ICES Journal of Marine Science, 62: 552–557 (2005)
doi:10.1016/j.icesjms.2004.12.014



# A framework for making qualities of indicators transparent

Kjellrun Hiis Hauge, Erik Olsen, Hilde Elise Heldal, and Hein Rune Skjoldal

Hauge, K. H., Olsen, E., Heldal, H. E., and Skjoldal, H. R. 2005. A framework for making qualities of indicators transparent. – ICES Journal of Marine Science, 62: 552–557.

A framework is presented to communicate different qualities of and uncertainties in indicators, and to enhance the transparency of conclusions and scientific advice based on them. The framework consists of four parts: Advice Statement, Power of Explanation, Robustness, and Performance Perspectives. The Advice Statement is descriptive and largely quantitative; the last three parts are more analytical and qualitative. Power of Explanation addresses the strength of the associated scientific knowledge. Robustness seeks to clarify how sensitive the indicator value and advice are to underlying assumptions. Performance Perspectives highlight prospects and drawbacks in a management context. The framework and its discussion are illustrated by examples on spawning-stock biomass of Northeast Arctic cod, technetium-99 concentrations in lobster, and harbour porpoise bycatch rates.



Keywords: indicators, qualities, transparency, uncertainty.

Received 1 April 2004; accepted 29 November 2004.

*K. H. Hauge, E. Olsen, H. E. Heldal, and H. R. Skjoldal: Institute of Marine Research, PO 1870 Nordnes, N-5817 Bergen, Norway. Correspondence to K. H. Hauge: tel: +47 55 238590; fax: +47 55 238687; e-mail: kjellrun@imr.no.*



## Introduction

There is currently great interest and emphasis on the use of indicators related to environmental and resource management (FAO, 1999; EEA, 2002, 2003). In addition, there is an increasing focus on participatory processes in fisheries management[1] and ecosystem-based management (Directorate of Fisheries, 2003). Selecting indicators, or developing management strategies based on indicators, demands clarity on quality aspects such as scientific rigour and relevance. By scientific rigour we mean the degree of coverage (in time and/or space), consistency, measures of uncertainty (known error distribution, bias), and power (whether changes in trend can be detected with statistical significance). Quality related to relevance would cover representativeness (correlation between an indicator and the issue for which it is supposed to be a proxy), sensitivity to underlying assumptions, and whether it is a suitable basis for management decisions.

ICES (2002a) has developed a set of criteria to help evaluate the usefulness of indicators of ecological quality: (i) relatively easy to understand by non-scientists and those who will decide on their use; (ii) sensitive to a manageable human activity; (iii) relatively tightly linked in time to that activity; (iv) easily and accurately measured, with a low error rate; (v) responsive primarily to a human activity, with low responsiveness to other causes of change; (vi) measurable over a large proportion of the area to which the metric is to apply; (vii) based on an existing body or time-series of data, to allow a realistic setting of objectives; and (viii) relationship to a state of wider environmental conditions. These criteria emphasize the practical aspects of strong linkages with manageable human activities through corrective measures. This is no doubt an important consideration. However, it is by no means certain that indicators fulfilling all demands can be found. Which criteria are the most important as a basis for management decisions is in the end a political choice, especially because the relevance and usefulness of an indicator are important considerations.

To judge the quality and relevance of an indicator, users need a transparent presentation of the scientific background and of the uncertainties involved. Of course, transparency is a basic requirement in science to allow for testing and validation of results, and to facilitate scientific debate. However, transparency also makes it easier for people with different backgrounds to participate in discussions about applications. While a detailed description of data and the

[1] See "Reforming the common fisheries policy" at http://europa.eu.int/comm/fisheries/reform/index_en.htm.

methods used may be sufficient for scientific colleagues, non-experts may have other needs. Knowing the underlying assumptions, simplifications, and other scientific judgements is useful, as is knowing how they affect the indicator and the objective to be agreed upon, and how well-founded is the underlying knowledge. In policy questions with conflicting interests, concealed uncertainty results in lack of trust in scientific advice. Transparency in conveying sources of uncertainty and how it affects advice and management should increase credibility and hence the quality of the advice. In contrast, lack of transparency may make complex issues look simple, but may also result in false expectations.

