Fisheries Research 168 (2015) 49–55

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

How much evidence is required for acceptance of productivity regime shifts in fish stock assessments: Are we letting managers off the hook?

Neil L. Klaer[a,*], Robert N. O'Boyle[b], Jonathan J. Deroba[c], Sally E. Wayte[a], L. Richard Little[a], Larry A. Alade[c], Paul J. Rago[c]

[a] CSIRO Oceans and Atmosphere, PO Box 1538, Hobart, Tasmania, Australia
[b] Beta Scientific Consulting Inc., 1042 Shore Dr. Bedford, Nova Scotia, Canada B4A 2E5
[c] NOAA Fisheries Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543, USA

ARTICLE INFO

Article history:
Received 12 November 2014
Received in revised form 25 March 2015
Accepted 25 March 2015
Handling Editor A.E. Punt
Available online 20 April 2015

Keywords:
Productivity shift
Regime shift
Fisheries

ABSTRACT

A difficult question often confronting fisheries assessment scientists and managers is whether or not to accept that a shift in stock productivity has occurred. This is particularly the case when a stock has remained at historically low biomass despite management intervention and when there is an expectation that there should have been a stock recovery. We outline a weight-of-evidence approach that provides a structured means to evaluate this question. The approach, which scores a range of attributes, was applied to five fisheries from the NW Atlantic and SE Australia, chosen to provide a range of supporting evidence, as well as different potential causal mechanisms for a productivity shift. Given the resulting scores for the example stocks, and whether a productivity shift has been accepted for those stocks, a score of between 7 and 12 indicated a level required for acceptance of a productivity shift. The approach has highlighted areas of future research that would improve individual species scores. It is hoped that the paper will encourage a more systematic examination of potential stock productivity shifts in assessments than has hereto been the case.

1. Introduction

A question often confronting fisheries assessment scientists and managers is whether or not a shift in stock productivity has occurred. This is particularly the case when a stock has remained at historically low biomass despite management intervention and when there is an expectation that there should have been a stock recovery.

For the purposes of this paper, a productivity shift is defined as a change over time in the biological characteristics of a fish stock that would lead to a change in biological reference points (such as maximum sustainable yield). In estimating biological reference points, it is often assumed that natural mortality, length-at-age, length–weight relationship, maturity-at-age/-length and the relationship of recruitment to spawning stock biomass are constant through time. A substantial temporal change in any of these factors would cause what we call a productivity or regime shift in the stock. Our focus is thus how best to interpret potential productivity shifts from a stock assessment and management advice viewpoint.

The decision on whether or not there has been a change in stock productivity is a difficult one. Accepting that there has been a shift in productivity moves the cause of the (usually) low biomass away from fishing to an external cause such as unfavourable environmental conditions. The responsibility for low stock status is thus removed from fisheries management as the cause is out of management control, and management is therefore "off the hook". On the other hand, if the stock biomass is low and productivity has not changed, there can be severe consequences for the future yield prospects of the resource – possible fishery closure or severe restriction of fishery effort.

Because of the potentially significant effect on biological reference points, a productivity shift might be described as an "extraordinary claim" and, as Sagan (1980) put it, "extraordinary claims require extraordinary evidence." Unfortunately, in our experience, decisions on whether or not there has been a productivity shift that should be taken into account in a stock assessment are often not based upon a systematic review of the available evidence but rather on expert opinion, which can be influenced by preconceptions. We contend that as acceptance of a productivity shift can have a very a profound influence on stock status and management

---

* Corresponding author. Tel.: +61 419107625; fax: +61 362325000.
E-mail address: neil.klaer@csiro.au (N.L. Klaer).

http://dx.doi.org/10.1016/j.fishres.2015.03.021

responsibility, an evidence-based approach is required to justify such acceptance.

A weight-of-evidence approach allows qualitative and semi-quantitative rating and assessment of the available scientific evidence in relation to some causal hypothesis or hypotheses (Krimsky, 2005). Importantly, hypotheses are articulated prior to the evaluation and then the evidence for – and against – each is evaluated. Note that this is similar to the Information–Theoretic approach used in quantitative model selection (Burnham and Anderson, 1998). Indeed, statistical model fitting involves the selection of the hypothesized processes which maximize the probability that the selected model gave rise to the observed data. However, in many assessments (if not most), models have not been developed that allow full quantitative exploration of the evidence for competing hypotheses. The default position has been that processes not included in a model are not considered. We suggest that it is prudent and necessary to evaluate support for competing hypotheses using a qualitative approach until such time as models can be developed to do so.

2. Methods

A weight-of-evidence approach can be used to decide whether a species has undergone a productivity shift that needs to be taken into account in a stock assessment. An advantage of the approach is that it provides an objective basis for what would otherwise be a subjective decision. A disadvantage is that the actual mechanism used to assign weightings to components and arrive at a decision may not be transparent (Krimsky, 2005). It is therefore important to test the proposed process(es) against verifiable examples.

We developed four criteria that we believe should be used to decide if there has been a productivity shift in a stock or population. These criteria are described below. Judgement using a weight-of-evidence approach is facilitated by assignment of a numerical score against each criterion.

