Retrospective investigation of assessment uncertainty for fish stocks off southeast Australia

Andre E. Punt[1,2], Jemery Day[1], Gavin Fay[3], Malcolm Haddon[1], Neil Klaer[4,1], L. Richard Little[1], Kristin Privitera-Johnson[2], Anthony D.M. Smith[1,4], David C. Smith[1], Miriana Sporcic[1], Robin Thomson[1], Geoffrey N. Tuck[1], Judy Upston[1] and Sally Wayte[5]

[1]CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart, TAS 7001, Australia

[2]School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA

[3]Department of Fisheries Oceanography, School for Marine Science & Technology, University of Massachusetts Dartmouth, New Bedford, MA 02744, USA

[4]Institute for Marine and Antarctic Studies, University of Tasmania, Private Bag 129 Hobart, Tasmania 7005, Australia

5: CSIRO retired

Abstract

There is a need to provide quantitative measures of uncertainty to support fisheries management decision making. A retrospective analysis of historical assessments for fish stocks off southeast Australia is conducted to quantify the extent of uncertainty associated with estimates of spawning stock biomass in absolute terms and when expressed relative to spawning stock biomass over a sequence of reference years. This approach to quantifying uncertainty captures more sources of uncertainty than alternative approaches, such as the estimate of the variance of terminal year spawning stock biomass from asymptotic methods, the extent to which estimates of spawning stock biomass vary among the sensitivity tests that form part of most assessments, and conventional retrospective analyses. By all measures, estimates of spawning stock biomass in absolute terms are much less certain than estimates of relative stock size (i.e. spawning stock biomass relative to a reference level), although application of most current harvest control rules rely on estimates of biomass in absolute terms. Overall, uncertainty in estimates of spawning biomass in absolute terms can be represented as a log-scale standard error of 0.37, while this standard error is 0.18 for estimates of spawning biomass in relative terms. There is considerable variation in among-assessment uncertainty in stock assessment outputs across species groups, with, for example, higher variation for assessments of chondrichthyans compared to other species.

Keywords: Meta-analysis; Retrospective analysis; Stock assessment; Uncertainty estimation

Corresponding author

A.E. Punt

Email: aepunt@uw.edu

Tel: 1-206-221-6319

Introduction

Management strategies for many of the world's major fisheries are based on harvest control rules, HCRs, which can be 'empirical' or 'model-based'. Empirical HCRs calculate management actions, such as limits on fishing effort or catch, as a function of data collected directly from the fishery (e.g., Butterworth and Punt, 1999; de Oliveira and Butterworth, 2004; de Moor et al., 2011). In contrast, 'model-based' HCRs use the outputs from stock assessments that fit population dynamics models to available monitoring data (e.g., IWC, 2012), and are by far the most common type of HCR implemented worldwide. The performances of model-based HCRs depend on the ability of the stock assessments to provide accurate (low bias) and precise (low variation) estimates of the quantities on which the HCR is based.

The HCRs on which fisheries management decisions for US fisheries are based include a buffer between the overfishing level (OFL) and the Acceptable Biological Catch (ABC) to account for scientific uncertainty. Various approaches have been developed to assess the extent of this source of uncertainty (Wiedenmann et al., 2017). The approach adopted for groundfish and coastal pelagic species off the US west coast involves calculating a ABC that is equal to the catch corresponding to F_{MSY} (the OFL) multiplied by a buffer that is less than 1.0 (i.e., ABC = (1-buffer)*OFL). The buffer depends on the quality of the assessment (Category 1: catch-at-age, catch-at-length, or other data that inform a relatively data-rich, quantitative stock assessment; Category 2: some biological indicators that may include a relatively data-limited quantitative stock assessment or non-quantitative assessment; and Category 3: few available data) and is calculated based on a percentile of a lognormal distribution centered on the OFL. The standard deviation (σ) of this distribution depends on the Category, and is selected by the Scientific and Statistical Committee of the Council (in this case the Pacific Fishery Management Council), while the percentile of the lognormal distribution is selected by the Council given their risk tolerance and the consequences of precaution for fisheries for other stocks (PFMC, 2016a). That is, there are two steps to the setting the buffer, the setting of σ, which is purely scientific, and the selection of the degree of risk tolerance, which is a policy decision. The value of σ for stocks in Category 1 is set to the maximum of a default value (0.36), the coefficient of variation of the estimate of biomass for the most recent year, and the log standard error between the estimate of current spawning output from a base model and a low state of nature model that is meant to be half as likely as the base model. The value 0.36 was based on a meta-analysis of errors in estimating biomass from a retrospective analysis (Ralston et al., 2011), while the σ values for Categories 2 and 3 are respectively set to twice (i.e., 0.72) and four times (i.e., 1.44) the default for Category 1 stocks given the presumed additional uncertainty associated with data-limited and data-poor stock assessments.

Stock assessments for many stocks off southeast Australia (Table 1) are based, particularly recently, on similar methods of stock assessment to those applied to groundfish and coastal pelagic species off the US west coast, i.e., integrated analysis based on age-structured population dynamics models (e.g., Methot and Wetzel, 2013; see review of these approaches by Maunder and Punt, 2013). In addition, the HCRs adopted for “data-rich stock assessments” (Tier 1 stocks whose assessments provide “robust assessment of fishing mortality and biomass” – Dowling et al., 2016) are similar to those applied in the US.

No explicit buffer is included in the management strategy for Tier 1 stocks, although the target reference point for biomass is B_{MEY} (the biomass corresponding to Maximum Economic Yield, with a proxy of 48% of unfished spawning stock biomass, i.e., $0.48B_0$) rather than the biomass corresponding to Maximum Sustainable Yield, the proxy for which is $0.4B_0$. Management strategies based on catch curves and trends in catch-per-unit effort data have been developed for stocks for which no model-based assessments are available, i.e., ‘data-poor’ stocks (Wayte and Klaer, 2010; Little et al., 2011; Dowling et al., 2016). While the proxies used for targets in these management strategies are assumed to relate to $0.48B_0$, buffers are supposed to be included explicitly in the management strategies for these ‘data-poor’ (Tiers 2+) stocks.

This paper synthesizes the outcomes from multiple stock assessments conducted through time of finfish stocks harvested off southeast Australia (a “historical retrospective analysis”). Variation in estimated spawning biomass (or depletion) for a given year among multiple assessments of the same stock can arise from multiple sources: 1) chosen model structure; 2) fixed parameter values and prior distribution selection for other parameters; 3) increases in data availability; 4) composition of the review panel; 5) version of software employed and hence how the assessment can be specified; and 6) members of the stock assessment team conducting the assessment (Ralston et al., 2011). The objective of this paper is to estimate the between-assessment variation in estimates of spawning stock biomass (or pup production for chondrichthyan species), and how this variation compares with that for stocks off the US west coast (Ralston et al., 2011). It also considers whether estimates of depletion (biomass relative to a reference point) for southeast Australian stocks are less variable among assessments than estimates of biomass in absolute terms, as might be expected given results of simulation studies of the performance of stock assessment methods (e.g., Punt 1995, 1997; Magnusson and Hilborn, 2007).