ICES has expressed the need for transparency and clarity in both fisheries (ICES, 1999, 2002b) and ecosystem-based (ICES, 2002a) management advice. However, transparency requires a methodological approach. Funtowicz and Ravetz (1990) constructed a notation scheme for multidisciplinary science projects. They address and categorize all sorts of uncertainty to achieve transparency in scientific results. Developing this scheme further, Van der Sluijs *et al.* (2002) designed diagnostic diagrams and sensitivity analyses to detect and assess key uncertainties in a $CO_2$-emission model, and how they influence management decisions. While impressive in its details and concerns, we believe that neither the users nor scientists are ready for such a detailed presentation of uncertainty. Awareness of its many facets and roles must first be awakened.

## A suggested framework

We regard four aspects as important for ensuring indicator transparency: a clear description in the context of associated knowledge, its scientific foundation, the robustness of its value, and its performance in a management context. Accordingly, our suggested framework consists of four parts and covers quantitative as well as qualitative uncertainty considerations: (a) Advice Statement, (b) Power of Explanation, (c) Robustness, (d) Performance Perspectives. Each part is split into several specific topics. The framework is illustrated by means of three summary tables to illustrate inherent differences in indicator quality: spawning-stock biomass (SSB) of northeast Arctic cod (*Gadus morhua*; Table 1), technetium-99 (Tc-99) concentration in lobster (*Homarus gammarus*; Table 2), and harbour porpoise (*Phocoena phocoena*) bycatch rate (Table 3).

The Advice Statement (a) covers what is usually found in scientific advice. It starts with the time-series of data (table or graph) to show historical development, preferably with uncertainty measures (error bars). An indicator function explains what the indicator is supposed to reflect, whether it is a proxy, and whether it is meant for management or monitoring purposes. Proxies are simplifications, implying added uncertainty. The user should be able to obtain a fair impression of how well a proxy reflects the issue of concern (addressed in the boxes b−d).

Indicators used for management purposes are often linked to a threshold or other types of (target or limit) reference points. Their values should be stated with a technical explanation. If a threshold has not been established, at least the specific concern and how it relates to the indicator should be described. Generally, there is supporting information on the issue, and a conclusion/advice. How this is presented may vary from a single qualitative statement to quantified predictions or suggested scenarios.

The Power of Explanation (b) reflects the soundness of the underlying scientific knowledge: what is known about cause and effect in respect of indicator values. The choice of a threshold may be somewhat arbitrary, like defining a risk at a certain probability level. Where different choices of thresholds affect stakeholders differently, this must be clearly stated. The Tc-99 threshold for what is considered to be a health threat is substantially higher than concentrations measured in fish in Norwegian waters. Still, fishers may be afraid of losing markets because of consumers' lack of trust in authorities. In such cases, openness to how science and/or politics have reasoned adds to transparency.

The ability to detect change may differ because of inherent delays (as for SSB; Table 1) or because of a low signal-to-noise ratio. In contrast, changes in Tc-99 concentrations (Table 2) can be measured immediately. If a statistical power analysis is carried out, the results should be presented. In a management context, it is useful to be able to separate the effects of different causes, such as the natural variation from man-made effects. Whether this aim is achievable varies, but it should be stated.

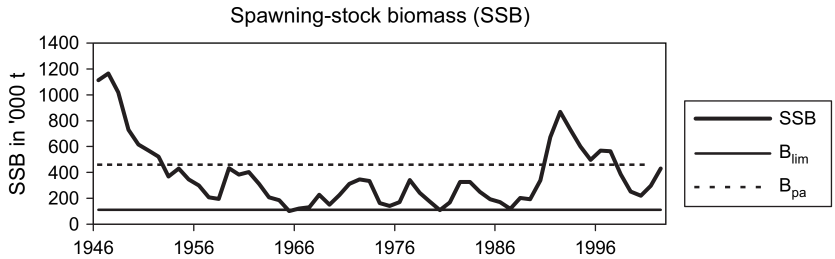
Scientific knowledge builds on assumptions. Although it is perfectly acceptable scientifically to restrict the validity of results to a certain set of assumptions, caution must be stated if the results are applied in a policy context. Assumptions may be critical as to whether the science is relevant or not, no matter how good the science is in itself. Identifying arbitrary choices (simplifications, judgements, generalizations) and underlying assumptions, together with their effects on the conclusions in terms of Robustness (c), improves the understanding of the uncertainty involved.