2.1. Criteria for judgement of productivity shift

2.1.1. Criterion 1. Observed change in a productivity indicator

A productivity indicator is an observation over time of change in some measure of the stock that potentially provides evidence for a change in productivity regardless of the level of fishing pressure. Such indicators may include recruitment estimates from egg, larval or young-of-the-year surveys, biomass estimates from fishery catch rates (CPUE) or fishery-independent surveys, or evidence of changing natural mortality from multi-species diet studies or observed fish kills.

A long period of change in an indicator such as available biomass despite management intervention, is usually the cause for first consideration of a productivity shift. Some fish stocks show more obvious signs of a productivity shift than others, so the degree of change in the relative level of the indicator should also be considered in the overall weight of evidence. Fish stocks often only receive attention when the observed shift is to an apparent state of lower productivity. Of course, the opposite is also possible and should also receive critical evaluation.

A long period is probably best interpreted from a stock management perspective. This would be a period of sufficient length for evidence of a disconnection between management expectations and the response of the fish stock. Often this would be related to the average generation time or maximum age of a species. For example, a high score would be given if fishing pressure was reduced to negligible levels on a stock for a period of multiple generations, but no increase in biomass has been observed with a high degree of precision (e.g. through intensive fishery independent surveys).

2.1.2. Criterion 2. Understanding of assessment model input data

This criterion applies to the quality of the observations on which a stock assessment model is based. Fundamental biological characteristics such as age-/length-at-maturity, the length–weight relationship, differences in growth/reproduction by sex are normally gained via targeted biological studies. Uncertainty in such information should be considered under this criterion.

In addition, for many stocks, there is uncertainty about spatial stock boundaries that affects many observation uncertainties. Catch history may also be uncertain due to lack of records, estimation of total catch based on discard mortality rates, or difficulty in separating similar species in commercial landings. Similarly, abundance indices may be noisy due to low sampling levels, show inconsistent trends, or are possibly biased if survey methodology has changed or sampling has occurred at the margins of the spatial or depth distribution of the stock. Similar sources of error also apply to sampling of age or length composition.

Low scores of this criterion apply when there is substantial uncertainty in the biology, total fishery catch, recent levels of fishing mortality, whether abundance indices are likely to be good indicators of true population abundance, or whether age/size sampling has been representative of the population.

If it is not possible to resolve such uncertainties, the score can be increased through the development of plausible ranges of alternative assessment inputs, at least for model sensitivity testing.

2.1.3. Criterion 3. Understanding of assessment model structural assumptions

This criterion applies to the extent it can be determined that an apparent time-varying shift in productivity is not a product of the structural assumptions of the assessment model. It is important to determine whether an apparent change in a parameter through time reflects an actual shift rather than the application of an inappropriate average relationship. For example, an inappropriate relationship might occur if recruitment is assumed to remain constant at all biomass levels when a model fit to data indicates a recruitment pattern better characterised by a Beverton and Holt (1956) stock–recruitment relationship. In a less extreme case, the model may assume a certain fixed value for Beverton and Holt steepness that results in apparent time-shifts in average recruitment residuals, which can be corrected using a different steepness value.

Not accounting for substantial shifts in fishery selectivity (e.g. dome shaped to logistic) over time may contribute to a perception of productivity shift, and should be closely examined if this is a possibility. Conversely, the introduction of a substantial selectivity change to an assessment model requires close scrutiny of the supporting evidence.

The lowest score level should result from the simple display by the model of an apparent change in the average of an important productivity parameter over a period of time. A higher score applies if the addition of time variation in a parameter is justified statistically (e.g. via Akaike information criterion or removal of a retrospective pattern), if alternative model structures that do not require time variation were considered and excluded, and if other possible sources of productivity change were also investigated. A more complete list of possible time-varying productivity parameters for a single species population model would include recruitment, fishery selectivity, natural mortality, growth, age/length at maturity, and fecundity.

Currently, most integrated assessment models routinely only allow for annual variation in recruitment, so this parameter is most often identified in stock assessments as potentially showing long periods of average change. As key population parameters are usually confounded, appropriate data are required to allow the assessment model to separately estimate time trends in several of these parameters simultaneously. However, if sufficient justification is provided and appropriate data are available, exploration of changes in other normally time-invariant parameters should be pursued whenever possible in stock assessments.

2.1.4. Criterion 4. Explanatory hypothesis

Have environmental and/or ecological studies provided independent evidence of a process that would support a hypothesised change in productivity? Are there indications from other species in the same system that point to the same hypothesis? At the simplest level, correlations of population parameters and environmental changes can be quantified (e.g. Blamey et al., 2012). The highest score would apply in cases where the hypothesised change was derived from a data-rich environment-coupled ecosystem model that produced abundance information across a wide range of important species that matches with indices from well-designed and applied field surveys. More than one plausible hypothesis for a productivity change may exist, and the evidence for and against each should be compiled and examined.

Scoring of the above four criteria should be made in a way that allows cross-species comparisons, so some guidelines are required for how the scores are derived. To assist in conducting these comparisons, interpretation of the criteria above was converted to explicit scoring guidelines in Table 1.