Most recent model-based assessments for finfish stocks off southeast Australia have been conducted using Stock Synthesis (Methot and Wetzel, 2013) (Table 1), while historical assessments (generally pre-2004) were based on modeling platforms developed for specific stocks (as is still common in Australia, Dichmont et al., 2016a). We therefore also examine whether adoption of a common assessment platform has reduced between-assessment variation. Unlike stock assessments for many US west coast fish stocks, the primary relative abundance index for assessments of fish stocks off southeast Australia is fishery-dependent standardized catch-per-unit-effort data. We consequently also explore whether estimates of biomass from stock assessments that use fishery-independent data are less variable than those that rely solely on fishery-dependent standardized catch-per-unit effort indices and fishery age- and length-composition data. Finally, we explore the reasons for major changes in estimates of spawning stock biomass (or pup production) in stock assessments from Australia's southeast.

2. Methods

2.1 The historical retrospective analysis

Ralston et al. (2011) considered three methods for calculating an among-assessment coefficient of variation for an assessed species based on different assumptions regarding which, if any, of the assessments provides estimates closest to the truth:

A. All biomass estimates are equally likely to be correct. The variation in biomass is quantified as the standard deviation of $\ell n(B_{i,t} / B_{j,t})$ for $i \neq j$ for all *i*, *j*, and *t* (restricted to be the most recent pre-defined X years) where $B_{i,t}$ is the estimate of spawning stock biomass for year *t* based on assessment *i*. As noted by Ralston et al. (2011) the raw standard deviation is positively biased as it is the ratio of two random variables so a bias correction factor of $1/\sqrt{2}$ is applied.
B. The mean biomass among assessments is the best estimate of central tendency. The variation in biomass is thus quantified as the standard deviation of $\ell n(B_{i,t} / \bar{B}_t)$ where $\bar{B}_t$ is the mean over assessments *i* of $\ell n(B_{i,t})$, again restricted to the most recent X years.
C. The most recent biomass estimates are most likely to be correct. The variation in biomass is thus quantified as the root mean square error between the $\ell n(B_{i,t})$ and

$\ell n(B_{I,t})$ where $B_{I,t}$ is the estimate of spawning biomass for year t based on the most recent stock assessment (I). The root mean square error calculation ignores the estimates of biomass for the assessment conducted in year I, as the associated errors would be zero.

Under the assumption of log-normal errors, the standard deviations, σs, from the three methods can be converted to CVs according to the formula:

$$CV = \sqrt{\exp(\sigma^2) - 1} \quad (1)$$

where σ^2 is the variance of the log-transformed biomass estimates (Evans et al., 2000).

2.1 Alternative ways to estimate σ

The reference analysis of this paper is based on method B (i.e. assuming that the mean biomass among assessment is the best estimate of central tendency), with X set to 20 years (as in Ralston et al., 2011). However, five alternative analyses are conducted to evaluate the sensitivity of the estimated value for σ to data set choice, in addition to use of methods A and C:

(a) Use of only the (currently) most valuable stocks. The stocks were ranked by their ex-vessel value for the most recent year with data (2014-15). Ex-vessel value by species was pro-rated to stock (e.g., east and west pink ling) based on the catches for 2014-15 (or earlier if the most recent assessment did not include the years 2014-15). Six stocks (indicated by asterisks in Table 1) account for more than 85% of the ex-vessel value of the stocks considered in this paper.

(b) Basing the estimate of σ on the results of assessments that have consistently used the results from fishery-independent surveys, as these might be expected to be more reliable than assessments based solely on fishery-dependent data. Recent assessments for tiger flathead (*Neoplatycephalus richardsoni*), and the eastern stock of jackass morwong, (*Nemadactylus macropterus*) have used fishery-independent data, as opposed to earlier assessments, so these stocks are not included in this sensitivity test. Bight redfish (*Centroberyx gerrardi*) and deepwater flathead (*Neoplatycephalus conatus*) are included in the set of species considered to have used fishery-independent data as use of these data started after the third assessment of each stock.

(c) Basing the estimate of σ on the results of assessments for shelf species, slope species and chondrichthyans individually (there is only one deepwater species, orange roughy).

(d) Basing the estimate of σ on increasing (X=25) or decreasing (X=15) the number of years of assessment output used to apply method B.

(e) Using only the assessments conducted using Stock Synthesis. It might be expected that use of Stock Synthesis has reduced discrepancies in assumptions among assessments and increased standardization, which would tend to reduce the (perceived) uncertainty in estimates of spawning stock biomass.

The values for σ are estimated for spawning stock biomass in absolute terms (c.f., Ralston et al. 2011) and when spawning stock biomass is expressed relative to a reference value, i.e., the average spawning stock biomass during 1986-1992. This was a period generally considered to be when many stocks, but by no means all, e.g. orange roughy, were close to ‘desirable conditions’ (e.g., Little et al., 2011), and perhaps more importantly a period for which almost all assessments provided an estimate of spawning stock biomass for at least one year. Spawning stock biomass in absolute terms feeds directly into management through the calculation of Recommended Biological Catches (RBCs; Rayns, 2007), while relative biomass impacts both the calculation of RBCs and also estimates of stock status relative to reference points.

2.3 Selection of stocks

Assessments were considered for inclusion in the retrospective analyses if they satisfied four criteria:

(a) the assessment was reviewed for use in management – this criterion excluded a recent assessment of the eastern stock of gemfish (*Rexea solandri*), which was conducted and not fully reviewed;

(b) there had to be at least one (approved) assessment for the period examined – this criterion eliminated stock assessments for three stocks (saw shark, *Pristiophorus cirratus* and *P. nudipinnis* and elephantfish, *Callorhinchus milii*, as well as assessments of the western stock of orange roughy, *Hoplostethus atlanticus*);

(c) it was possible to obtain the detailed outputs from the assessment in machine readable form – unfortunately, this eliminated some of the very earliest assessments of eastern orange roughy for which assessment reports exist (e.g., CSIRO and TDPIF, 1996; Bax 1997, 2000a,b,c), but model output was not readily available, as well as some early assessments of pink ling (*Genypterus blacodes*); and

(d) the stock structure underlying the assessment matched that for the most recent assessment (e.g., the earliest assessment of blue warehou, *Seriolella brama* [Punt, 1998] assessed this species as a single stock rather than two stocks, while the same is the case for pink ling).

There are assessments that did not satisfy more than one of these criteria. For example, single assessments were developed for the western stock of gemfish (Chambers et al., 2014) and silver trevally *Pseudocaranx dentex* (Day et al., 2007), and they were not approved for management. Also, assessments that included additional data that would inform parameters, i.e., ‘update assessments’ that simply projected the population ahead under recent catches (e.g., the update assessment for silver warehou, *Seriolella puncata* conducted in 2006; G. Tuck, pers. comm.) were omitted.