Underlying assumptions should be stated for indicators and thresholds separately. If the indicator is a straightforward measurement with a controlled error, a sensitivity analysis is irrelevant (Table 2). However, indicators may be calculated on the basis of some model of interacting factors. For example, predictions of spawning-stock biomass (SSB; Table 1) rest on numerous sampling schemes and assumptions, where the advice may be quite sensitive to assumptions on maturity, confidence in catch statistics, and interactions with other stocks. If a sensitivity analysis is lacking, the indicator is not validated properly, and the associated uncertainty must be regarded as unknown. Also,

Table 1. Framework example: SSB level of Northeast Artic cod.

a) **Advice statement** (ICES, 2003)

*Indicator time-series*:




*Indicator function*: SSB: proxy for recruitment potential; used in fisheries management to avoid low recruitment

*Threshold*(*s*): $B_{lim}$ = 220 000 t (changepoint regression), $B_{pa}$ = 460 000 t (lowest estimate having >90% probability of being above $B_{lim}$)

*Supporting information*: Majority consists of first-time spawners. Uncertainty considered larger than previous years. Concerns about under-reporting of catches. Time-series of catches, fishing mortalities, etc.

*Conclusion/advice*: Stock is outside safe biological limits. Reduce F considerably to less than $F_{pa}$ (0.40). (Catch forecast table could be included here with predicted SSB)

b) **Power of explanation**

*Cause/effect in indicator*: Fishing affects SSB. Environmental conditions affect growth and hence maturity and total biomass. Environment affects recruitment, and later SSB

*Cause/effect in threshold*: Threshold based on historical averages and does not present a distinct danger

*Ability to detect change*: To some extent. Indicator value and changes in trend are established after some years

*Ability to separate effects*: To some extent

c) **Robustness**

*Underlying assumptions indicator*: No unreported catches (at least the past years)

Assumptions on maturity

Assumptions in data interpretation (scrutinizing echograms, age reading, etc.)

Parameters and weighting of data sources correct

Assumptions on stock interactions (with capelin, cannibalism, etc.)

Mathematical/statistical model is well defined and numerical model is stable

*Sensitivity analysis of indicator*: SSB estimate sensitive to assumptions on maturation, stock interactions, and reliability of catch statistics. No sensitivity analysis is presented in official advice

*Underlying assumptions in threshold*:

Historical measures of SSB and recruitment of sufficient precision

Constant reference points valid, independent of today's fluctuations in biological parameters and in environmental conditions

SSB is a good proxy for reproductive potential

Reduction in F will eventually rebuild stock, independent of present stock

Uncertainty in estimated SSB constant from year to year

*Sensitivity analysis of threshold*: Done to some extent, but not presented in official advice

d) **Performance perspectives**

*Ability to adjust indicator level*: Reduced/increased fishing affects SSB immediately, but extent difficult to estimate. Natural variation too large to keep SSB stable by management regulations

*Reversibility of danger*: To some extent, but difficult to predict quantitatively. Irreversible situations may occur

*Possibility of reducing uncertainty*: To some extent, e.g. improve methods, monitoring technology, catch control (improve presentation and form of advice)

---

decisions about thresholds are rarely unequivocal. A sensitivity analysis should indicate the robustness in the choice of threshold.

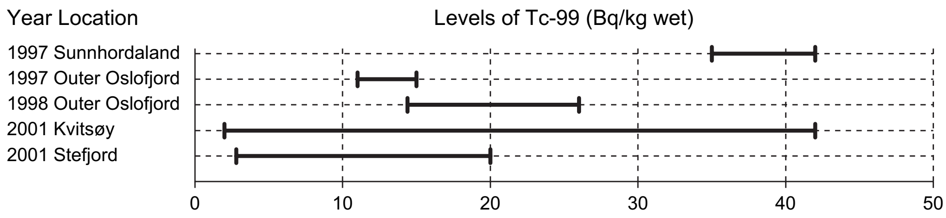
In Performance Perspectives (d), we address the applicability of the indicator in the management and decision-making context in a general way, by indicating the ability to adjust indicator level by managing human activities. It may be possible to adjust SSB by management measures, but it is not possible to manage the fisheries so that SSB does not fluctuate (Table 1).