In this study, we applied these criteria to a number of example fish stocks that have shown potential shifts in productivity:

2.2. Example fish stocks

2.2.1. Southern Gulf of St. Lawrence Atlantic cod (Gadus morhua, Gadidae)

Atlantic cod are distributed widely in coastal areas across the eastern and western North Atlantic Ocean. They are mainly confined to depths of 150–200 m on the continental shelf, but can be found at depths of more than 600 m. Twelve separate stocks are recognised for management purposes (Shelton et al., 2006), one of which resides in the southern Gulf of St. Lawrence. The growth of southern Gulf cod has changed considerably over time, with the mean weight of a 7-year-old reducing from 2.5 kg in the 1960s to less than 1.0 kg by 2010 (Swain et al., 2012). In the 1950s, female cod generally matured at 6–7 years of age, but now do so between ages 4 and 5. Cod typically live up to 20 years but there are now few cod older than age 10.

Excessive fishing pressure was the main reason for the decline of the northwest Atlantic cod stocks to the mid-1990s, but the stocks have failed to recover since the introduction of severe catch restrictions (Shelton et al., 2006). The southern Gulf cod stock is currently at the lowest level observed in the 61-year record and is still declining. Abundance of mature cod during 2008–2010 is estimated to be 37% of the average during the mid to late 1990s and 10% of the average in the mid-1980s (Swain et al., 2012). The stock has been well studied, and there has been general acceptance that while overfishing caused the stock decline, a reduction of population productivity is the main factor contributing to the lack of recovery (Swain and Chouinard, 2008; Swain et al., 2012). Reasons proposed for the unexpectedly low recovery rates both in the southern Gulf and elsewhere are: increased natural mortality, decreased body growth, reduced recruitment rates in some stocks, and possibly the effects of continued directed and bycatch fishing (Shelton et al., 2006). The most important factor affecting recovery is elevated natural mortality, estimated for different time periods using the method of Sinclair (2001) applied to age-disaggregated survey indices (Shelton et al., 2006) as well as within an ADAPT assessment model (Swain et al., 2012).

2.2.2. Atlantic Herring (Clupea harengus, Clupeidae)

Herring in the Gulf of Maine/Georges Bank area have a complex stock structure with fish originating from several separate spawning aggregations (Overholtz et al., 2004). Herring from each aggregation are mixed in unknown proportions as they migrate from southern overwintering areas to northern summer feeding grounds, and back (Overholtz et al., 2004). Fish from each spawning aggregation only segregate during spawning activity in the fall. Subsequently, fishery and survey catches are largely composed of fish from multiple spawning aggregations, which precludes assessing each spawning aggregation independently (NEFSC, 2012).

Herring growth in the region declined during the late 1980s and early 1990s. For example, the mean weight of an age-5 fish during 1964–1985 was 0.22 kg, but was 0.15 kg during 1995–2010 (NEFSC, 2012). The age at 50% maturity is between 2 and 3 years of age with 100% maturity at ages 4–5 (NEFSC, 2012).

The Georges Bank component of the herring stock collapsed in the early 1980s, largely due to overfishing. The stock complex has since recovered and landings have been relatively stable at approximately 100,000 mt for ~15 years (NEFSC, 2012).

Table 1
Scoring guidelines.

Score	Observed change in a productivity indicator	Understanding of assessment model input data	Understanding of assessment model structural assumptions	Explanatory hypothesis
0	Short period less than one generation	Model input uncertainties are unknown	Key population parameters affected have not been identified	The mechanism is unknown
1	More than one generation	A number of model inputs are uncertain and the extent of uncertainty has not been characterised	Modeled changes in one or more key population parameters have fitted with observed biomass changes	A plausible mechanism for productivity shift has been developed from general knowledge of biophysical processes
2	Multiple generations and across several assessment/management cycles	Uncertain model inputs have been characterised and plausible ranges for those uncertainties have been investigated	Modeled changes in key production parameters have been somewhat validated by investigation of alternative model structures and/or improved model behaviour such as the removal of retrospective patterns	Output from a limited biophysical or multispecies model is consistent with observed patterns of change in productivity
3	Multiple generations and across many regular assessment/management cycles in the same timeframe	The character of model inputs is well understood and uncertainty has largely been eliminated or well estimated statistically	Validated modeled changes are consistent with output from a biophysical or multispecies model	Output from a comprehensive biophysical multispecies model is consistent with observed patterns of change in productivity

A recent assessment of the herring stock (NEFSC, 2012) included an increase in natural mortality starting in 1996 and continuing through 2011 (the terminal year in the recent assessment). This increase in natural mortality resolved a retrospective pattern, but estimates of consumption of Atlantic herring were also used to inform this decision. Consumption estimates are not usually considered as a mechanism for adjusting natural mortality, but the treatment in this case appears justified by the annual availability of stomach contents for a wide range of herring predators (e.g. fish, mammals, birds) and annual abundance estimates of these predators. The consumption estimates, however, were relatively imprecise, and much of this imprecision was caused by low stomach sampling rates for some predators and poor spatial sampling coverage, rather than predator abundance estimates.

Thus, the modelled increase in natural mortality was an uncertain decision with strong implications for the expected long-term productivity of the stock. For example, maximum sustainable yield (MSY) from the assessment model with a time-varying increase in natural mortality was 45,000 mt and the fishing mortality rate in 2011 was 70% of the fishing mortality rate at MSY, while those values from an assessment model with time-invariant natural mortality were 122,000 mt and 35%, respectively. Although both models suggest fishing mortality rates below that associated with MSY, this convenience should not detract from the result that the two models provide different perceptions of relative stock status and yield outlooks for the fishery. Management decisions and perceptions of success are therefore dependent on the modelled change in natural mortality.