The time-trajectories of spawning stock biomass on which the bulk of the analyses are based (see Section 2.4 for an exception) are those for the base-case (or best) analysis. This is usually the analysis on which management advice was based.

2.4 Alternative measures of uncertainty

The results from applying the historical retrospective analysis are contrasted with alternative ways to quantify uncertainty. Specifically, conventional retrospective analyses (in which recent years of data are removed in turn from the most recent assessment and the parameters re-estimated, e.g. Mohn, 1999; Hurtado et al., 2015), estimates of the logarithms of terminal year biomass based on inverting the Hessian matrix (i.e., asymptotic methods), which can be converted into a value for σ, and the variation in estimates of biomass and depletion among all sensitivity tests for the most recent assessment are explored as potential measures of uncertainty. Assessment uncertainty is based on the variance in estimates from all sensitivity tests for the 'sensitivity test method', which differs from how sensitivity tests are used to quantify uncertainty for US west coast groundfish, which involves only the base model and a model that is considered to be half as likely as the base model. Assessments conducted using Bayesian estimation methods also provide estimates of uncertainty via the posterior distribution for spawning stock biomass. However, very few assessments of stocks off southeast Australia have been implemented as Bayesian analyses and so this possibility was not evaluated.

3. Results

3.1 Stock assessments

Table 2 summarizes the Tier 1 stock assessments considered in this paper (based on 19 stocks). All but two of the stock assessments are based on forward projection integrated analysis methods. The exceptions are school whiting (*Sillago flindersi*) assessments conducted in 1999 and 2003, which were based on ADAPT-VPA methods (e.g., Punt, 1999). Most of the assessments before 2006 were based on either stock assessment methods developed for specific stocks ("Case" in Table 2) or Coleraine (Hilborn et al., 2000; "Col" in Table 2). Subsequently, assessments were based on Stock Synthesis ("SS" in Table 2) for the bulk of the stocks (SS2 until about 2008 and SS3 thereafter). Stock Synthesis has been relatively stable in terms of its basic functionality since the introduction of SS2, but there have been some major developments which would impact the results of assessments. The exceptions to the use of Stock Synthesis since 2006 have been the assessments of gummy shark (*Mustelus antarcticus*) and school shark (*Galeorhinus galeus*). Both assessments are based on spatially structured population dynamics models that share parameters among stocks and use tagging data to estimate fishing mortality (both species) and movement (school shark) (Pribac et al., 2005; Punt et al., 2000). These features are not currently included in Stock Synthesis (and likely any other current stock assessment packages; Dichmont et al., 2016a).

3.2 Time trajectories of spawning stock biomass and depletion

Time trajectories of spawning stock biomass (or pup production for chondrichthyan species) from the base-case (or best) assessments for each stock from each assessment show that there is considerable variation in the extent to which the assessments provide consistent estimates of spawning stock biomass (Fig. 1). Some assessments, for example for tiger flathead, orange roughy, and gummy shark (TAS), and to a lesser extent blue warehou (east and west) and eastern gemfish, are very consistent in their estimates of spawning stock biomass among assessments. Other assessments (i.e., jackass morwong east, redfish, deepwater flathead, blue grenadier, *Macruronus novaezelandiae*, and school shark) are consistent over periods of time and then change markedly. There is a third group of assessments - those that seem to imply the same trajectory over time in a relative sense, but where the absolute scale of biomass (or pup production) changes substantially among assessments (i.e., Bight redfish, *Centroberyx gerrardi*, jackass morwong west, gummy shark in Bass Strait and off South Australia).

The ability to estimate spawning stock biomass in a relative sense is consistent with that to estimate absolute spawning stock biomass, except for the third group of species for which performance in a relative sense is markedly better than in absolute sense (Fig. 2). Relative performance is also better for other assessments such as those for deepwater flathead and redfish.

3.3 Quantification of among assessment variation

Table 3 lists the estimates of σ across all species for each method of estimation as well as the eight sensitivity scenarios that involve calculating σ from subsets of the available stocks. The base-case estimates of σ use the assessments for all of the stocks and method B. The base-case estimate of σ is 0.37 for spawning stock biomass in absolute terms (henceforth referred to as "SSB_{abs}") and is 0.18 for spawning stock biomass in a relative sense (henceforth referred to as "SSB_{rel}"); the difference in values for σ for SSB_{abs} vs SSB_{rel} is consistent with the visual expectations from Figs 1 and 2. The lower among-assessment variation in SSB_{rel} compared to SSB_{abs} is confirmed when method B is compared by stock category (Fig. 3).

The estimates of σ are not appreciably sensitive to assuming that all assessments are equally likely to be "best" (method A), but they are substantially higher (0.60 for SSB_{abs} and 0.31 for SSB_{rel}) when the most recent assessment is taken to provide the "best" estimates (method C). This is not unexpected given Figs 1 and 2 where the most recent assessments for Bight redfish, jackass morwong west, redfish, eastern gemfish, pink ling west, gummy shark (TAS) and school shark lead to the lowest (or, less commonly, highest) estimates of SSB_{abs} (Fig. 1). This pattern is less obvious when spawning stock biomass is expressed relative to the average over 1986-1992 (Fig. 2).

The estimates of σ are not lower for the valuable species nor for those species that have consistently used fishery-independent data (rows “Valuable species only” and “With consistent fishery-independent data only”; Table 3). The results are also not very sensitive to changing the number of years used to compare σ (rows “X=15” and “X=25” in Table 3). The value for σ differs among species groups, with the among-assessment variation in SSB_{abs} being lower for slope species than the full set of species, while the chondrichthyans have greater among-assessment variation in SSB_{abs}. However, this conclusion does not hold for SSB_{rel}, with the assessments for chondrichthyans showing very low variation in SSB_{rel}, compared to in SSB_{abs}.

Basing the estimate of σ on assessments conducted using only Stock Synthesis suggests that use of Stock Synthesis as an assessment platform has led to less among-assessment variation in SSB_{abs}. However, this is potentially misleading because none of the chondrichthyan assessments have been conducted using Stock Synthesis, and the Stock Synthesis assessments all occurred from 2006 onwards, implying they included more data than earlier assessments. This caveat is confirmed if σ is calculated ignoring the chondrichthyan assessments (0.324), which is lower than the base value, but not substantially. In addition, the change to Stock Synthesis coincided with the application of a structural adjustment to the SESSF fleet and the imposition of the current harvest strategy policy with its $0.48B_0$ proxy for a target.

3.4 Stock specific results and alternative measures of uncertainty

It is possible to apply method B to each stock individually, resulting in stock-specific values for σ (Table 4). As expected from Figs 1 and 2, there is considerable variation among stocks in the value for σ, particularly when SSB_{abs} is analyzed.

