen's Indicators

Fishermen's observations provide a source of indicators of stock status, fishery conditions and regulatory compliance largely independent from DFO data. Development of indicator series through some sort of survey was viewed as the only way to get consistently collected multi-year data. The imperative of satisfactorily incorporating fishermen's views on stock status into assessments was recognised. The Traffic Light method facilitates introducing fishermen's indicators directly into the calculation of stock status.

An end-of-season survey of groundfish harvesters from the southern Gulf of St. Lawrence conducted since 1996 by telephone was reviewed (Hurlbut). Some of the results have been considered during stock assessments. In the last couple of years the FRCC has conducted a mail-in questionnaire for fishermen engaged in eastern and western Scotian Shelf and Georges Bank fisheries. This is used for their guidance in making recommendations on harvesting. DFO Science in Newfoundland has conducted similar surveys, e.g. for capelin, and information on their performance needs to be obtained. Performance of these surveys should be fully evaluated as a basis for further efforts to develop fishermen's indicators that can be used in Traffic Light analyses.

There is an increasing literature on fishermen's traditional knowledge that could provide insights on survey design and reliability of results. It is important to understand what observations by fishermen form the basis of their perceptions so that questions can be phrased appropriately and their answers interpreted correctly.

6. Form of Scientific Advice/Interface with Clients

Reactions to November 2000 Groundfish RAP Traffic Light Output

At the groundfish RAP 2000 a formal paragraph structure was adopted in SSRs to communicate an 'Outlook' from Traffic Light tables. The first four paragraphs summarised the results for the three population characteristics of abundance, production and fishing mortality, and for status overall (overall summary), in the contexts of immediate and longer-term prospects. Subsequent paragraphs addressed specific issues such as data deficiencies, bycatch or other regulatory problems. A Traffic Light table was included that showed a 'summary' line but 'characteristic' lines were omitted for simplification. No information was provided in the SSR on boundaries or weighting of indicators in order to simplify the presentation further. All indicators shown as series of lights in the Traffic Light table were supported in the SSR with graphs illustrating the underlying data in continuous form.

No reactions to the advice are yet available from the FRCC or the fishing industry. There was no discussion of the methodology during the course of the FRCC's planning meetings and no science presentations were required at FRCC public consultations. Departmental regional Groundfish Advisory Committee meetings are upcoming and these may provide some industry reactions. Fisheries Management Branch attendees at the present meeting thought the Traffic Light table summarising all the indicators was helpful and that it should be easily interpretable by a novice. They had not yet seen the final versions of the SSRs but indicated the need to be able to translate the advice into a bottom line with regard to TACs, i.e. the need for decision rules (see below).

Form of Scientific Advice

Application of the Traffic Light method to data-poor stocks does not introduce new limitations on the precision of advice. Indeed, it requires that the advice be formulated in a rigorous manner that gives explicit weights to all of the available information. Current status in relation to that of earlier years is presented in a quantitative form in the Traffic Light table and this should be an improvement over the subjective evaluations of the past.

The particular formulation used for the advice in November 2000 may, nonetheless, give less precise guidance than that offered previously for data-poor stocks. Clients are familiar with the previous form of advice and have developed their implicit decision rules on the basis of their understanding of it. While the advice given in November 2000 may reflect more accurately the limitations of the scientific information available, managers have to have the necessary guidance to make appropriate practical decisions on the basis of it.

In cases where SPAs have been accepted as the basis for advice, the advice has included projections of the consequences of allowing various levels of catch on F and stock size. The question was raised whether, in these cases, the form of advice would be the same in future. Reconciliation between SPA and Traffic Light analyses was discussed under Case Studies (see above), but the issue remains unresolved. There is a widespread perception that, in providing scientific advice, ways need to be found to take into account all sources of information on stock status. Combinations of the two methods may be possible. The question of the use of projections, even if SPA results are accepted within a Traffic Light table, also remains at issue. It was recognized that some quantification of advice is necessary, although not necessarily in the same form as in the past.

The issue of the use of continuous versus discrete scales, discussed previously under the topic of data integration, was raised again in relation to presentation of Traffic Light results. The view was expressed that if it were possible to present results in a continuous state, this would be more informative than only three discrete states. It was pointed out that current practice was to do both in that each indicator is shown in the SSR in its continuous state as well as in its discrete state in the Traffic Light table. However, the merit of presenting the data both ways was questioned. This matter was unresolved.

Decision Rules

It was accepted that the Traffic Light methodology would be incomplete until the interface with clients was established through sets of decision rules. The decision rules for each application need to be worked out amongst all parties (scientists, managers, fishermen's representatives and FRCC in case of groundfish). It was the WG view that decision rules should be rigid only in specifying that a particular result triggers a type of action, but should leave full flexibility in decisions on measures adopted to implement the action. A formalistic decision rule where a particular result triggers a prescriptive response was to be avoided.

The new OBFM initiative requires formulation of decision rules as part of the IFMP process. The WG was informed that the Scotia-Fundy Region had put forward groundfish and inshore scallop fisheries for pilot applications of OBFM. In the case of groundfish, the DG has approved a two year Regional IFMP extending to March 2002, in anticipation that a pilot application of the OBFM initiative will provide a longer term plan for succeeding years. Thus the groundfish fishery plan development presents an obvious choice for a case study for exploring and determining decision rules that link Traffic Light results to managerial actions.

It was decided that the WG should conduct some case studies that addressed the issues of interfacing the Traffic Light output with the rest of the management system. Groundfish stocks provide well-worked examples and results would complement concurrent Regional plan development. As well as providing advice on the type/direction of action required, ways of quantifying the consequences of different strengths of the action are also required. The relationship of Traffic Light and SPA results also needs to be resolved. It is necessary to have FM and FRCC involved in these case studies. In this regard, the FRCC initiative to produce long-term fisheries resource conservation plans makes it essential to the success of groundfish management planning that the current Science, FM and FRCC initiatives come together to form a logical whole. The proposed case study discussions could provide a mechanism for resolving the technical issues related to this.

The Precautionary Approach

It is necessary, in IFMP implementation, to meet the requirements of the precautionary approach. The FMSWG discussion paper gave a set of requirements for operationalizing the precautionary approach as: set quantifiable objectives and strategies to meet them, establish indicators of performance in relation to objectives, define unacceptable outcomes, have pre-agreement on decision rules that trigger corrective actions when unacceptable conditions are approached, and take uncertainties into account.

All of these concepts are embodied in the OBFM initiative. Concern was expressed about the rigidity of decision rules relating to the precautionary approach in drafts of OBFM guidelines produced to date. It was explained that the guidelines were drafted using a NAFO/ICES framework. However, the Traffic Light method could be used to the same purpose if the lights were referenced to the precautionary approach concepts. For example, the objective could be to keep a stock in green or fluctuating between green and yellow with decision rules becoming increasingly restrictive if there were sequences of yellow or occurrences of red.

7. Next Steps

Indicator Descriptions

Ways to have indicator descriptions drafted and reviewed were discussed. A good start has been made in engaging the interest of scientific experts to provide drafts and it was proposed to build on that base by encouraging those that have made commitments, to produce drafts that are in conformity with the revised guidelines. Peer review of the indicator descriptions was seen as a necessity. The review process is a stock assessment activity and as such could be considered as outside the mandate of the FMSWG. Initially, review requires some special arrangement to be made. Subsequently, evaluations of new indicators, and revisions to established ones, could become part of the annual assessment review process.

If the Traffic Light method comes to be widely used in other regions, co-ordination will be necessary to maintain consistency. Putting indicator descriptions on a Traffic Light web site is the obvious solution. There is a need to retain a history of indicator descriptions so that the versions referenced in particular assessment documents are retrievable at a later date. This can be accomplished using an electronic revision control system. A gatekeeper will be required to ensure that reviewed material only is posted.

The Draft July 2000 Workbook

The general description of the Traffic Light method and guidelines for its use provided by the workbook was considered to have some lasting value. It was suggested that its authors update it and publish it as a Research Document. It could subsequently be posted on the Traffic Light web site. The latter version could be updated periodically to incorporated new methodological developments. The view was also expressed that there is a need to inform the broader scientific community by publishing in the primary literature an account of the progress made in developing and evaluating this method.