*2.2.3. Southern New England/Mid-Atlantic Bight yellowtail flounder (*Limanda ferruginea*, Pleuronectidae)*

The southern New England/Mid Atlantic (SNEMA) yellowtail flounder stock represents the southern-most range of the yellowtail flounder stock off the northwest Atlantic. Historically, the stock has supported a substantial fishery dating back to the early 1940s (post-WWII), and in the late 1960s when catches of yellowtail were dominated by distant-water fleets. Average catches in the late 1960s averaged over 38,000 mt, peaking to over 44,000 mt in 1969. Despite the elimination of the distant-water fisheries in the early 1970s, followed by management restrictions on catches in the mid-1990s, spawning stock biomass has been very low since 1991, with recruitment below the long-term average. The spatial range of stock biomass has also experienced changes in distribution, contracted from that associated with large recruitments observed during the 1970s and 1980s to being concentrated at very low abundances at the northern range of the stock.

The North Atlantic Oscillation and its effects on stratification of the water column and temperature were identified as potential reasons that the stock has not recovered (Sullivan, 2005). In the most recent assessment (NEFSC, 2012), the mid-Atlantic cold pool, a seasonal cold water mass that forms in the region, was modelled as a possible mechanism for reduced recruitment. However, the mechanistic effects of such oceanographic processes on recruitment success in yellowtail flounder are still poorly understood, and an environmental series that can reliably explain the recruitment pattern remains unknown. Hence, biomass reference points and stock status are highly influenced by the period of recruitment used to forecast the stock into the future; that is, conducting forecasts including historical recruitments that are relatively high or only more recent recruitment that are relatively low. In spite of management's effort to reduce fishing pressure, the stock has not recovered to historical levels, so it is suggested that broad ecosystem changes may be responsible for the apparent observed shifts in stock productivity.

*2.2.4. SE Australia jackass morwong (*Nemadactylus macropterus*, Cheilodactylidae)*

Jackass morwong is a moderately long-lived (up to 30 years) demersal fish species that occurs in the continental shelf waters of the southern hemisphere. Stock abundance in south-eastern Australia has declined recently, despite lower catches than in the past. This decline is attributed to mostly below-average recruitment since 1985; detailed examination of the data inputs and assessment diagnostics has confirmed that this pattern does not appear to be an artefact of the model fit to the data (Wayte, 2013). Abundance index data derives from commercial CPUE, and observer sampling that provides information on length and age of the catch, but generally covers less than 5% of fishing operations annually. Lower steepness of the stock–recruitment relationship was considered unlikely, and cannot account for the recruitment decline.

The unusual early life history of jackass morwong makes it likely that recruitment is environmentally driven. The species is atypical among temperate finfish species in that it has an extended offshore pelagic post-larval stage, spending 9–12 months in the offshore surface waters of south-eastern Australia (Vooren, 1972; Bruce et al., 2001). The observed southward shift of the ocean current system off south-eastern Australia (Ridgway, 2007) may have reduced the suitability of the offshore environment for survival of post-larval jackass morwong. The enrichment, concentration, and retention processes that combine to provide favorable reproductive habitat for many types of fish (Bakun, 1998) have all likely become weaker in the oceans off south-eastern Australia.

In 2011, based on an informal application of the procedure outlined in this paper, and the results of a management strategy evaluation, Australian fisheries managers accepted a shift in jackass morwong productivity that had immediate implications for the perceived status of the stock, as well as longer term implications for future catches. The new stock assessment models two productivity regimes: the first from 1915 when the fishery commenced; and the second, lower productivity, regime from 1988. This new assessment gave a more optimistic view of current stock status than the previous assessment (35% of unfished as compared to 23%), as current biomass was compared to a lower unfished equilibrium biomass. The new assessment also gave a higher Recommended Biological Catch (RBC) for the following year, and led to a change in the overfishing classification of the stock. Previously, jackass morwong had been classified by fishery managers as 'subject to overfishing', because current catches were greater than the RBC estimated by the assessment, but with the new assessment, this was no longer the case. The new assessment estimated the long-term RBC under the lower productivity regime to be less than half that of the long-term RBC estimated by the previous assessment.

*2.2.5. SE Australia eastern gemfish (*Rexea solandri*, Gempylidae)*

Eastern gemfish is a midwater species caught along the edge of the continental shelf off southern Australia and New Zealand. It is a carnivorous species, feeding mainly on fish such as family macrouridae (whiptails) and deepwater cardinalfish (*Apogonops anomalus*, Acropomatidae), as well as royal red prawns (*Haliporoides sibogae*, Solenoceridae) and squid (Pogonoski et al., 2002). Eastern gemfish have a moderately fast growth rate, reaching 50 cm in length and 1 kg in weight after 3 years.

During the 1970s and 1980s, eastern gemfish comprised a large proportion of trawl landings off south-eastern Australia, with trawling targeting winter pre-spawning aggregations of mature fish. Catches peaked in 1980 at 5000 mt. Fishing effort actively targeted the pre-spawning migration, when large aggregations of the species were predictable, and there were concerns that the CPUE index of the pre-spawning aggregations was hyperstable. Observer sampling that provides information on length and age of the catch generally covers less than 5% of fishing operations annually. Catches fluctuated during the early 1980s, but the stock experienced recruitment failure in the late 1980s.