The estimates of among-assessment variation in SSB_{abs} from method B generally exceed the estimates of standard errors of log-biomass for the terminal year, in some cases substantially (e.g., Bight redfish and redfish; Fig. 4a). Gummy shark off Tasmania is the only case for which the standard error of log-biomass for the terminal-year (0.171) is larger than the among-assessment variation in SSB_{abs} (0.086), although there are only two assessments for gummy shark off Tasmania and both were conducted quite recently.

The situation for among-sensitivity-test variation differs between SSB_{abs} and SSB_{rel} (Figs 4b,c). The variation in estimates of SSB_{rel} among sensitivity tests is generally of the same order of magnitude as the among-assessment variation in SSB_{rel} (Fig. 4c). However, this is not the case for SSB_{abs}, for which the sensitivity tests are unable to capture that different assessments may lead to different estimates abundance in absolute but not relative terms (see Fig. 1).

4. Discussion

The results of the historical retrospective analysis suggest that estimates of spawning stock biomass (or pup production) from stock assessments exhibit considerable among-assessment variation (a stock-averaged value for σ of 0.37), but that this variation differs among taxa. The value for among-assessment variation in spawning stock biomass is very similar to that for the other region where this type of analysis has been undertaken (the US west coast). For most stocks, this variation exceeds that based on other measures for quantifying assessment uncertainty (such as asymptotic standard errors and sensitivity tests), raising methodological questions as well as posing a challenge for managers regarding how best to allow such uncertainty to flow in decision making.

4.1 Why do some assessment results change markedly among years?

Several of the stocks have very high values for σ, particularly for spawning stock biomass in absolute terms. The reasons for this are fairly diverse, for example:

- Bight redfish. The most recent results are less uncertain than those from previous assessments as only now are the fishery catches starting to induce sufficient depletion for the fishery data to become informative. This has also coincided with a marked increase in the number of data sources included in the assessment. Nevertheless, important sources of uncertainty derive from the relatively poor estimates of natural mortality and seemingly variable size at maturity that may reflect an unknown degree of spatial structure in biological properties of the stock (Haddon, 2016).
- Gummy shark. The index of abundance pertains to the abundance of animals that are available to the gear (mainly gillnets), which is a relatively small component of the population and does not include many mature animals. The scaling from the exploited component of the population to the mature component (and hence pup production) depends on the value of natural mortality (which is estimated as part of the assessment) as well as inferences regarding what proportion of the population is in the region fished using gillnets, and this proportion can change between assessments.
- Jackass morwong east. The trend in the spawning stock biomass of this stock changed between the 2004 and 2006 assessments. This change is attributable to a change to the way catch and effort data were standardized to develop the key tuning index for the assessment. Earlier standardizations ignored catch and effort records when the catch was 30kg or less while later standardizations used such data. This, along with an increasing trend in small catches, changed the trend in the index of abundance.

- School shark. The large change in trend between the 1996 and 2000 assessment resulted from a major change to the structure of the assessment from one that treated all of southern Australia as a single homogeneous region in which one stock was located (Punt and Walker, 1998) to an assessment method that modeled school shark as two “movement stocks” found in eight regions (Punt et al., 2000). The latter assessment disaggregated the data used in the earlier assessment spatially, and made use of additional data sources, in particular length-composition and tagging data.

The greatest sensitivity arises for “lightly fished” stocks such as Bight redfish (i.e., spawning stock biomass above $0.48B_0$) for which the trend can be well determined, but abundance in an absolute sense cannot be. This is the case even when a fishery-independent index of abundance may be available (e.g., Bight redfish). However, this sensitivity is common to all stock assessments (e.g., Ralston et al., 2011), and not unique to southeast Australia.

Changes, particularly in results for Tier 1 assessments of main species in the SESSF fishery from one assessment to the next are of concern for all involved – scientists, managers and industry. There is a direct linkage of those assessment results to the total allowable catches (TACs) set for the fishery if the assessment is accepted. Such changes receive scrutiny when reviewed at fishery Resource Assessment Group meetings that include representatives from all of these groups. Since about 2006 and especially for SS assessments, bridging analyses have normally been provided with each assessment that separates differences from the last assessment due to improvement in historical data records, adding of recent data or data not previously considered by the assessment, and changes to assumptions in the assessment model. Data changes are expected and normally justified, but changes to structural assumptions of the assessment model require justification based on scientific merit, and are examined on that basis. Such a process has probably led to more consistency in assessment procedures, for example, across four different analysts for tiger flathead from 2004 to 2016. A possible disadvantage of this procedure may be to discourage a completely new assessment with completely new set of assumptions from a new analyst (but see orange roughy assessment by Upston et al. 2015).

4.2 Methodological conclusions and caveats

Several methods to quantify uncertainty are applied in this paper. Reporting asymptotic standard errors of model outputs is a common practice and is indeed a requirement in some jurisdictions (e.g., PFMC 2016b). However, such standard errors were very poor predictors of uncertainty in spawning stock biomass, as quantified by the historical retrospective analysis (Fig. 4b; also Fig. 4 of Ralston et al. 2011). This is one reason that the buffer to account for scientific uncertainty by the Pacific Fishery Management Council is based multiple measures of uncertainty. It is not surprising that asymptotic standard errors do not reflect among-assessment variation in estimates of spawning stock biomass well because they are conditioned on the selected model structure, data set choices, assumed parameters, and weights assigned to data, aspects that change between assessments and that may be key drivers of among-assessment variation.

It might be expected that the variation in estimates of terminal year spawning stock biomass among sensitivity tests may better capture the “true” extent of uncertainty. However, that is generally not the case, at least under the assumption that the historical retrospective analyses estimates of uncertainty are most accurate and when viewing spawning stock biomass in absolute terms (Fig. 4b). In addition, it may have been anticipated that assessments that included many sensitivity tests would have better captured uncertainty as reflected by the historical retrospective analysis, but that is not the case (results not shown). A difficulty with basing measures of assessment uncertainty on the results of all sensitivity tests is that there is a considerable range in the number and type of such tests – in fact some of the assessments did not report results for any sensitivity tests (Table 4 “N/A”). Moreover, the sensitivity tests in Australian assessments are not designed to capture a pre-specified amount of uncertainty unlike the states of nature used to develop decision tables for US west coast groundfish. The level of uncertainty inferred from sensitivity tests might have been higher had this been the case.

The preferred approach of this paper is ‘historical retrospective analysis”, and provision of this output is required in some jurisdictions (e.g., PFMC, 2016b), but not currently in Australia. Part of the reason for this is the need to retain model input and output files, a practice not implemented formally in Australia until recently. However, “historical retrospective analysis” can only be performed when multiple assessments are available for the assessed stocks, which reduced the number of stocks that could be included in this paper (see Section 2.3). Conventional retrospective analyses could mimic the results of a historical retrospective analysis, but that is seldom the case (see Fig. 5 for a comparison between a historical retrospective analysis (upper panel) and conventional retrospective analysis (lower panel) for Bight redfish). The reasons that the results from a historical retrospective analysis differ from those of a conventional retrospective analysis are the same as outlined for asymptotic standard errors of terminal year biomass.