Further Case Studies

Integration of Indicators: Four different ways of combining indicators were proposed during the discussions. It was agreed that these should be tested using common data sets to establish the advantages and limitations of each and how sensitive results are to the method used. No arrangements were made for this during the meeting.

Decision Rules: Several inter-related problems remain outstanding associated with the translation of Traffic Light results into management actions. Case studies need to be examined in a multi-disciplinary working group. No arrangements were made for this during the meeting.

Future Meetings

A proposal was made to hold the next meeting of the WG in the second half of June, 2001 contingent on there being sufficient business.

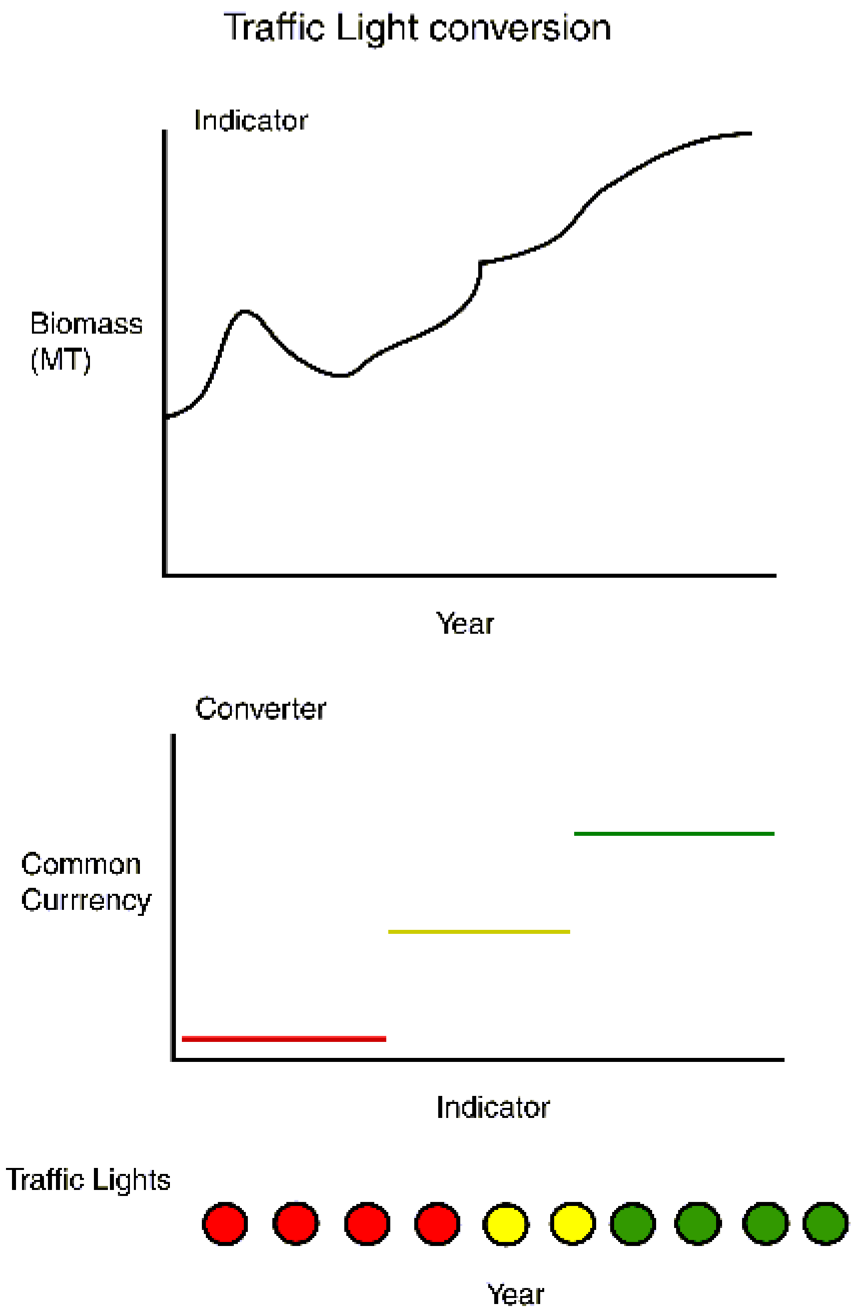


Figure 1. Schematic of strict traffic light conversion of an indicator (Biomass) into Traffic Lights. The top portion is the input indicator, the middle is the indicator converted to a common currency and the resultant lights are shown at the bottom.

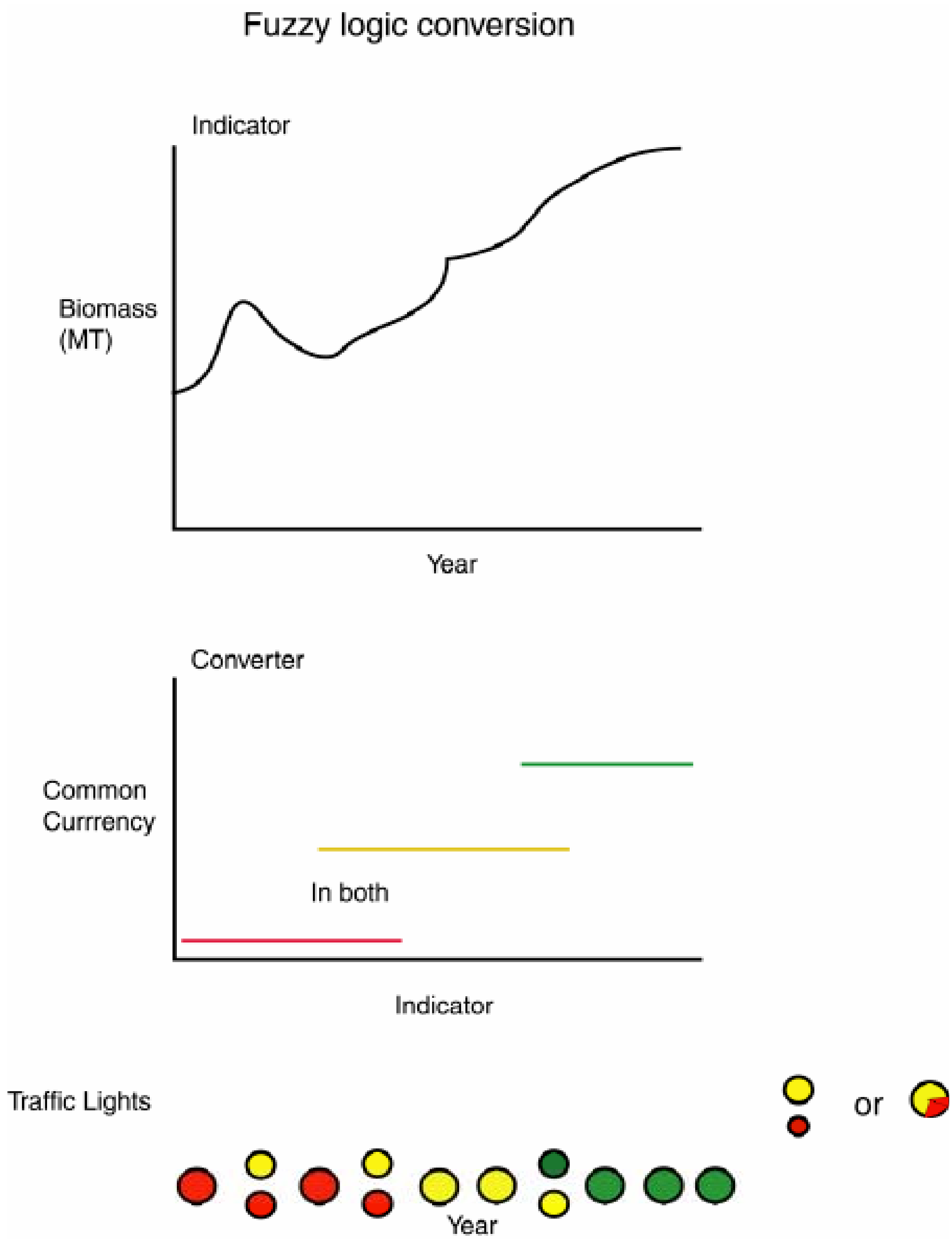


Figure 2. Schematic of fuzzy logic conversion of an indicator (Biomass) into Traffic Lights. The top portion is the input indicator, the middle is the converted to a common currency and the resultant lights are shown at the bottom.