Since the early 1990s, the spawning stock has declined substantially because of a series of very poor cohorts (Punt and Smith, 1999), and the TAC was progressively reduced and was set at zero in 1993, where it has remained. A number of possible causes for low recruitment have been suggested including environmental factors such as changing currents and winds due to climate change, depression of larval production due to a threshold biomass being passed, reproductive regime shift, and disruption of spawning behaviours by fishing activity (Little and Rowling, 2008). However, the reasons for poor recruitment are not known, as the life cycle stages at which the disruptive processes are operating have not been identified.

3. Results

From a maximum possible score of 12, scores for the five example stocks ranged from 4 to 10 (Table 2). This highlights that even for a stock such as southern Gulf of St. Lawrence Atlantic cod, where a productivity shift has been generally accepted for some time, there are still uncertainties related to model structural assumptions and incomplete understanding of influential ecosystem processes.

All example stocks scored 2 or 3 for the observed change in a productivity indicator, as would be expected for stocks where a possible productivity shift has been suspected. On the understanding of assessment input data, there was considerable variation among stocks, with the lowest scores for those with uncertain stock boundaries. No example stock scored at the highest level for understanding of assessment model structural assumptions as none have been validated against output from a biophysical or multispecies model for productivity change. The lack of a comprehensive biophysical multispecies model that is consistent with observed productivity changes also explains why no example stock scored at the highest level under the explanatory hypothesis criterion.

Scores for the example stocks match reasonably well with current management acceptance of productivity shift. Gemfish is an example stock where a productivity change has not yet been accepted and stock reference points have not been revised for management purposes. In the case of yellowtail flounder, the current management acceptance of stock productivity shift is in contrast with the low weight of evidence score in Table 2. Although the stock reference points for yellowtail flounder are uncertain, the stock is now considered to be rebuilt (NEFMC, 2014). Given the scores for the example stocks, and whether productivity shift has been accepted for those stocks, a score of about 7 is (empirically) indicative of the level required for acceptance of a productivity shift.

4. Discussion

An informal application of the procedure outlined here was used to evaluate the likelihood of a productivity change for jackass morwong in Australia in 2011 (Wayte, 2013). A similar process was also applied during an independent review of the Atlantic herring and SNEMA yellowtail flounder stock assessments conducted by the Northeast Fisheries Science Center (O'Boyle et al., 2012). For these two stocks, the independent peer review panel accepted that a change in natural mortality was justified in the assessment of Atlantic herring, and suggested there was a 60/40 likelihood that there had been a shift in average recruitment in the SNEMA yellowtail flounder stock. For southern Gulf cod, there is general acceptance that elevated natural mortality has been the cause for a lack of stock recovery (Swain et al., 2012). As mentioned above, the acceptance of a shift in each case had strong implications for stock status and management response.

An important outcome of the scoring process is highlighting of research required to improve the score for a particular stock. For example, work is required to improve characterisation of the uncertainties in input data for yellowtail flounder, and to closely examine

Table 2
Summary weight of evidence per example species.

Stock	Observed change in a productivity indicator	Understanding of assessment model input data	Understanding of assessment model structural assumptions	Explanatory hypothesis	Weight of evidence
Southern Gulf of St. Lawrence Atlantic cod	18 years of low biomass (3)	Stock biology well characterized. Catch and its composition is well known and the survey exhibits low CVs (3)	Modeled values of total mortality suggest that natural mortality has increased for a number of sub-stocks (2)	Increase in natural mortality is mostly attributed to increased predation by seals, but no direct measurement from field studies (2)	10
Gulf of Maine/Georges Bank Atlantic herring stock complex	16 years of relatively high natural mortality (2)	Some uncertainty in stock boundaries and catch levels that have not been characterised (1)	A modeled increase in natural mortality since 1995 eliminates a strong retrospective pattern and better matches consumption estimates (2)	A multispecies model that estimates herring consumption based on a long time series of stomach contents across a range of species was used to inform the *M* increase hypothesis (2)	7
SNE/Mid-Atlantic yellowtail flounder	21 years of low biomass (2)	Some uncertainty in stock boundaries and catch levels that have not been characterised (1)	Modeled annual recruitment residuals below average for past 22 years (1)	Effect of oceanographic processes on recruitment is poorly understood. Environmental correlate with poor recruitment sought but not found (0)	4
Jackass morwong	12 years of low biomass (2)	Stock boundaries well defined, total catch well estimated and uncertainty in CPUE accounted for. Percentage of catch covered by size/age sampling is relatively low (2)	Modeled annual recruitment residuals below average for past 19 years. Fit to the data is improved by productivity shift; lower steepness cannot account for decline (2)	Some knowledge of larval distribution and recent changes in ocean circulation conditions lead to a plausible mechanism. Corroboration from some other species (1)	7
Gemfish	22 years of low biomass (3)	Stock boundaries well defined, total catch well estimated and uncertainty in CPUE accounted for. Percentage of catch covered by size/age sampling is relatively low (2)	Modeled annual recruitment residuals generally below average for past 24 years (1)	Larval biology poorly understood. Mechanism for poor recruitment not understood (0)	6

assessment model structural assumptions for yellowtail flounder and gemfish.