Historical retrospective analysis is not without concerns however, and these concerns could lead to both over- and under-estimation of measures of assessment uncertainty. The former can arise if assessments change in structure (and hence outcomes are due to a change in understanding and stock assessment best practice), which is the case for school shark when the assessment moved from being based on a spatially aggregated to a spatially disaggregated population dynamics model. In addition, over-estimation could arise if the (true) extent of uncertainty is changing over time, as might be expected for stocks that initially had few data, but now have more data.

The concerns with over-estimation of uncertainty should be balanced with those of under-estimation. Under-estimation can arise for several reasons, most important of which is that the assessments are not “independent” in any real sense. For example, the bulk of the data used in the most recent assessment of a stock will be the same as those data used in several previous assessments. This is particularly a concern for stocks (such as blue grenadier) that have been assessed frequently (often annually in the early 2000s; Table 2). Similarly, lack of independence arises because the same analysts often conduct a single assessment for multiple years, and often will start an assessment update from the previous assessment, only making changes to key assumptions (e.g., fleet structure, value for natural mortality) when “needed”. Even changes to assessment software may not be expected to result in very different estimates because in most cases the underlying population dynamics equations are very similar (e.g., most of the assessments analyzed were based on statistical single stock-age-structured population dynamics models) (see Deroba et al. [2014] for comparisons of assessment outcomes based on different assessment platforms). Peer-review processes will also tend to “stabilize” assessments by encouraging continuity of assessments.

The number of stock assessments conducted annually is unbalanced, with many more assessments conducted during the period 2004-2010 than either previously (although some of the assessments conducted before 2004 were excluded under our criteria) and particularly subsequently (Table 2). This imbalance will impact estimates of σ, especially if assessments conducted in sequential years differed little in terms of assumptions, but were largely “repeat assessments” with minor revisions to data series.

All things equal, the value of σ will underestimate estimates of sustainable catches based on harvest control rules that involve projecting the population ahead, owing to uncertainties in quantities such as the target harvest rates (e.g. Punt et al., 2014) as well as uncertainties caused by variation in natural mortality rates and uncertainty about poorly-recruited year-classes.

4.3 Management implications

The estimate of among-assessment uncertainty for southeast Australian fish stocks based on application of method B (0.37) is almost identical to the value estimated using the same method for groundfish and coastal pelagic species off the US west coast (0.36; Ralston et al., 2011). The almost exact similarity is, of course, coincidental, but suggests that considerable assessment uncertainty is not unique to southeast Australia. A value for σ of 0.37 corresponds (quite closely), under the assumption of log-normality, to the estimate of spawning stock biomass for an (average) assessment differing from the expected estimate of spawning stock biomass from -50% to +100% with 95% certainty. This level of possible error is accounted for in the US west coast groundfishery by adopting a buffer of between 4.4% and 8.7% for data-rich stocks to avoid fishing mortality unintentionally exceeding the proxy for F_{MSY} with probability more than 40%. Many of the assessments in Table 1 would not qualify as data rich in the US west coast groundfishery (see Appendix E of PFMC, 2016b), and would have a buffer closer to 8.7-17.2%. The management system for southeast Australian fish stocks differs from that in the US as the target biomass for southeast Australian stocks is the biomass that corresponds to Maximum Economic Yield, which is conventionally assumed to be 20% larger than the biomass corresponding to Maximum Sustainable Yield (0.48B_0 vs 0.4B_0; Punt et al., 2014; Rayns, 2007). There is consequently an implicit buffer in the southeast Australian system to prevent fishing mortality exceeding F_{MSY}, which is the conventional definition for overfishing.

The value for σ of 0.37 is an average over many stocks, and it is clear that some stocks, particularly the chondrichthyans and those stocks that have been fished relatively lightly, are subject to much greater uncertainty, at least in relation to the absolute scale of spawning stock biomass (or pup production). Whether the current (implicit) buffer is large enough to prevent fishing mortality exceeding the proxy for F_{MSY} with acceptable uncertainty remains unclear, but could be the subject of management strategy evaluation type analyses.

The estimates of spawning stock biomass in relative terms are more robust than the estimates of absolute abundance, a well-known result (e.g., Punt 1995, 1997; Magnusson and Hilborn, 2007). However, if the scaling (which was set to 1986-1992, a period for which data are available for most stocks) had been to unfished biomass, B_0, the inferred uncertainty would have been larger. Such results are not reported here because several of the older assessments did not estimate B_0. Many management systems, including those in Australia, have implemented management strategy systems in which the catch limit from data-rich assessments depends critically on estimates of absolute abundance, with the management systems often simulation tested. The results of this paper suggest that previous evaluations of the performance of management strategies that fail to account for changes over time in factors such as the assessment software being used, the availability of new data streams and changes to pre-specified parameters may have over-estimated the performance of such systems and that management strategies that rely on trends in abundance (such as Australia's Tier 4 harvest control rule; Little et al., 2011) may be more robust to assessment uncertainty. An evaluation of this hypothesis is beyond the scope of the present paper, but should be considered in future work.

Conclusions and recommendations

The results of this paper (re)highlight that 1) even assessments of data-rich stocks are subject to not-inconsiderable uncertainty; 2) estimates of biomass in absolute terms are more uncertain than estimates of biomass scaled to biomass in a reference year; and 3) the level of uncertainty varies considerably among stocks. The estimate of among-assessment variation in biomass is found to be larger than other (simpler to obtain) measures of uncertainty, but this estimate could either under- or over-estimate uncertainty.

The results of this paper suggest a need for other jurisdictions to evaluate assessment uncertainty using historical retrospective analyses, which should be a standard output of any assessment (i.e., each assessment should report outputs such as Figs 1 and 2 for the stock under consideration). In addition, the difficulty in understanding the reasons for some of the changes in assessment results suggests that assessment analysts should carefully document and highlight changes to assessment assumptions and data inputs. Sampson et al. (in press) provide an excellent example (their Tables 11 and 12) of documenting the consequences of changes in assessment assumptions and data.

The estimate of σ could be used directly in harvest control rules, as is the case for US west coast fisheries or less directly in management strategy evaluations. The latter could involve setting the level of error in estimating biomass to σ or ensuring that simulated assessment error is consistent with the value of σ for the stock under consideration or the group of stocks within which the stock falls.

Finally, the results of this paper also provide guidance on implementing a buffer system in the Australian context (Dichmont et al., 2017a), with the intention for them to be used to attain risk equivalency across assessment methods or tiers (Dichmont et al., 2016b, 2017b; Fulton et al., 2016).