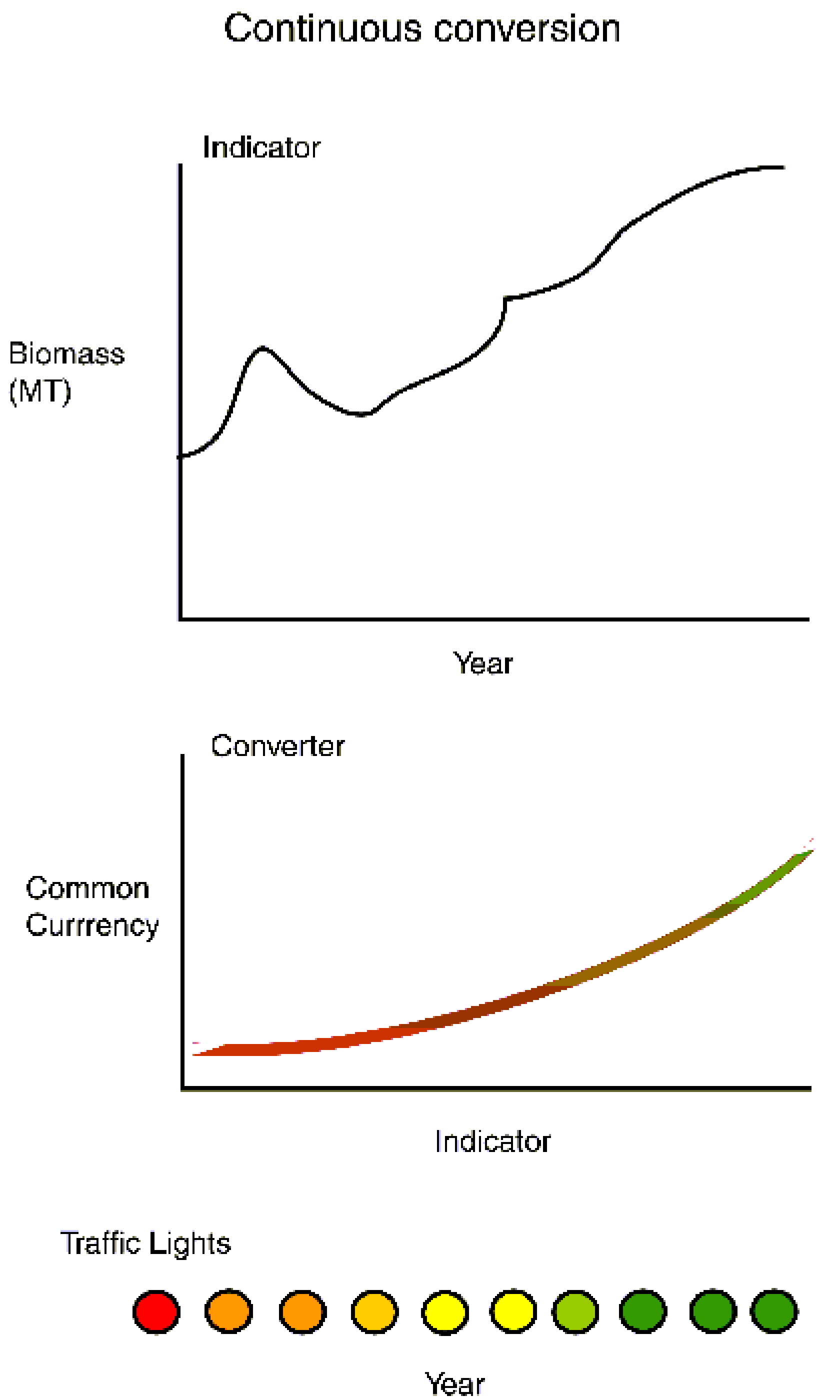


Figure 3. Schematic of continuous conversion of an indicator (Biomass) into Traffic Lights. The top portion is the input indicator, the middle is the converted to a common currency and the resultant lights are shown at the bottom.

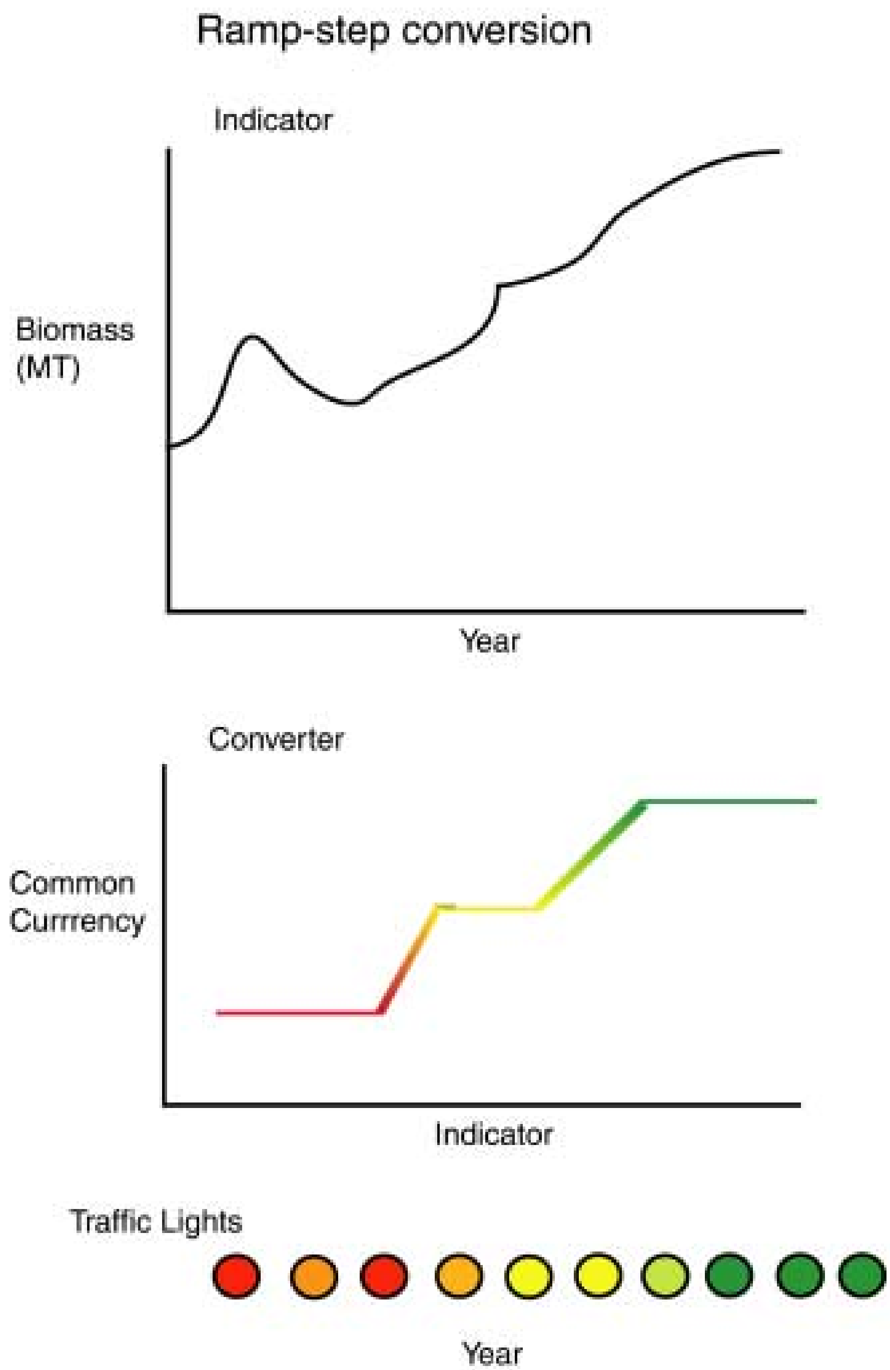


Figure 4. Schematic of ramp-step conversion of an indicator (Biomass) into Traffic Lights. The top portion is the input indicator, the middle is the converted to a common currency and the resultant lights are shown at the bottom.