Scoring as outlined here, and the suggested level for acceptance of productivity shift assumes that the weighting of each criterion is equal, and that the scores against each criterion were objectively evaluated. The primary purpose of applying the approach is to ensure that all sources of available evidence for a possible productivity shift have been examined. The weightings are secondary and in reality would differ among cases, although equal weighting is a useful starting point. It would be an improvement if a more detailed examination provided evidence for how the scoring should be adjusted. Objective evaluation of the criteria would preferably be made by a group that is uninfluenced by the outcome of the decision but knowledgeable about available forms of evidence for productivity shift, such as an independent stock assessment review panel. Such a panel would ideally represent a variety of institutions (governmental, non-governmental and academic).

In addition to scoring against the four suggested criteria, an important consideration is the evaluation of the risk of making an incorrect decision about a productivity shift. Such a risk evaluation can be made using management strategy evaluation (MSE: Smith et al., 1999; Bunnefeld et al., 2011), where the effects of assessment assumptions are tested using an operating model representing the 'true' underlying dynamics of the stock. The types of data usually collected by the fishery are generated from the 'true' population each year, and then used in a stock assessment within the MSE framework. The results of the assessment are used in decision rules to calculate management advice (e.g. future catch levels) that is introduced back into the operating model. This feedback loop is critical to evaluating assessment assumptions in terms of their effects on the 'true' population.

To perform a full productivity shift risk analysis, two versions of the operating model are required: one with the shift, and one without the shift. For each version of the operating model, projections are conducted with two assessments (or harvest control rules) that either incorporate the shift or not. Examination of performance measures from the MSE, such as 'true' stock status, the risk of the stock falling below specified levels, and future catch levels, can elucidate the trade-offs and risks associated with the assessment assumptions. For example, for the jackass morwong stock, Wayte (2013) used MSE to show that the consequences of mis-specifying the assessment model were riskier to the stock if the assessment model had assumed that no productivity shift had occurred.

Many harvest strategy policies implemented worldwide include reference points that assume that stocks will return to a fixed virgin biomass, B_0, on cessation of fishing. At present, operational stock assessments generally only account for the population dynamics of the species in question, and rarely account for multi-species ecosystem interactions. This is counter to our established knowledge about dynamic ecosystems. The notion that ecosystems undergo regular cycles and changes was first expressed in 1960 by Ramon Margalef: "Ecosystems result from the integration of populations of different species in a common environment. They rarely remain steady for long, and fluctuations lie in the very essence of the ecosystems and of every one of the . . . populations [that compose the system]" (Smith, 1994). It has been long recognised that many fisheries undergo multi-year and decadal fluctuations in recruitment (Steele, 1996), and that equilibrium productivity of wild fish stocks is a potentially misleading concept (Caddy and Seijo, 2005). A meta-analysis of 230 fish stocks by Vert-pre et al. (2013) found that about 40% of the stocks showed trends consistent with productivity shifts unrelated to stock abundance. A more recent meta-analysis (Szuwalski et al., in press) found that 61% of 224 stocks in the RAM Legacy Stock Assessment Database (Ricard et al., 2012) showed no positive relationship of recruitment and spawning biomass over the observed range of stock sizes, and that the environment appears to more strongly influence recruitment than spawning biomass for many stocks.

Managers should evaluate the type of harvest control rules used to define fishing mortality or catch levels when productivity shifts occur. For example, constant fishing mortality rate control rules have been shown to attain yields more similar to constant escapement control rules when productivity shifts occurred in the form of autocorrelation in the asymptote parameter of a Beverton–Holt stock recruit model, which was counter to when no such productivity changes occurred and constant escapement attained much higher yield (Parma, 1990; Walters and Parma, 1996). However, much of the research on this topic has focused exclusively on yield-based objectives and ignored errors in the assessment and management process (Deroba and Bence, 2008). The relative weight given to competing fishery objectives, the degree of compensation in the stock–recruitment relationship, and the consideration of errors in the assessment and management process all affect relative control rule performance (Deroba and Bence, 2008). How these performance-affecting factors interact in the presence of productivity shifts is unknown and is a topic for future research (Deroba and Bence, 2008). Similarly, most current fisheries management assumes that reference points, such as B_0 that are often used to define the shape of harvest control rules, are fixed through time. Productivity shifts will require re-estimation of B_0 and other reference points, and these estimates will contain error (Vert-pre et al., 2013). Szuwalski et al. (in press) suggest that significant periods of environmentally forced changes in average recruitment levels (climate regimes) need to be carefully considered for recalculation of stock management reference points, and that risk analysis be carried out on the acceptance of such regimes and their continuation into the future.

Within single species dynamics, a simple first step can be taken by calculating the "dynamic B_0" as first proposed by MacCall et al. (1985). Where annual deviations in life history characteristics are fitted by a model, it is possible to calculate the theoretical biomass that would have occurred in any historical year if there was no fishing. Some recognition of the dynamic nature of species productivity would be realized if the dynamic B_0 was used to set target and limit reference points for fisheries management. Simulation testing has demonstrated improvement in management performance of dynamic B_0 under productivity change (A'mar et al., 2009). However, all of the uncertainties outlined here, particularly in relation to criteria 2 and 3, would also apply, and should still be considered.