Acknowledgements

This research has been supported by CSIRO Oceans & Atmosphere, the Australian Fisheries Management Authority (AFMA) and the Fisheries Research and Development Corporation (FRDC). The authors are also grateful to the review efforts made by the various Resource Assessment Groups in which these assessments were first presented. Natalie Dowling (CSIRO), Steve Ralston (NMFS retired), and an anonymous reviewer are thanked for their comments on an earlier version of the paper. GF was supported by a grant from the Cooperative Institute for the North Atlantic Region, Woods Hole Oceanographic Institution, pursuant to NOAA Award NA14OAR4320158. KPJ was funded by the National Science Foundation Graduate Research Fellowship Program. AEP was [partially] funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement and NA15OAR4320063, Contribution No. ______.

References

Bax, N., 1997. Stock Assessment Report 1997: Orange Roughy. Report for the South East Fishery Stock Assessment Group. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart. 64 p. (Unpublished report held by CSIRO, Hobart).

Bax, N., 2000a. Stock Assessment Report: Orange Roughy 1995, Stock Assessment Report, South East Fishery Stock Assessment Group, Australian Fisheries Management Authority, Canberra, 55 p. (Unpublished report held by AFMA, Canberra).

Bax, N., 2000b. Stock Assessment Report: Orange Roughy 1996, Stock Assessment Report, South East Fishery Stock Assessment Group, Australian Fisheries Management Authority, Canberra, 37 p. (Unpublished report held by AFMA, Canberra)

Bax, N., 2000c. Stock Assessment Report 2000: Orange Roughy. Report for the South East Fishery Stock Assessment Group. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart. 81 p. (Unpublished report held by CSIRO, Hobart).

Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985–998.

Chambers, M., Moore, A., Barnes, B., 2014. Quantitative assessment of the western gemfish stock. Australian Bureau of Agricultural and Resource Economics and Sciences research report, ABARES, Canberra. 33pp.

CSIRO and TDPIF, 1996. Orange roughy 1994, Stock Assessment Report, South East Fishery Assessment Group. Australian Fisheries Management Authority, 204 p. (Unpublished report held by AFMA, Canberra).

Day, J., 2010. School whiting (*Sillago flindersi*) stock assessment based on data up to 2008. pp. 190-249. In: Tuck, G.N. [Ed.] *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2009, Part 1*. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research.

Day, J., In press. Tiger flathead (*Neoplatycephalus richardsoni*) stock assessment based on data up to 2015. pp x-x. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery*. Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Day, J., Rowling, K., Punt, A.E., 2007. Updated stock assessment for silver trevally (*Pseudocaranx dentex*) in the south east fishery: August 2006. pp. 296–319. In: Tuck, G.N. [Ed.] *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2006-2007. Volume 1: 2006*. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research.

Day, J., Thomson, R., Tuck, G., 2016. Silver Warehou (*Seriolella punctata*) stock assessment based on data up to 2014. pp 74-126. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2015, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

De Oliveira, J.A.A., Butterworth, D.S., 2004. Developing and refining a joint management procedure for the multispecies South African pelagic fishery. ICES J. Mar. Sci. 61, 1432–1442.

De Moor, C.L., Butterworth, D.S., De Oliveira, J.A.A., 2011. Is the management procedure approach equipped to handle short-lived pelagic species with their boom and bust dynamics? The case of the South African fishery for sardine and anchovy. ICES J. Mar. Sci. 68, 2075–2085.

Deroba, J.J., Butterworth, D.S., Methot Jr, R.D., De Oliveira, J.A.A., Fernandez, C., Nielsen, A., Cadrin, S.X., Dickey-Collas, M., Legault, C.M., Ianelli, J. Valero, J.L., 2014. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES J. Mar. Sci. 72, 19–30.

Dichmont, C.M., Deng, R.A. and A.E. Punt. 2016a. How many of Australia's stock assessments can be conducted using stock assessment packages? Mar. Pol., 74, 279–287.

Dichmont, C,M., Punt, A.E., Dowling, N., De Oliveira, J.A.A., Little, L.R., Sporcic, M., Fulton, E., Gorton, R., Klaer, N., Haddon, M., Smith, D.C., 2016b. Is risk consistent across tier-based harvest control rule management systems? A comparison of four case-studies Fish and Fish. 17, 731–747.

Dichmont, C.M., Fulton, E., Punt, A.E., Little, L.R., Dowling, N., Gorton, R., Sporcic, M., Smith, D.C., Haddon, M., Klaer, N., 2017a. Operationalising the risk-cost-catch trade-off. CSIRO Oceans and Atmosphere. Fisheries Research and Development Corporation Project 2012-202, Brisbane, October 138pp.

Dichmont, C.M., Fulton, E., Gorton, R., Sporcic, M., Dowling, N., Little, L.R., Punt, A.E., Haddon, M., Smith, D.C., 2017b. From data rich to data-limited harvest strategies – does more data mean better management? ICES J. Mar. Sci. 74, 670–686.

Dowling, N.A., Punt, A.E., Little, L.R., Dichmont, C., Smith, D.C., Haddon, M., Sporcic, M., Fulton, E.A., Gorton, R.J., 2016. Assessing a multilevel tier system: the role and implications of data quality and availability. Fish. Res. 183, 588–593.

Evans, M., Hastings, N., Peacock, B., 2000. *Statistical Distributions Third Edition*. John Wiley & Sons Inc. New York. 221 pp.

Fulton, E.A., Punt, A.E., Dichmont, C.M., Gorton, R., Sporcic, M., Dowling, N., Little, L.R. Haddon, M., Klaer, N., Smith, D.C., 2016. Developing risk equivalent data-rich and data-limited harvest strategies. Fish. Res. 183: 574-587.

Haddon, M., 2016. Bight redfish (*Centroberyx gerrardi*) stock assessment using data to 2014/2015. pp 9-50. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2015, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Haddon, M. In press. Deepwater flathead (*Platycephalus conatus*) stock assessment using data to 2015/2016. GAB RAG Dec. 2016. Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere. 41 p.

Hilborn, R., Maunder, M.N., Parma, A., Ernst, B., Paynes, J., Starr, P.J., 2000. Documentation for a general age-structured Bayesian stock assessment model: code named Coleraine. Fisheries Research Institute, University of Washington, Seattle, Washington (FRI/UW 00/01).

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R.R., McGillaird, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-pre, K.A., Whitten A.R., Punt, A.E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci. 72, 99–110.

International Whaling Commission. 2012. The Revised Management Procedure (RMP) for Baleen Whales. J. Cet. Res. Manage (Suppl.) 13, 485–494.

Little, L.R., Rowling, K. 2009. Eastern Gemfish (*Rexea solandri*) stock assessment based on 2008 survey data. pp. 167-208. In: Tuck, G.N. [Ed.] *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2009, Part 1*. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research.

Little, L.R., Wayte, S.E., Tuck, G.N., Smith, A.D.M., Klaer, N., Haddon, M., Punt, A.E., Thomson, R., Day, J., Fuller, M., 2011. Development and evaluation of a cpue-based harvest control rule for the southern and eastern scalefish and shark fishery of Australia. ICES J. Mar. Sci. 68, 1699–1705.