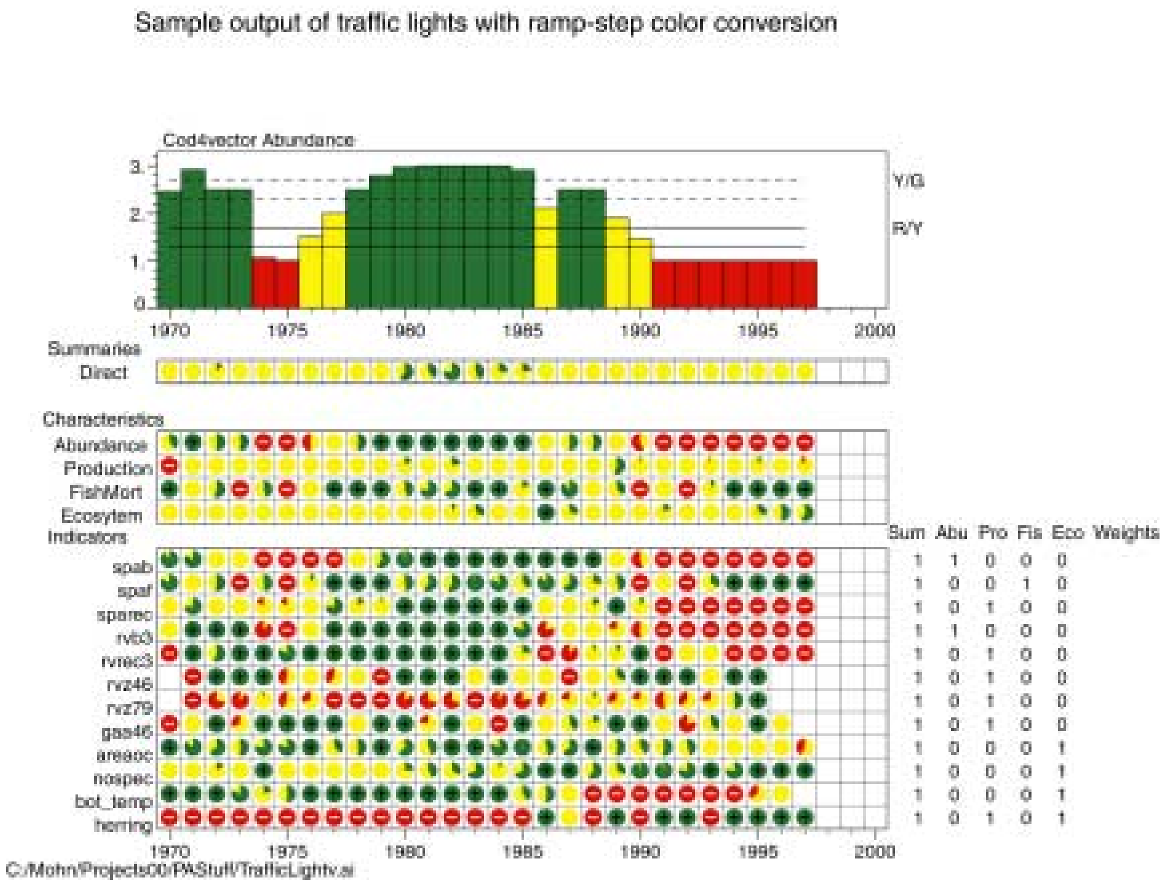


Figure 5. Example of ramp-step conversion using 4VsW cod data.

Annex 1. Agenda.

FMSWG MEETING

8-11 January, 2001

TOPIC: Use of the Traffic Light Method for Application of the Precautionary Approach to Fishery Management Planning.

PURPOSES OF MEETING:

- Learn from experience of the pilot application of the Traffic Light method at November 2000 groundfish RAP and assess client reactions
- Evaluate possible solutions to outstanding issues and assess Traffic Light performance through examination of test cases
- Establish a plan of action for production of a citable version of the Methods Workbook
- Make recommendations for further research and on future applications

AGENDA:

1. **Opening remarks**
2. **Introduction**
    a) *Origin and history of the topic in Maritimes Region*
    b) *Review of current status of the methodology*
3. **Approval of agenda and schedule**
4. **Indicators of the status of attributes**

a) *Indicator description content*
b) *Descriptions of new indicators*
    i. Stock status
    ii. Fishery status

5. **Criteria for Assigning Lights to Indicator Values**
6. **Integration of Indicators**
7. **Form of Scientific Advice/Interface with Clients**

a) *Presentation of Traffic Light table results*
b) *The Precautionary Approach*

8. **Traffic Light Case Studies**
9. **Next Steps**
10. **Review of Proceedings**

Annex 2. List of Participants

Annand, C.	Resource Management, Scotia-Fundy
Black, J.	MFD-BIO, Science, Scotia-Fundy
Branton, R.	MFD-BIO, Science, Scotia-Fundy
Bundy, A.	MFD-BIO, Science, Scotia-Fundy
Campana, S.	MFD-BIO, Science, Scotia-Fundy
Clark, D.	MFD-St. Andrews, Science, Scotia-Fundy
Evans, G.	NWAFC, St. John's, Newfoundland
Fanning, P.	MFD-BIO, Science, Scotia-Fundy
Fowler, M.	MFD-BIO, Science, Scotia-Fundy
Fu, C.	MFD-BIO, Science, Scotia-Fundy
Gavaris, S.	MFD-St. Andrews, Science, Scotia-Fundy
Halliday, R.	MFD-BIO, Science, Scotia-Fundy
Hurlbut, T.	MFD-GFC Moncton, Science, Gulf
Hurley, P.	MFD-BIO, Science, Scotia-Fundy
Koeller, P.	IMPD-BIO, Science, Scotia-Fundy
Liew, D.	Policy and Economics, Scotia-Fundy
McRuer, J.	MFD-BIO, Science, Scotia-Fundy
Mohn, R.	MFD-BIO, Science, Scotia-Fundy
Morin, R.	MFD-GFC Moncton, Science, Gulf
O'Boyle, R.	RAP Secretariat, Maritimes
Parsons, D.	NWAFC, St. John's, Newfoundland
Peacock, F. G.	Resource Management, Scotia-Fundy
Rivard, D.	Science, Ottawa
Showell, M.	MFD-BIO, Science, Scotia-Fundy
Silvert, W.	IPIMAR, Portugal
Simon, J.	MFD-BIO, Science, Scotia-Fundy
Sinclair, M.	Science, Scotia-Fundy
Smith, S.	IMPD-BIO, Science, Scotia-Fundy

Annex 3.

Report of Traffic Light Workshop
Moncton, Sept. 20, 2000

Attendees:

Gerry Black
Alida Bundy
Ghislain Chouinard
Don Clark (Rapporteur)
Paul Fanning,
Stratis Gavaris
Ralph Halliday,
Lei Harris
Robert Mohn (Chair)
Doug Pezzack
Gloria Poirier
Heath Stone
Doug Swain

Note: Items or comments that are from RAP will appear in Italics and will be mainly used to point out usage made obsolete subsequent to this Workshop. One exception is the category called characteristics. Although defined after the meeting, this word will be used extensively in this document without Italics. Similarly, although the term "Summary Attribute" was used at the Working Group, it has subsequently been denoted simply as Summary below.

Purpose: (Excerpts from invitational e-mail)

Reason 1. We need to have done our homework on traffic lights and the related topics of uncertainty and risk before the next RAP series. If not a common, at least a non-contradictory, approach should be formulated before going public.

Further to this point is FMSWG Ad-hoc Task Group on Implementation of the PA.