The risk of irreversibility is an important aspect, because it relates to the threshold chosen. As long as the indicator is kept

Table 2. Framework example: technetium-99 in lobster (based on Brown *et al.*, 1998; Kolstad and Rudjord, 2000; Kolstad and Lind, 2002).

a) **Advice statement**

*Indicator time-series*:




*Indicator function*: SSB: Tc-99 level in lobster a proxy for level in exploited marine organisms in Norwegian coastal waters. Indicator used for monitoring purposes. Concern: health threat by consuming seafood from these areas

*Threshold*(*s*): EU: action level for prospective nuclear accident of 1 250 Bq $kg^{-1}$. Consuming food with higher concentrations involves possible health threat. Special dietary advice given in emergency situation

*Supporting information*: Discharges temporarily stopped. Among marine organisms consumed, lobsters have highest concentrations. However, higher concentrations are found in seaweed, and time-series is longer (Dahlgaard *et al.*, 1997)

*Conclusion/advice*: Concentrations well below threshold. Considered safe to eat seafood from Norwegian waters

b) **Power of explanation**

*Cause/effect in indicator*: For 10 years, Sellafield, UK, has been the only significant source. Concentrations and time lags depend on amount discharged and ocean currents. Other sources have minor effects

*Cause/effect in threshold*: Based on epidemic studies and general knowledge of physical properties of beta emitters

*Ability to detect change*: Changes of more than 5−10% detectable

*Ability to separate effects*: Other sources than Sellafield negligible (except in case of an accident on other nuclear power plants)

c) **Robustness**

*Underlying assumptions indicator*: Based on direct measurements. Underlying assumptions only linked to correctness of measurements

*Sensitivity analysis of indicator*: Not relevant

*Underlying assumptions in threshold*: Authors suppose that threshold is based on epidemic studies to identify concentrations that are dangerous for human consumption. Underlying assumptions also relate to knowledge of physical properties of beta emitters

*Sensitivity analysis of threshold*: Unknown to authors whether sensitivity analysis has been carried out

d) **Performance perspectives**

*Ability to adjust indicator level*: Reductions in discharges lead to reduction of concentration in lobster. Background concentration of Tc-99 in seawater fixed

*Reversibility of danger*: Probably reversible in marine organisms. Background concentration in seawater irreversible (Dahlgaard *et al.*, 1995)

*Possibility of reducing uncertainty*: Measurement uncertainty sufficiently low. Non-quantifiable uncertainty linked to accidents and other unforeseen events

away from the threshold, we deal (it is hoped) with reversible states. Tc-99 concentrations in lobster may be reversible, but not beyond the background radiation (Table 2). Adding related dangers will make the picture more complete.

Managers may want to know whether further research or additional control measures can help to reduce uncertainty. For instance, existing data on the bycatch rate of harbour porpoise are poor (Table 3), so uncertainty may be reduced by collecting appropriate information. In fish stock assessment, the gain of further research may not be so obvious.

## Discussion

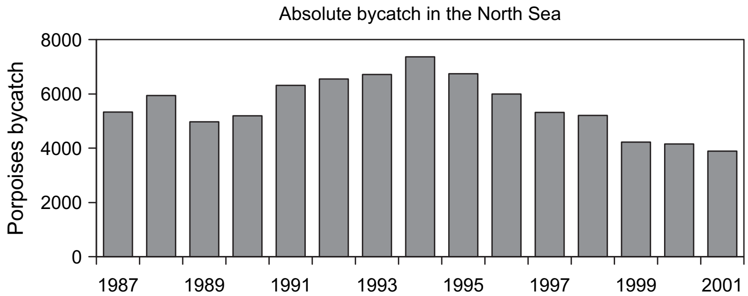
Application of the proposed framework to the three case studies of environmental indicators helped us formulate the information categories that should be included. The framework adds perspectives for enhanced transparency, compared with what is usually provided (e.g. ICES, 2003). For SSB (Table 1), for example, both the sources of uncertainty and the lack of a comprehensive uncertainty assessment are spelled out, allowing the user either to accept the uncertainty, to suggest alternative ways of producing, presenting or evaluating the advice, or to investigate alternative management strategies.

The framework manages to reflect fundamental differences in uncertainty among the three indicators. Determining Tc-99 concentrations introduces little uncertainty other than a low measurement error, whereas the level of SSB rests on a myriad of assumptions, all associated with more or less uncertain scientific

Table 3. Framework example: bycatch of harbour porpoise.

a) **Advice statement**

*Indicator time-series*: Bycatch rate of harbour porpoise in all North Sea fisheries, in absolute numbers and expressed as a percentage of total population size (ICES, 2002a)




Population estimates based on surveys in 1994/1995, and bycatch rate therefore only estimated for those 2 years (ICES, 2002a).

| Year | Population size | Bycatch rate | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|
| 1994 | 348 044 | 2.1% | 3.2% | 1.3% |
| 1995 | 348 044 | 1.9% | 3.0% | 1.2% |