Magnusson, A., Hilborn, R., 2007. What makes fisheries data informative? Fish. Fish. 8, 337–358.

Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. 142, 61–74.

Methot, R.D., Wetzel, C.R., 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99.

Mohn, R., 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488.

Pacific Fishery Management Council (PFMC) 2016a. Pacific Coast Groundfish Fishery Management Plan for the California, Oregon, and Washington Groundfish Fishery. http://www.pcouncil.org/wp-content/uploads/2017/03/GF_FMP_FinalThruA27-Aug2016.pdf

Pacific Fishery Management Council (PFMC) 2016b. Terms of Reference for the groundfish and coastal pelagic species assessment review process for 2017-2018. http://www.pcouncil.org/wp-content/uploads/2017/01/Stock_Assessment_ToR_2017-18.pdf

Pribac, F., Punt, A.E., Walker, T.I., Taylor, BL., 2005. Using length, age and tagging data in a stock assessment of a length selective fishery for gummy shark *(Mustelus antarcticus)*. J. Northw. Atl. Fish. Sci. 35, 267–290.

Punt, A.E., 1995. The performance of a production-model management procedure. Fish. Res. 21, 349–374.

Punt, A.E. 1998. *Ad hoc* tuned VPA analyses for the blue warehou resource off southern Australia for 1998. Document submitted to the 30 – 31 May 1998 meeting of BWAG. 12 pp.

Punt, A.E., 1997. The performance of VPA-based management. Fish. Res. 29, 217–243.

Punt, A.E., 1999. ADAPT and integrated analysis assessments of school whiting. Document submitted to the 17-18 June 1999 school whiting workshop. 18 pp.

Punt, A.E., 2008. Data analysis and preliminary updated stock assessment of blue warehou (*Seriolella brama*) based on data up to 2008. SlopeRAG document. 53 pp.

Punt, A.E., Walker, T.I., 1998. Stock assessment and risk analysis for the school shark (*Galeorhinus galeus*) resource off southern Australia. Mar. Freshw. Res. 49, 719–731.

Punt, A.E., Thomson, R., Sporcic, M., In press. Gummy shark assessment update for 2016, using data to the end of 2015. pp x-x. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Punt, A.E., Smith, A.D.M., Smith, D.C., Tuck, G.N., Klaer, N.L., 2014. Selecting relative abundance proxies for B_{MSY} and B_{MEY}. ICES J. Mar. Sci. 71, 469–483.

Punt, A.E., Pribac, F., Walker, T.I., Taylor, B.L, Prince, J.D., 2000. Stock assessment of school shark *Galeorhinus galeus* based on a spatially-explicit population dynamics model. Mar. Freshw. Res. 51, 205–220.

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J., Conser, R.J., 2011. An approach to quantifying scientific uncertainty in stock assessment. Fish. Bull. 109, 217–231.

Rayns, N., 2007. The Australian government's harvest strategy policy. ICES J. Mar. Sci. 64, 596–598.

Sampson, D.B., Hamel, O.S., Bosley, K., Budrick, J. Cronin-Fine, L., Hillier, L.K., Hinton, K.E., Krigbaum, M.J., Miller, S., Privitera-Johnson, K.J., Ramey, K.L., Rodomsky, B.T., Solinger, L.K., Whitman, S.D., In press. 2017 Assessment Update for the US West Coast Stock of Arrowtooth Flounder. Pacific Fishery Management Council. 7700NE Ambassador Place, Portland, OR 97220. http://www.pcouncil.org/groundfish/stock-assessments/

Savage, J. 2015. Australian fisheries and aquaculture statistics 2015, Fisheries Research and Development Corporation project 2016-246, ABARES, Canberra. http://data.daff.gov.au/data/warehouse/9aam/afstad9aamd003/2015/AustFishAquacStats_2015_v1.0.0.pdf

Thomson, R., Punt, A., 2010. Stock assessment update for school shark *Galeorhinus galeus* based on data up to 2008: re-analysis and final draft. pp 295-311. In: Tuck, G.N. [Ed.] *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2009, Part 1.* Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research.

Tuck, G.N., 2014. Stock assessment of blue grenadier *Macruronus novaezelandiae* based on data up to 2012. pp 61-115. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2013, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Tuck, G.N., 2015. Stock assessment of redfish *Centroberyx affinis* based on data up to 2013: Supplement to the October 2014 Shelf RAG paper. pp 148-168. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2014, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Tuck, G.N., Day, J., Wayte, S., 2016a. Assessment of the eastern stock of Jackass Morwong (*Nemadactylus macropterus*) based on data up to 2014. pp 156-204. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2015, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Tuck, G.N., Day, J., Wayte, S., 2016b. Assessment of the western stock of Jackass Morwong (*Nemadactylus macropterus*) based on data up to 2014. pp 221-243. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2015, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Upston, J., Punt, A.E., Wayte, S., Ryan, T., Day, J., Sporcic, M., 2015. Orange Roughy (*Hoplostethus atlanticus*) Eastern Zone stock assessment incorporating data up to 2014. pp. 10–81 In: Tuck, G.N. [Ed.] *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2014. Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere Flagship.

Wayte, S., Klaer, N., 2010. An effective harvest strategy using improved catch-curves. Fish. Res. 106, 310–320.

Whitten, A.R., Punt, A.E., 2014. Pink ling (*Genypterus blacodes*) stock assessment based on data up to 2012. pp 116-142. In: Tuck, G.N. [Ed.]. *Stock Assessment for the Southern and Eastern Scalefish and Shark Fishery 2013, Part 1.* Australian Fisheries Management Authority and CSIRO Oceans and Atmosphere.

Wiedenmann, J., Wilberg, M., Sylvia, A., Miller, T., 2017. An evaluation of acceptable biological catch (ABC) harvest control rules designed to limit overfishing. Can. J. Fish. Aquat. Sci. 74, 1028–1040.

Table

Table 1. Summary of Tier 1 assessment types applied to selected stocks in southeast Australian fisheries, including temporal coverage employed in assessments. Stocks indicated by an asterisk are those that are most valuable in the fishery and those indicated by & are above the target biomass. Stock Synthesis (SS).