Minutes, Meeting of 17 May 2000:

> Item 6. A computer tool for examining the **integration of Traffic Light results (weighting of indicators)** and the implications of **reference point** decisions was demonstrated (MOHN, BLACK). It was agreed that this should be used in evaluation of several case studies in a workshop environment. Case studies should be selected for which there is convincing evidence, independent if the indicators being tested, for a major change in stock status. Div. 4VsW cod provides one example, Div.4T cod and Georges Bank herring are others, where the external reference provided by fishing experiences provide clear evidence of an "event" i.e. a large and rapid change in abundance. A limited range of indicator weightings will be considered, e.g. 0, 0.5, and 1.0. The intention is to reach conclusions that are general to the application of the Traffic Light method, not to determine "optimal" formulations for

particular stocks. The conclusions will be incorporated into the Workbook as guidance for users of the Traffic Light method. While it is necessary to engage the interest of those analysts responsible for assessment of the stocks that fit the selection criteria the workshop should be open to anyone who wishes to actively participate (within the constraint of practicality). The workshop is to be held no later than early September, as the results will be needed for the application of the Traffic Light method to groundfish assessments in the October RAP. Preparatory work could be co-ordinated over the summer.

Item 7. <u>The form and scope of advice</u> that can be given based on traffic light results should be part of the agenda for the workshop. In this context the related issues of uncertainty and risk could be addressed.

Presentations:

Two Working Papers were presented. The first by Gavaris was an example focussing on when linking advice to indicators. The second, by Mohn, was a case study using 4VsW cod data.

1. Gavaris paper

This MS focused on how to structure indicators in order to facilitate fisheries management decisions. This work investigated indicators with reference to risk analysis and the Precautionary Approach. The emphasis was to demonstrate how the consequences of alternative management actions, e.g. quotas, could be evaluated against established reference points for the indicator by projecting the future state of the indicator.

The resulting discussion posed the question “How can we use indicators to fit into the existing management framework?” We need to tell people what the impacts are of potential actions; what the consequences are.

It was alternatively suggested that we should be trying to explain with a number of indicators how we have arrived where we are, and what ‘state’ we might be in. Rather than suggesting a catch level, give a framework, yet to be defined, in which to make a sensible decision. Whatever the current management is, this suggests which direction they should go from here.

2. Mohn paper

This WP presented a case study based on 4VsW cod. This stock was chosen because of its dynamic history; near collapse in late 1970s, strong recover then collapse and closure in the early 1990s. Use an aggregation of indicators to provide a red/yellow/green measure of overall status of a resource. The indicators may be interdependent, and may be integrated as a whole and also separated into groups representing a given characteristic (See discussion of attribute vs. characteristic vs. summary given below).

3. Discussion:

Discussion followed on how less formal indicators (e.g. fishermen's experience) are to be used. They can be entered as an indicator by getting a time series of this information. How to weight indicators is then difficult, both in terms of how to weight more subjective indicators and also the problem of weighting short time series.

This led to the general question of how will we deal with new indicators. Given three or four years of a relative index on fishing success this could be input, once the meaning of the indicator is determined in relation to stock status. The weighting of the indicator might start low, then gain weight once its relation to stock status becomes more apparent. *How to deal with short time series was dealt with at RAP by down weighting them in either the direct summary or into Characteristics. However, this was viewed as a stop-gap measure pending a more satisfactory solution.*

The related question of how to deal with non-independent indicators was also addressed. If there are three measures of biomass, i.e. three indicators of an "attribute"; they could each be given less than full weighting so that other characteristics with fewer measured indicators would not be overwhelmed in a direct summary. If characteristics are defined and used, this number of underlying indicators is not a factor in the definition of an indirect summary.

General Discussion:

Issue: Should economic indicators or fishery indicators enter here? Are we measuring status of the stock, or a broader fishery status? Should status of the resource be a separate module within a decision making process which might also include socio-economic measures.

Conclusion: Status of the stock would be a constraint to managers in addressing the status of the fishery. In the context of providing biological advice on stock status, at RAP, indicators such as catch may not be directly relevant. This might be relevant at another, management related, venue. Perhaps other indicators would be added to the same table at a fisheries management venue. Thus, as the resource progresses through the review and management systems, new indicators may be added.

Issue: Should we show an indicator if we do not know how it is linked? Does this abrogate the responsibility of making a scientific conclusion to a body that may lack the expertise to make it?

Conclusion: Perhaps the list of indicators might be more restricted in the SSR, but other indicators included in a Working Paper for discussion at RAP.

Issue: Debate of merits of different approaches to summarizing information; directly or by linking like indicators into Characteristics.

Conclusion: Consensus could not be reached and it was proposed that both methods be tried to get more experience with how each performs. Some argued that low exploitation (green) and low abundance (red) should be distinguished from high exploitation (red) and high abundance (green). A single, overall indicator in this case would be yellow, even though different management response is indicated. This illustration argues against the integration of fundamentally different indicators. It should be possible to devise decision rules that account for a few fundamental characteristics in a more intelligent fashion than simply averaging. *In the build-up to RAP both were tried (a simple sum for the Direct Summary and a combination of Abundance and Production Characteristics for the Indirect Summary) and in the majority of cases they performed quite similarly. For RAP only Direct Summaries were presented*.

Issue: Whether to use discrete (3-state traffic lights) or semi- discrete (more than 3 states) or continuous variables in analysis and advice. The related question of when to discretize, especially with regards to integration was also posed.

Can a linkage be found between discretized overall summary space and the terms of advice? The coarseness on the input should dictate the coarseness of the advice. In other word if the Summary has only 3 states this might be equivalent to block of 100 kt of biomass, how to generate advice on removals at a finer scale?

Conclusion: General consensus was not reached on this topic and it is hoped that future work (perhaps a workshop) could address it. One proposal was that indicators could be presented in the context of a risk analysis, as done in Gavaris's paper. This is a different format than the traffic lights. Give the value of the indicator as well as its status (colour). For some indicators this could be done in a risk analysis format. Having both values and traffic lights would allow those who make management advice to decide which format is the more useful.

Issue: Formulation of advice.

Discuss the biomass indicators, the recruitment indicators, the production indicators, then provide advice on each to some degree (protect recruitment, removals at a low level). Where stock is dynamic, appearing to increase or decrease substantially, we may provide advice on a magnitude of change in harvest, or conservation levels, and indicate close monitoring is required (annual) while if stock 'status' appears static, longer term monitoring at a status quo management could be advised. If we are satisfied that the indicators fit with the estimates of stock status, give the numbers that come from projection. If indicators are unclear and contradictory, then the VPA is used as an indicator, not presented as an absolute estimate that provides concrete advice on harvest levels.

What happens close to a boundary? Do we give categorical advice, as indicated by categories, or do we indicate where within the category we are? This is turning categories back into continuity. Does saying large change not indicate a scale which is similar to giving specific advice? It implies a proportional change on top of a current exact value.

Conclusion: A traffic light table should include all the indicators, the integrated indicators (*later given the name Characteristics*), then an overall integrated estimator (*later given the name Summary*).

Issue: Definition of Boundaries.

Lots of discussion of where boundaries should go and what they mean. The use of boundaries that are based on the range of observations in the available time series compounds the problems associated with averaging over indicators that do not all span the same time period. Particular care should be taken in circumstances where indicators span different time periods and the boundaries are determined by the range of observations in each time series.

If there is no integration, then those using the table may integrate in some fashion. Also, it was proposed that managers could use a 2 or higher dimensional matrix to assist in making decisions. Where an indicator is measured imprecisely, but has a fairly direct linkage to the characteristic of concern and another is precisely measured, but is less clearly linked to the characteristic can we integrate these rationally? The rejoinder is “what is the purpose of the advice”, do we need to phrase advice directly in relation to biomass, or can we provide analyses of a variety of biological characteristics which when integrated suggest the general state of the stock, and the nature of management which needs to be considered?

Conclusion: Consensus was not reached on this issue.

Issue: How many Characteristics?

A characteristic is some aspect of a resource that is inferred from indicators. Should characteristics be those things we want to specifically comment on in describing stock status; things which will be directly useful for management considerations; those characteristics we will use in formulating our advice?

Conclusion: No. The top lines in the traffic lights display will be called summaries. “Characteristics” consisting of a set of indicators may feed into the summaries, or summaries may be computed directly from individual indicators.

A traffic light table should include all the indicators, the integrated indicators (*characteristics*), and then an overall integrated estimator (*Summary*).

If two indicators of the same characteristic diverge it may suggest that the status of the characteristic is uncertain and an average may be inappropriate suggesting that although we have indicators a characteristic cannot be resolved. On the other hand the integration may be performed but say the characteristic is poorly known.