While current population status may better be described by the dynamic B_0 approach, there is a question about how to make future stock predictions after productivity change. Whether environmental factors and projections of them from climate models should be included in management strategies was reviewed by Punt et al. (2013). They concluded that the ability of the management strategy to achieve management goals is not usually improved unless the process for how the environmental factors affect the system is well known. In the normal situation where the process mechanism is not well understood, it is most appropriate to investigate the robustness of the management strategy to a broad plausible range of future change in biological parameters.

In a perfect world, fisheries management would be based on fish stock conditions derived from an integrated ecosystem model constructed using representative monitoring data for all important modelled processes. Operational fisheries management is still far from this utopia; complex ecosystem models are currently used for big picture, strategic direction-setting and conceptual management purposes, but not for short-term tactical management decision making, such as setting annual fishery catches (Plagányi et al., 2012). In addition, outputs from single species stock assessments are commonly used for the construction of the more complex ecosystem model, normally without consideration of uncertainty, biases and correlations in that assessment output (Brooks and Deroba, in press). For complex multi-species fisheries such as the Australian Southern and Eastern Scalefish and Shark Fishery, recovery of management costs according to the value of the fishery appears to rule out the intensive data collection that would be required to support more comprehensive system models that potentially could be used for tactical decisions.

In the medium-term, it is clear that single species population models and management approaches that assume fixed values for B_0 and other stock parameters will remain common, which is where the protocol in this paper could be applied.

Acknowledgements

The authors thank Tony Smith and Beth Fulton from CSIRO Oceans and Atmosphere, Fred Serchuk from the NOAA Northeast Fisheries Science Center, and two anonymous reviewers for their useful comments on earlier versions of this manuscript.

References

A'mar, Z.T., Punt, A.E., Dorn, M.W., 2009. The evaluation of two management strategies for the Gulf of Alaska walleye pollock fishery under climate change. ICES J. Mar. Sci. 66, 1614–1632.

Bakun, A., 1998. Ocean triads and radical interdecadal variation: bane and boon to scientific fisheries management. In: Pitcher, T.J., Hart, P.J.B., Pauly, D. (Eds.), Reinventing Fisheries Management. Kluwer Academic Publishers, Dordrecht, pp. 331–358.

Beverton, R.J.H., Holt, S.J., 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. Rapp. P.v. Réun. Cons. Int. Explor. Mer 140, 67–83.

Blamey, L.K., Howard, J.A.E., Agenbag, J., Jarre, A., 2012. Regime-shifts in the southern Benguela shelf and inshore region. Prog. Oceanogr. 106, 80–95.

Brooks, E.N., Deroba, J.J., 2015. When data are not data: the pitfalls of post-hoc analyses that use stock assessment model output. Can. J. Fish. Aquat. Sci. 72, http://dx.doi.org/10.1139/cjfas-2014-0231 (in press).

Bruce, B.D., Evans, K., Sutton, C.A., Young, J.W., Furlani, D.M., 2001. Influence of mesoscale oceanographic processes on larval distribution and stock structure in jackass morwong (*Nemadactylus macropterus*: Cheilodactylidae). ICES J. Mar. Sci. 58, 1072–1080.

Bunnefeld, N., Hoshino, E., Milner-Gulland, E.J., 2011. Management strategy evaluation: a powerful tool for conservation? Trends Ecol. Evol. 26, 441–447.

Burnham, K.P., Anderson, D.R., 1998. Model Selection and Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York, USA.

Caddy, J.F., Seijo, J.C., 2005. This is more difficult than we thought! The responsibility of scientists, managers and stakeholders to mitigate the unsustainability of marine fisheries. Philos. Trans. R. Soc. B 360, 59–75.

Deroba, J.J., Bence, J.R., 2008. A review of harvest policies: understanding relative performance of control rules. Fish. Res. 94, 210–223.

Krimsky, S., 2005. The weight of scientific evidence in policy and law. Am. J. Public Health 95 (Suppl. 1), S1.

Little, L.R., Rowling, K., 2008. Eastern gemfish (*Rexea solandri*) stock assessment based on 2008 survey data. In: Tuck, G.N. (Ed.), Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery: 2006–2007, vol. 2. CSIRO Marine and Atmospheric Research.

MacCall, A.D., Klingbeil, R.A., Methot, R.D., 1985. Recent increased abundance and potential productivity of Pacific mackerel (*Scomber japonicus*). Calif. Co-op. Ocean. Fish. Invest. Rep. 26, 119–129.