*Indicator function*: monitoring; porpoises are entangled in gear, and most die from drowning; concern is population decline

*Threshold*(*s*): annual bycatch of 1.7% of best population estimate; higher rates can lead to population decline (Kock and Benke, 1996; Tregenza *et al.*, 1997) to <80% of carrying capacity (ICES, 2001)

*Supporting information*: stranding of dead porpoises and other indications of high and unreported bycatch; data quality poor, especially for current population size

*Conclusion/advice*: bycatch higher than threshold, especially for specific North Sea areas; advice has been to identify those areas and to suggest mitigating measures; no rules in place to determine action at high bycatch rates

b) **Power of explanation**

*Cause/effect in indicator*: Indicator value depends on fishing. Data quality regarded as limited. Incomplete reporting may cause artificial trends

*Cause/effect in threshold*: expected population decline dependent on correct estimate of population growth rate

*Ability to detect change*: to some extent; depends on quality of reported data

*Ability to separate effects*: may be difficult to separate effects of limited reporting and actual bycatch rate

c) **Robustness**

*Underlying assumptions indicator*: (i) annual bycatch data correct; (ii) annual estimates of population size correct

*Sensitivity analysis of indicator*: Unknown, but would have been relevant

*Underlying assumptions in threshold*: (i) maximum population growth rate 4%; sustainable maximum bycatch rate 1.7% (ICES, 2001); (ii) other mortality factors estimated correctly (ICES, 2001); (iii) carrying capacity estimated correctly (ICES, 2001); (iv) knowledge of separation of stocks (Tolley *et al.*, 1999)

*Sensitivity analysis of threshold*: Unknown, but would have been relevant. N.B. Fundamental question is if goal of reaching 80% (ICES, 2001) of carrying capacity is realistic and can be integrated with other goals in ecosystem-based management. One may not expect a decimated population of top predators to return to its historical carrying capacity given increased human impact through fishing, shipping, recreation, pollution, etc. Rather, population goals should be set in agreement with other ecosystem aspects, and existing human activities

d) **Performance perspectives**

*Ability to adjust indicator level*: Yes, through regulations of fishing effort and technical regulations (Trippel *et al.*, 1999)

*Reversibility of danger*: Reversible at "normal" population levels; at low levels, bycatch at or above the threshold may push population to extinction

*Possibility of reducing uncertainty*: Yes. Lack of data; by initiating comprehensive monitoring schemes and conducting regular and comprehensive assessment and studies of demography

judgements. In the case of TC-99, the main uncertainty may be related to the threshold, but we have not been able to find the document describing its basis. This indicator is also associated with indeterminacy, because future levels depend mainly on human behaviour (political decisions on production or unforeseen future accidents). The harbour porpoise example suggests that the main problem is data quality. Reported bycatch numbers are not considered reliable and the last surveys date back to 1994 (ICES, 2002a). Because the question on stock separation does not seem to be settled (Tolley *et al.*, 1999), non-quantifiable uncertainty is a major part of total uncertainty in this case.

A detailed and systematic description of indicator qualities has many applications in ecosystem-based management. First, the framework may increase awareness among scientists of the scope for enhancing the relevance of their advice by calling attention to different uncertainty perspectives that they may not be used to (or trained to) communicate, or upon which to reflect. Furthermore, the framework may help to open up the black box often underlying scientific advice based on indicators, by facilitating judgement about their relevance, uncertainties, limits, and strengths. For instance, in selecting a suite of indicators to measure the state of an ecosystem, decision-makers have numerous time-series from many fields of science from which to select. Because non-experts may have limited insight into how well an indicator reflects the issue and how well-founded the knowledge in this field is, the transparency provided by the framework may facilitate the selection process. Ultimately, this should make it easier to find effective management strategies.

At this stage, the framework should be regarded as a guide to what features to address, rather than a rigid scheme of boxes and short texts. Our examples have not been worked up in full detail because their main purpose was to demonstrate the approach. In fact, it was not always clear how to allocate information to each of the boxes. More case studies on completely different indicators should help to uncover potential weaknesses and to make further improvements. However, because the framework is about communicating knowledge and advice and raising awareness among all parties, it eventually needs to be evaluated by its users.

## References


Brown, J., Kolstad, A. K., Lind, B., Rudjord, A. L., and Strand, P. 1998. Technetium-99 contamination in the North Sea and in Norwegian coastal areas, 1996 and 1997. StrålevernRapport, 1998:3. Norwegian Radiation Protection Authority, Østerås, Norway.