Proposed Summary Headings *(Characteristics):*

1) Abundance

2) Mortality

Should M and F be considered separately? Should M go into productivity? This allows guidance to come through directly on how to manage exploitation.

3) Production

Should this be hierarchical? Do we need to summarize all growth, all condition, all recruitment indicators before you go to production? This can be handled by weighting, with all measures (indicators) present.

4) Environment

How does this feed into stock status? Should those ecosystem indicators that influence productivity simply go into a productivity characteristic.

Issue: How to handle summaries.

A suggestion is to produce 2 'Summary" lines. One will be a multinomial integration of all indicators, and one will be produced from the 3 or 4 "summary characteristics" that have been selected for the stock. These are defined as Direct and Indirect Summaries respectively.

In a 3x3x3x3 summary space of 4 characteristics, rather than a strict integration, there could be 'if then else' rules. If A low, B low, C med, D med, then…. A final option would be to make the final traffic light a summation of the summary of the indicators. This is a loose approximation of the 'If A low, B low, C med, D med, then…' rule structure, for providing a light. *This has been done and is the basis of the Indirect Summary used in preparation for RAP*.

Conclusion: A set of integrations of indicators which works for all stocks may not be feasible at this stage, but coming up with a variety of integration rules for different stocks does not seem appropriate.

Issue: If an analytical assessment is available, should projection products be done? How will they be used in a traffic light structure?

- Perhaps project only with the current management (status quo TAC). This would provide information on which way management should go.
- Could we give prognosis on the state of a characteristic (production will remain green/red next year) for use in a traffic light to go along with projections from SPA?
- If we can give a subjective interpretation of what is "significant" change, would this come from a projection that is not shown?

If projections (from an analytical model) of some nature might be included, why not give them in the form of an 'Armstrong' plot with *status quo* removals?

Conclusion: There was no consensus that in the majority of analytical assessments we can say with any certainty what the specific F would be for any given catch. Among the views discussed were:

- At most we should indicate if projections at the status quo fishery appears to leave the resource status good, bad, indifferent, and then suggest direction of change required.

- If you do not feel you can do an adequately reliable projection, do not. If you do, why not do a projection at Status Quo plus 10 or 20%?

Issue: Default Boundaries for Indicators.

Conclusion: When comparing against the mean in cases where more is better (like biomass), green/yellow is at the mean, and yellow/red is 60% of mean. If indicator is one for which more is worse (like mortality) yellow/red is at 140% of the mean.

Issue: Should the red/yellow boundary in a Summary be called a Limit (in the PA sense)?

Conclusion: No, for this year red can be referred to as a serious situation rather than a Limit.

Furthermore, the following definitions for each colour were proposed:

Red – severe situation, intervention required, continued or augmented
Yellow – adjustments may be required to control fishing; check characteristics for details
Green – status quo or increased removals

May need to differentiate between when yellow is a sum of yellows and when it is a combination of conflicting colours. There will probably always be a need for a verbal synthesis.

A stock that is in a red or yellow condition may take some years to move to green. The necessary management actions may have been initiated but the colour will not change for some time. There is a need to recognize this in formulating advice and perhaps to evaluate if the desired effect is being achieved, i.e. can and should we distinguish changing state within a colour.

Issue: Determination of weightings.

Conclusion: Weighting (sliders) should be shown, and a rationale given on why we have selected these weightings. This will be open to discussion at RAP, and must be available for consideration. *Weights were limited to 0, .1, .5 and 1. These defaults carried through RAP although in most cases the partial weights were used for shorter time series.*

Issue: Uncertainty:

A subjective assessment of uncertainty is available visually from the table of indicators, how much are colours mixed as opposed to areas of homogeneity. A subjective expression of uncertainty could be included based on the number of indicators and how well they are felt to reflect the underlying stock characteristics.

Conclusion: There was no consensus on how to express uncertainty.

Issue: Communication with RAP and FRCC.

Conclusion: An introduction to the methods will be presented at RAP to introduce the vocabulary and the general methods. Before RAP it may be appropriate to send to the FRCC a brief summary of the methods we will follow this year in presenting our stock assessments. This may not be the only format of analysis; in that it will be an important one, here is how it works.

Issue: Status of the Draft Traffic Light Methods Workbook.

Conclusion: The Workbook was presented to this Workshop as an information item. We have not been asked to approve it, nor was approval considered.

Issue: Post-RAP agenda.

Conclusion: A workshop is proposed for early January, which would give GFC the time to incorporate results before their RAP. There is still a need for work to answer the technical questions. The November RAP could provide data for future case studies. Also need people willing to try some ideas and criteria to judge success.

The upcoming WG should also address uncertainty – how is it dealt with here how to capture and communicate in the Traffic Light context.

Glossary

Actions. Mathematical operations used in defining Boundaries. The most common would be, Mean, Maximum, Geometric Mean, Percentile or a fixed value (F0.1).

Attributes. An attribute is a basic property of a fish, fish stock or fishery, e.g. a fish stock has an abundance measurable by its numbers or biomass, a growth pattern may be measured by its instantaneous growth rate or its length at age.

Boundaries. Decision points for indicators that separate the indicator into green, yellow and red. The decision points are used in conjunction with Actions. For example, the green/yellow boundary might be set at 80% of the mean, set at 40% of the maximum observed. In first example the boundary is 0.8 and the action is Mean; in the second it's 0.4 and Maximum.

Characteristic. A characteristic is a conceptual entity based upon a number of indicators. Its purpose is to aggregate similar indicators for further analysis or discussion. A characteristic is empirically based and may or may not reflect an attribute of the resource. Characteristics may also be combinations of other characteristics.

Indicator. A time series of measurements, observations or model outputs. They may be either quantitative or qualitative.

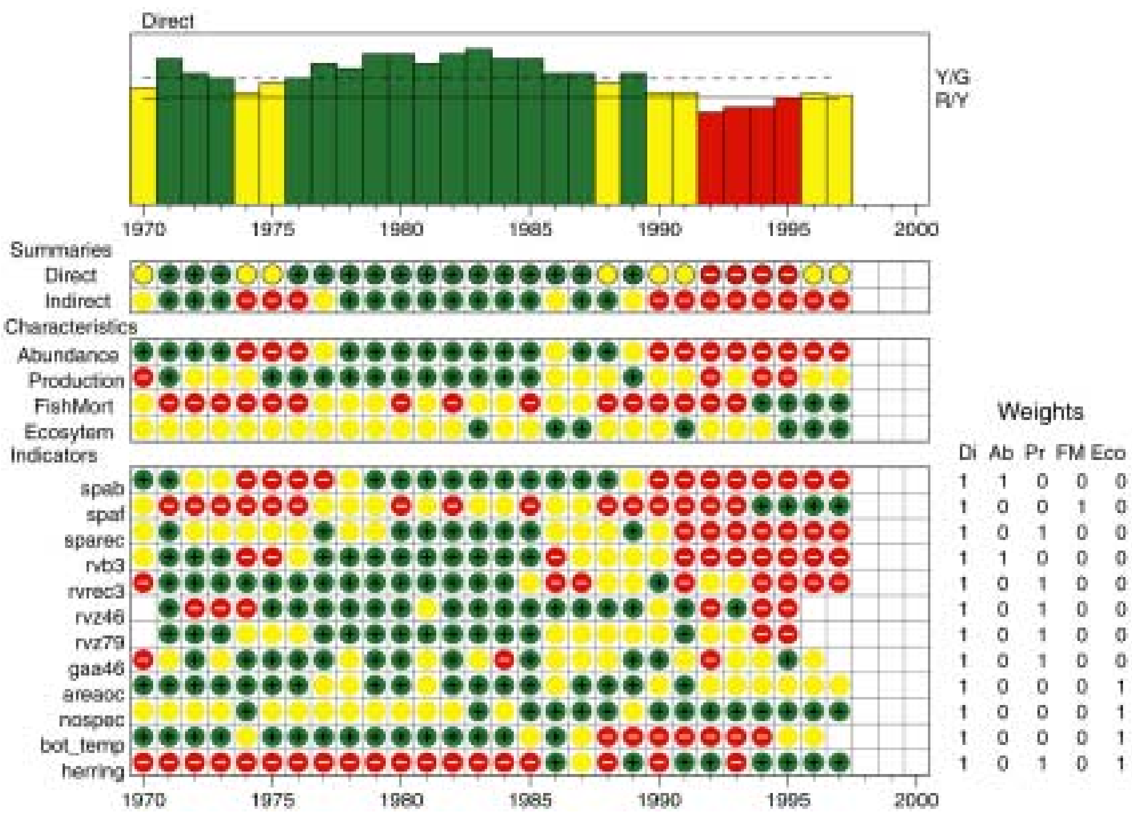
Summary. An overall index of stock status in the traffic light system. Two types of summary were proposed. A weighted sum of indicators is called a Direct Summary. A combination of Characteristics is called an Indirect Summary.