New England Fishery Management Council, 2014. Framework Adjustment 51 to the Northeast Multispecies Fishery Management Plan, http://www.nefmc.org/nemulti/index.html

Northeast Fisheries Science Center, 2012. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. In: Northeast Fisheries Science Center, Woods Hole, Massachusetts, 5–9 June, 2012, US Dept of Commerce, Northeast Fisheries Science Center Reference Document 12–18, http://www.nefsc.noaa.gov/publications/crd/crd1218/

O'Boyle, R., Francis, C., Hall, N., Klaer, N., 2012. Stock Assessment Review Committee 54 Panel Summary Report. In: 54th Northeast Regional Stock Assessment Workshop Stock Assessment Review Committee Meeting, Northeast Fisheries Science Center, Woods Hole, Massachusetts, 5–9 June, 2012, http://www.nefsc.noaa.gov/saw/saw54/

Overholtz, W.J., Jacobson, L.D., Melvin, G.D., Cieri, M., Power, M., Libby, D., Clark, K., 2004. Stock assessment of the Gulf of Maine–Georges Bank Atlantic Herring Complex, 2003. Northeast Fisheries Science Center Reference Document 04-06, http://www.nefsc.noaa.gov/publications/crd/crd0406/

Parma, A.M., 1990. Optimal harvesting of fish populations with non-stationary stock recruitment relationships. Nat. Resour. Model. 4, 39–77.

Plagányi, É.E., Punt, A.E., Hillary, R., Morello, E.B., Thébaud, O., Hutton, T., Pillans, R.D., Thorston, J.T., Fulton, E.A., Smith, A.D.M., Smith, F., Bayliss, P., Haywood, M., Lyne, V., Rothlisberg, P.C., 2012. Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. Fish Fish. 15, 1–22.

Pogonoski, J.J., Pollard, D.A., Paxton, J.R., 2002. Conservation Overview and Action Plan for Australian Threatened and Potentially Threatened Marine and Estuarine Fishes. Environment Australia, Canberra, http://www.environment.gov.au/system/files/resources/ca415225-5626-461c-a929-84744e80ee36/files/marine-fish.pdf

Punt, A.E., A'mar, T., Bond, N.A., Butterworth, D.S., de Moor, C.L., De Oliviera, J.A.A., Haltuch, M.A., Hollowed, A.B., Szuwalski, C., 2013. Fisheries management under climate and environmental uncertainty: control rules and performance simulation. ICES J. Mar. Sci. 71, 2208–2220.

Punt, A.E., Smith, A.D.M., 1999. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). ICES J. Mar. Sci. 56, 860–875.

Ricard, D., Minto, C., Jensen, O.P., Baum, J.K., 2012. Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish Fish. 13, 380–398.

Ridgway, K.R., 2007. Long-term trend and decadal variability of the southward penetration of the East Australian current. Geophys. Res. Lett. 34, L13613.

Sagan, C., 1980. "Encyclopaedia Galactica". Cosmos. Episode 12. 01:24 minutes in. PBS.

Shelton, P.A., Sinclair, A.F., Chouinard, G.A., Mohn, R., Duplisea, D.E., 2006. Fishing under low productivity conditions is further delaying recovery of Northwest Atlantic cod (*Gadus morhua*). Can. J. Fish. Aquat. Sci. 63, 235–238.

Sinclair, A.F., 2001. Natural mortality of cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. ICES J. Mar. Sci. 58, 1–10.

Smith, A.D.M., Sainsbury, K.J., Stevens, R.A., 1999. Implementing effective fisheries management systems – management strategy evaluation and the Australian partnership approach. ICES J. Mar. Sci. 56, 967–979.

Smith, T.D., 1994. Scaling Fisheries. The Science of Measuring the Effects of Fishing. New York, Cambridge University Press, pp. 1855–1955.

Steele, J.H., 1996. Regime shifts in fisheries management. Fish. Res. 25, 19–23.

Sullivan, M.C., Cowen, R.K., Steves, B., P, 2005. Evidence for atmosphere–ocean forcing of yellowtail flounder (*Limanda ferruginea*) recruitment in the Middle Atlantic Bight. Fish. Oceanogr. 14 (5), 386–399.

Swain, D.P., Chouinard, G.A., 2008. Predicted extirpation of the dominant demersal fish in a large marine ecosystem: Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. Can. J. Fish. Aquat. Sci. 65, 3215–2319.

Swain, D.P., Savoie, L., Aubry, É., 2012. Recovery Potential Assessment for the Laurentian South designatable unit of Atlantic Cod (*Gadus morhua*): the southern Gulf of St. Lawrence cod stock (NAFO Div. 4T-4Vn(Nov-Apr)). DFO Canadian Science Advisory Secretariat Research Document 2012/052. iii + 51 p.

Szuwalski, C.S., Vert-Pre, K.A., Punt, A.E., Branch, T.A., Hilborn, R., 2015. Examining common assumptions about recruitment: a meta-analysis of recruitment dynamics for worldwide marina fisheries. Fish Fish., http://dx.doi.org/10.1111/faf.12083 (in press).

Vert-pre, K.A., Amoroso, R.O., Jensen, O.P., Hilborn, R., 2013. Frequency and intensity of productivity regime shifts in marine fish stocks. Proc. Natl. Acad. Sci. U. S. A. 110, 1779–1784.

Vooren, C.M., 1972. Postlarvae and juveniles of the tarahiki (Teleostei: Cheilodactylidae) in New Zealand. N. Z. J. Mar. Freshw. Res. 6, 602–618.

Walters, C., Parma, A.M., 1996. Fixed exploitation rate strategies for coping with effects of climate change. Can. J. Fish. Aquat. Sci. 53, 148–158.

Wayte, S.E., 2013. Management implications of including a climate-induced recruitment shift in the stock assessment for jackass morwong (*Nemadactylus macropterus*) in south-eastern Australia. Fish. Res. 142, 47–55.