Dahlgaard, H., Bergan, T. D. S., and Christensen, G. 1997. Technetium-99 and caesium-137 time series at the Norwegian coast monitored by the brown alga *Fucus vesiculosus*. Radioprotection-Colloques, 32(C2): 353–358.

Dahlgaard, H., Chen, Q., Herrmann, J., Nies, H., Ibbett, R. D., and Kershaw, P. J. 1995. On the background level of $^{99}$Tc, $^{90}$Sr and $^{137}$Cs in the North Atlantic. Journal of Marine Systems, 6: 571–578.

Directorate of Fisheries. 2003. Forslag til program for utredning av konsekvenser av fiskeri I området Lofoten-Barentshavet. March 2003, Bergen, Norway. 22 pp.

EEA. 2002. Testing of indicators for the marine and coastal environment in Europe. 1. Eutrophication and integrated coastal zone management. European Environment Agency Technical Report, 84. 48 pp.

EEA. 2003. Testing of indicators for the marine and coastal environment in Europe. 3. Present state and development of indicators for eutrophication, hazardous substances, oil and ecological quality. European Environment Agency Technical Report, 86. 64 pp.

FAO. 1999. Indicators for sustainable development of marine capture fisheries. FAO Technical Guidelines for Responsible Fisheries, 8. 68 pp.

Funtowicz, S. O., and Ravetz, J. R. 1990. Uncertainty and quality in science for policy. Kluwer Academic, Dordrecht. 229 pp.

ICES. 1999. Report of the Coordinating Group on ICES Advice. ICES Document, CM 1999/Del: 21. 40 pp.

ICES. 2001. Report of the ICES Advisory Committee on Ecosystems, 2001. ICES Cooperative Research Report, 249. 75 pp.

ICES. 2002a. Report of the ICES Advisory Committee on Ecosystems, 2002. ICES Cooperative Research Report, 254: 129 pp.

ICES. 2002b. The ICES Strategic Plan. International Council for the Exploration of the Sea, Copenhagen.

ICES. 2003. Report of the ICES Advisory Committee on Fishery Management, 2003. ICES Cooperative Research Report, 261: 975 pp.

Kock, K., and Benke, H. 1996. On the bycatch of harbour porpoise (*Phocoena phocoena*) in German fisheries in the Baltic and the North Sea. Archive of Fishery and Marine Research, 44: 95–114.

Kolstad, A. K., and Lind, B. 2002. Radioactivity in the Marine Environment 2000 and 2001. Technetium-99 concentrations in Norwegian coastal waters and biota. StrålevernRapport, 2002:6. Norwegian Radiation Protection Authority, Østerås, Norway.

Kolstad, A. K., and Rudjord, A. L. 2000. Radioaktiv forurensning i marint miljø. $^{99}$Tc i sjøvann og biota fra norske kystområder i 1998 og 1999. StrålevernRapport, 2000:7 (in Norwegian). Norwegian Radiation Protection Authority, Østerås, Norway.

Tolley, K. A., Rosel, P. E., Walton, M., Bjørge, A., and Øien, N. 1999. Genetic population structure of harbour porpoises (*Phocoena phocoena*) in the North Sea and Norwegian waters. Journal of Cetacean Research and Management, 1: 256–274.

Tregenza, N. J. C., Berrow, S. D., Hammond, P. S., and Leaper, R. 1997. Harbour porpoise *(Phocoena phocoena)* bycatch in set gillnets in the Celtic Sea. ICES Journal of Marine Science, 54: 896–904.

Trippel, E. A., Strong, M. B., Terhune, J. M., and Conway, J. D. 1999. Mitigation of harbour porpoise (*Phocoena phocoena*) bycatch in the gillnet fishery in the lower Bay of Fundy. Canadian Journal of Fisheries and Aquatic Sciences, 56: 113–123.

Van der Sluijs, J. P., Potting, J., Risbey, J., van Vuuren, D., de Vries, B., Beusen, A., Heuberger, P., Corral Quintana, S., Funtowicz, S., Kloprogge, P., Nuijten, Petersem, A., and Ravetz, J. 2002. Uncertainty assessment of the IMAGE/TIMER B1 $CO_2$ emissions scenario, using the NUSAP method. Dutch National Research Program on Climate Change, Report 410 200 104. 227 pp. Available from www.nusap.net.