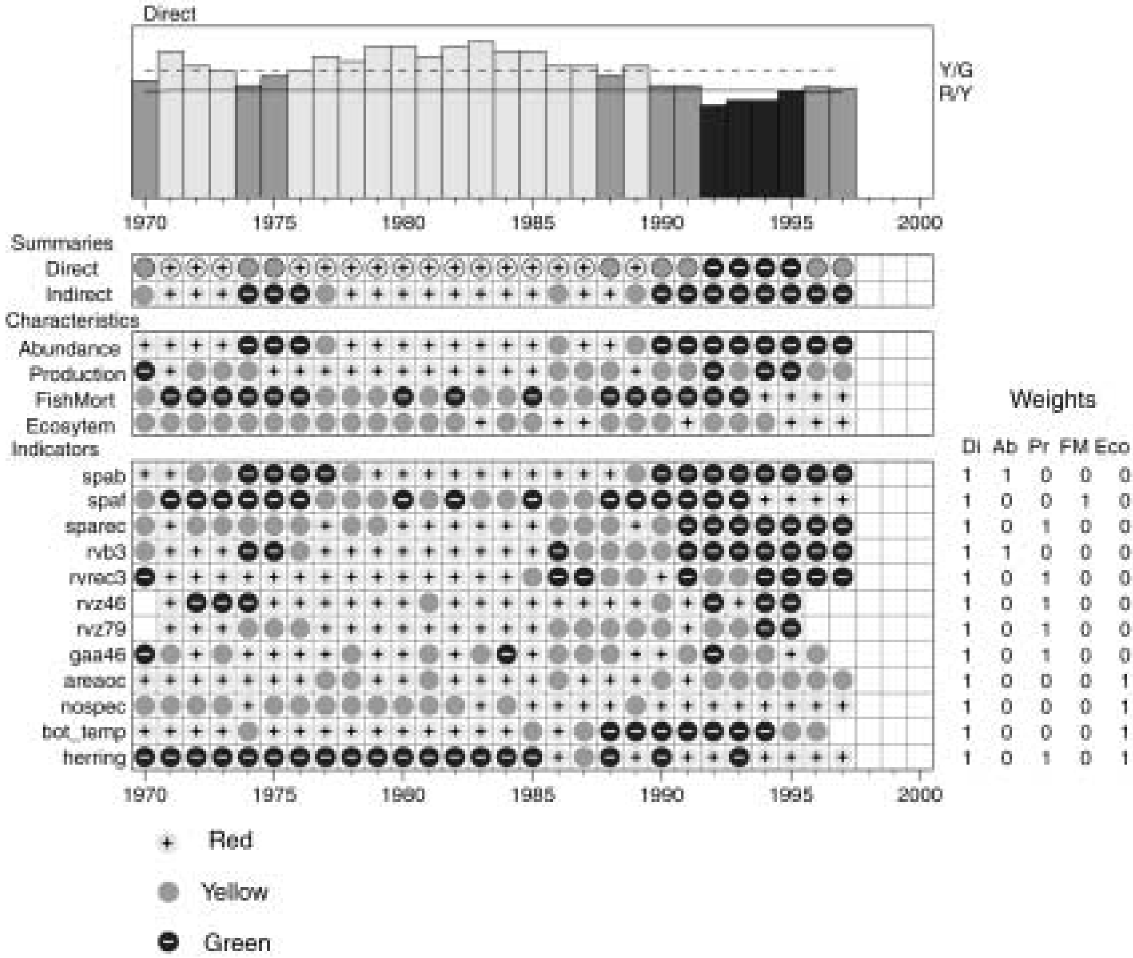
Weights. Values between 0 and 1 that are assigned to each indicator and scale their relative importance when rolled up into Characteristics or a Direct Summary.

<u>Appendix A. Sample Traffic Light Output</u>

The following figures (colour and gray) show sample traffic light output using data from 4VsW cod. The bottom block contains 12 indicators from SPA, survey, environmental and ecological sources. The Indicators, spab, spaf and sparec are respectively the biomass, fishing mortality and recruitment from an SPA. Similarly, rvb3, rvrec3, rvz46, rvz79, gaa46 are survey based biomass, recruitment, total mortality rates over two age ranges and growth rate. The area occupied (areaoc) is derived from proportion of non-zero sets in the summer survey as well is the number of species reported for each tow (nospec). Bottom temperature (bot_temp) and herring abundance (herring) are also from the summer survey series.

To the right of the indicator block are the weights for each indicator used to form the Direct Summary (Su) and the four Characteristics (Abundance, Ab; Productivity, Pr; FishMort, FM; and Ecosystem, Eco). In this example, the characteristic Abundance is the average of the Indicators spab (SPA biomass) and rvb3 (3+ biomass from the survey). The uppermost block of lights provides the two Summaries. The direct is the average of all the indicators as shown by the column Su in the weights. The Indirect is a combination of the Abundance and Production Characteristics. The histogram at the top of the figure is the Direct Summary and the boundaries between yellow and green (Y/G) and red and yellow(R/Y) are shown as dotted lines.




Direct
Y/G
R/Y
1970
1975
1980
1985
1990
1995
2000
Summaries
Direct
Indirect
Characteristics
Abundance
Production
FishMort
Ecosytem
Indicators
spab
spaf
sparec
rvb3
rvrec3
rvz46
rvz79
gaa46
areaoc
nospec
bot_temp
herring
Weights
Di Ab Pr FM Eco
Red
Yellow
Green

Annex 4.

REVIEW OF

"Use of the Traffic Light Method for Application of the Precautionary Approach to Fishery Management Planning"

DRAFT July 2000

[Comments by Dr. Esra Karasakal, Visiting Post Doc–Faculty of Administration, University of Ottawa]

- ✓ Statistical Process Control (Q-charts, P-charts) can be employed to analyze the system and to diagnose the problems. Also time series analysis should also be considered to forecast the future periods.

- ✓ In this problem, there are actually many decision makers which introduces a great deal of complexity into the analysis. We can put this problem into a general category of problems known as 'group decision making problems' (see, e.g., Keeney and Raiffa,1993, pp. 515-547).

- ✓ Before starting to solve the problem, criteria and alternatives should be defined clearly. In group decision making, the criteria for evaluating the alternatives can be chosen using Delphi-technique (see, e.g., Hwang and Lin, 1987). For structuring criteria, one should refer to the views expressed in Keeney and Raiffa (1993, pp. 31-65). A hierarchical model for criteria is usually suggested for dealing with complex real-life problems (see, e.g., Saaty, 1995).

- ✓ The basic methodology is not appropriate for analysing the multicriteria problem because decisions will be very much dependent on the historical data (i.e. reference points are determined based on the historical data). This may result in biased decisions. For example, if historical data are very low, then the level of the reference points will be low. Thus an overestimation about the current condition may be possible.

- ✓ In my opinion, this approach does not provide enough resolution. In some approaches, some criteria are measured on interval scales, because it is not possible to measure the values of those criteria. For ex, in some approaches, you are allowed to measure some criteria on interval scales. But the levels of scales are determined considering the properties of the criteria.

- ✓ In the traffic light approach, weights of the criteria are calculated considering
  - Precision of the estimator
  - Independence of criteria
  - Etc.

  However, weights are not used to handle above issues in MCDM. It only shows one criteria is either more or less important than the another.

Some basic methods that are normally used to calculate weights are:

- Successive comparisons (Knoll and Engelberg, 1978)
- Eigenvalue method (Saaty, 1977)
- Ranking (Kendall, 1970)
- Cardinal evaluation (Edwards, 1977)
- Etc.

It is worth noting that criteria data obtained from more than one source should be combined before determining weights in order to avoid overweighing of criteria (for ex., ranges can be defined for the values of criteria).

There are some methods developed for multicriteria problems (Stochastic Multicriteria Acceptability Analysis- SMAA), where criteria data are uncertain or inaccurate and where for some reason it is impossible to obtain accurate or any weight information from the DM(s). (see, e.g., Lahdelma et al., 1998, Lahdelma and Salminen, 1999).

✓ If data are dependent, a method that handles dependency in the data (ANP) (see work by Saaty) can be used.

References

Edwards W. (1977) Use of Multiattribute Utility Measurement for Social Decision Making, in Conflicting Objectives in Decisions, Bell D. E., Keeney R. L., Raiffa H., Eds., 1977, Wiley, pp. 247-276.

Hwang C-L, Lin. M-J (1987) Group Decision Making under Multiple Criteria. Springer-Verlag, Berlin Heidelberg

Keeney R. L., Raiffa H. (1993) Decisions with Multiple Objectives- Preferences and Trade-offs. Cambridge University Press, New York, USA.

Kendall M. (1970) Rank Correlation Methods, Charles Griffin, London, 4rt Edition.

Knoll A. L., Engelberg A. (1978) Weighting multiple Objectives. The Churchman-Ackoff Technique Revisited. Computers and Operations Research, 5, 165-177.

Lahdelma R, Hokkanen J, Salminen P (1998) SMAA - Stochastic multiobjective acceptability analysis. European Journal of Operational Research, 106, 1, 137-143

Lahdelma R, Salminen P (1999) SMAA-2: Stochastic Multicriteria Acceptability Analysis for Group Decision Making. Forthcoming in Operations Research

Saaty T. L. (1977) A Scaling Method for Priorities in Hierarchical Structures. Journal of Mathematical Psychology, 15, 3, 234-281

Saaty T.L. (1995) Decision Making for Leaders: The Analytic Hierarchy Process for a Complex World. RWS Publications, 3rd Edition. Pittsburgh, USA.

Annex 5.

Template outlining the Sections of a Traffic Light Indicator Account

(with a sample account for the RV index of recruitment for 4VsW cod).

Indicator: Descriptive name of indicator including short form used on Traffic Light output.

Characteristic: Name of the characteristic of stock status the indicator reflects. Current choices include abundance, production, fishing mortality and ecosystem.

Description: Detailed description of the indicator of an attribute, including data source(s), range selections, computation including transformations, smoothing etc.

Boundary Point(s): Basis for setting boundary points between colour ranges. Although statistics based on historical data series have been widely used it is preferable that some external basis for determining the ranges be applied, especially in cases where there is a short data series or little dynamic range in the data. Intercalibration of data series may be an option in some cases.

Properties:

Measurability: (our confidence in estimate of the indicator)

- statistical properties of estimator e.g. variability, bias, skewness
- transformations if required
- standard or alternative formulations for estimator, if non-standard, why?
- consistency with other estimators of indicator

Interpretability:

- how does this indicator reflect stock status or the identified characteristic?
- what caveats exist?
- how well are the colour boundaries related to changes in stock status or characteristic?

Sensitivity:

- how rapidly does indicator respond to changes in stock status?
- is there adequate time for management interventions?
- does the natural variability likely mask real changes?

Weight:

The overall value of an indicator, on a scale relative to the other available indicators. This is based primarily on the strength of the individual attributes and the qualities of the proposed reference points. At least four types of uncertainty have been identified:

1. Statistical uncertainty due to sampling error in the indicator related to measurability
2. Boundary point uncertainty related to interpretability
3. Importance of the indicator, again related to interpretability

4. Structural uncertainty, the sensitivity of the indicator to changes in the formulation of the estimator.

Review of Performance or **Validation of Indicator:**

If data exist, an assessment of the performance of the indicator over a time series of estimates is desirable. An alternative approach would be to validate the indicator and boundary points against external information, e.g. $F_{0.1}$, or critical values for fish condition. In either case the review must demonstrate the adequacy of the indicator, its estimator and the selected boundary values.

Indicator: 4VsW cod Research Vessel Index of Recruitment (RVrec3)

Characteristic: Production.

Description: The July RV stratified log transformed mean catch per tow for a cohort at age 3.

Boundary Point(s): Arbitrary boundary points, based on the available historical series, are proposed with the GY boundary at the geometric mean of the series and the RY boundary set at 0.6•GY. This approach is not suitable in the case of a short time series, or a series that does not include a substantial fraction of the full dynamic range.

Properties:

Measurability:

- Research vessel surveys are noted for high variance and substantial 'year effects'.
- Given the generally skewed nature of RV survey data, a log transformation to stabilise the variance has been used.
- Several alternate formulations have been used in various stocks which may alter results
  - In some cases more than one observation is available for a pre-recruit year class, either from more than one survey series in the same year or at more than one age in successive annual surveys.
  - Catchability weighting and smoothing are optional in situations of high noise.
- Confidence in the index will vary from stock to stock depending on the suitability of the survey for measuring recruitment. Once a time series has been established, comparisons can be made with other indices of year class strength, e.g. from SPA, to estimate confidence limits on predictions.

Interpretability:

- This is a simple and direct measure of recent and current recruitment with respect to past observations based on designed and controlled direct sampling of the stock. Recruitment *per se* may fluctuate widely and individual year classes may be very small without indicating an overall decline in stock size or productivity. Much greater weight should be given to several successive small year classes than to a single observation.
- The arbitrary boundaries are a weakness but, given the wide range seen over the 30 years of this series, they do seem to represent the bounds of good and bad recruitment.

Sensitivity:

- The relative year class strength is measured annually. Changes in the index are known a year or more before year classes are entering the fishery. Thus, the index can provide several years' forewarning of changes in status for this stock.

Weight:

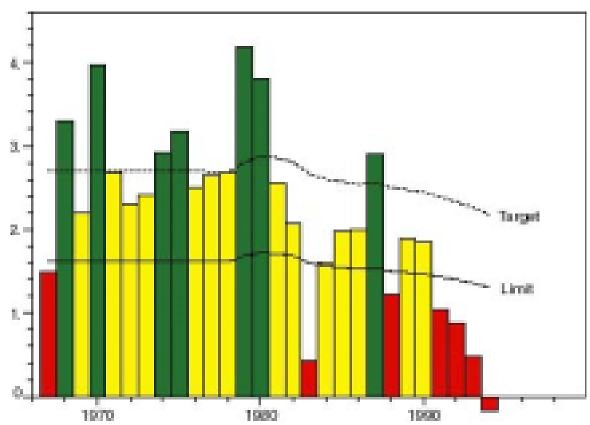
i) This indicator is subject to large sampling variation including large 'year effects'. The distribution is highly skewed, requiring the log transformation. Statistical uncertainty is relatively high.
ii) The boundary points are arbitrary leading to uncertainty about their appropriateness. This is offset to some degree by the long time series and large range observed throughout it. Thus, uncertainty is moderate.
iii) This is one of the most important indicators of stock production and hence future potential yields. There is little or no uncertainty about the importance of this indicator.
iv) A variety of recruitment indices from RV data could be used to produce a smoother index, e.g. averaging several estimates over years. Although smoother, they are not likely to differ in trend and suggest little uncertainty about the structure of the estimator.

Overall, this indicator should get high weight.

Review of Performance:

This indicator clearly shows that the year classes since 1988 are all small to dangerously small. However, it also shows that several year classes were estimated to be very small (1967 and 1983) at times when the stock was not declining. Although a red light would have appeared in 1991 (1988 yc at age 3), waiting for a strong warning signal of successive red lights would have taken until 1995. The single negative estimate (the 1995 cohort) is simply the log of a number less than 1, making it the smallest in 30 years.

The indicator is very effective at identifying small cohorts and all of the 'red light' cohorts proved out to be small in subsequent years.



Yearclass at age 3