Common name (stock)	Value (2014-15)[1] ('000 AUD)	Type of assessments	Number of assessments used	With consistent fishery-independent data	Range of years	Current depletion (based on the most recent assessment)	Most recent assessment
Shelf species							
Bight redfish	1,266	Case-specific, SS	6	Yes	1960-2014	0.621[&]	Haddon (2016)
Deepwater flathead	4,230*	Case-specific, SS	7	Yes	1980-2015	0.448[&]	Haddon (In press)
Jackass morwong (east)	399	Coleraine, SS	8	No	1915-2014	0.094	Tuck et al. (2016a)
Jackass morwong (west)	27	SS	5	No	1986-2014	0.630[&]	Tuck et al. (2016b)
Redfish	232	Case-specific, SS	2	No	1915-2013	0.090	Tuck (2015)
School whiting	2,513*	Case-specific, SS	5	No	1947-2008	0.434	Day (2010)
Tiger flathead	15,428*	Case-specific, SS	7	No	1915-2015	0.425[&]	Day (in press)
Slope species							
Blue grenadier	1,854	Case-specific, SS	12	Yes	1960-2012	0.777[&]	Tuck (2014)
Blue warehou (east)	15	Case-specific, SS	4	No	1986-2008	0.153	Punt (2008)
Blue warehou (west)	15	Case-specific, SS	4	No	1986-2008	0.173	Punt (2008)
Gemfish (east)	224	Case-specific, SS	3	No	1968-2998	0.153	Little and Rowling (2009)
Ping ling (east)	195	SS	6	No	1970-2013	0.199	Whitten and Punt (2014)
Ping ling (west)	2,071	SS	6	No	1970-2013	0.432	Whitten and Punt (2014)
Silver warehou	2,450*	Case-specific, SS	8	No	1980-2014	0.316	Day et al. (2016)
Deep species							
Orange roughy (east)	0 (fishery closed)	Case-specific, SS	3	Yes	1980-2014	0.226	Upston et al. (2015)
Shark species							
Gummy shark (Bass Strait)	9,085*	Case-specific	5	No	1927-2016	0.530[&]	Punt et al. (In press)
Gummy shark (South Australia)	4,460*	Case-specific	5	No	1927-2016	0.632[&]	Punt et al. (In press)
Gummy shark (Tasmania)	1,026	Case-specific	2	No	1927-2016	0.750[&]	Punt et al. (In press)
School shark	1,740	Case-specific	4	No	1927-2008	0.099	Thomson and Punt (2010)

1: Source: Savage (2015), pro-rated based on catches by stock for 2014 and 2015 where the value was given by species in total (2013 for pink ling)

Table 2. Stocks considered, and the software employed in each available assessment by type and year. "Case" denotes that assessment was conducted using case-specific software, "Col" that the assessment was based on Coleraine, and "SS" that it was based on Stock Synthesis.

Species(stock)	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Shelf species (generally < 200m)																					
Bight redfish, *Centroberyx gerrardi*									Case		SS	SS		SS		SS				SS	
Deepwater flathead, *Neoplatycephalus conatus*									Case		SS	SS			SS		SS	SS			SS
Jackass morwong, *Nemadactylus macropterus* (east)									Col		Col	SS	SS	SS	SS	SS				SS	
Jackass morwong, *Nemadactylus macropterus* (west)													SS	SS	SS	SS				SS	
Redfish, *Centroberyx affinis*							Case												SS		
School whiting, *Sillago flindersi*				Case					Case			SS	SS	SS							
Tiger flathead, *Neoplatycephalus richardsoni*									Case	Case	SS			SS	SS			SS			SS
Slope species (generally 200-700m)																					
Blue grenadier, *Macruronus novaezelandiae*			Case			Case	Case							SS		SS		SS			
Blue warehou, *Seriolella brama* (east)									Case	Case	SS		SS								
Blue warehou, *Seriolella brama* (west)									Case	Case	SS		SS								
Gemfish, *Rexea solandri* (east)					Case							SS	SS								
Ping ling, *Genypterus blacodes* (east)												SS	SS	SS	SS		SS	SS			
Ping ling, *Genypterus blacodes* (west)												SS	SS	SS	SS		SS	SS			
Silver warehou, *Seriolella puncata*					Case		Case					SS	SS	SS			SS			SS	
Deep species (generally >700m)																					
Orange roughy, *Hoplostethus atlanticus* (east)							Case				SS								SS		
Shark species																					
Gummy shark, *Mustelus antarcticus* (Bass Strait)					Case				Case	Case								Case			Case
Gummy shark, *Mustelus antarcticus* (South Australia)					Case				Case	Case								Case			Case
Gummy shark, *Mustelus antarcticus* (Tasmania)																		Case			Case
School shark, *Galeorhinus galeus*	Case				Case								Case	Case							

Table 3. Sensitivity of the estimates from two alternative methods for computing among-assessment variation (σ) in biomass (SSB_{abs}; SSB_{rel}) and sub-sets of data. Refer to Methods Section for description of methods A, B and C.

Data set choice	Method	σ	
		SSB_{abs}	**SSB_{rel}**
All species	A	0.391	0.291
All species	B	0.370	0.176
All species	C	0.596	0.307
Valuable species only	B	0.412	0.185
With consistent fishery-independent data only	B	0.455	0.218
Shelf species	B	0.413	0.170
Slope species	B	0.263	0.195
Shark species	B	0.533	0.109
X=15 years	B	0.360	0.193
X=25 years	B	0.371	0.156
Only Stock Synthesis assessments	B	0.293	0.140

Table 4. Estimates of σ from the meta-analysis (method B) and the asymptotic standard error of terminal year biomass for spawning stock biomass (SSB) in absolute terms, and estimates of σ from the meta-analysis and from the among-assessment variation in estimates of terminal year depletion from sensitivity tests. N/A denotes that either no asymptotic standard error was estimated or no sensitivity tests were conducted for the assessment

Species(stock)	Absolute SSB (SSB_{abs})			Relative SSB (SSB_{rel})	
	Meta-analysis	**Asymptotic value**	**Sensitivity-based**	**Meta-analysis**	**Sensitivity-based**
Shelf species					
Bight redfish	0.702	0.158	0.332	0.100	0.192
Deepwater flathead	0.554	0.234	0.163	0.243	0.116
Jackass morwong (east)	0.218	0.099	0.117	0.159	0.182
Jackass morwong (west)	0.226	0.214	0.432	0.047	0.185
Redfish	0.739	0.155	0.223	0.332	0.251
School whiting	0.320	0.132	0.191	0.240	0.146
Tiger flathead	0.132	0.104	0.181	0.088	0.158
Slope species					
Blue grenadier	0.185	0.147	0.366	0.250	0.277
Blue warehou (east)	0.382	0.216	N/A	0.177	N/A
Blue warehou (west)	0.309	0.261	N/A	0.152	N/A
Gemfish *(east)*	0.267	0.107	0.165	0.199	0.171
Ping ling *(east)*	0.248	0.180	N/A	0.220	N/A
Ping ling *(west)*	0.280	0.219	N/A	0.097	N/A
Silver warehou	0.228	0.135	0.172	0.190	0.078
Deep species					
Orange roughy (east)	0.179	0.093	0.234	0.212	0.309
Shark species					
Gummy shark (Bass Strait)	0.543	0.137	0.247	0.095	0.180
Gummy shark (South Australia)	0.620	0.137	0.175	0.169	0.198
Gummy sharks (Tasmania)	0.086	0.171	0.221	0.012	0.093
School shark	0.602	N/A	0.163	0.078	0.179

Figure

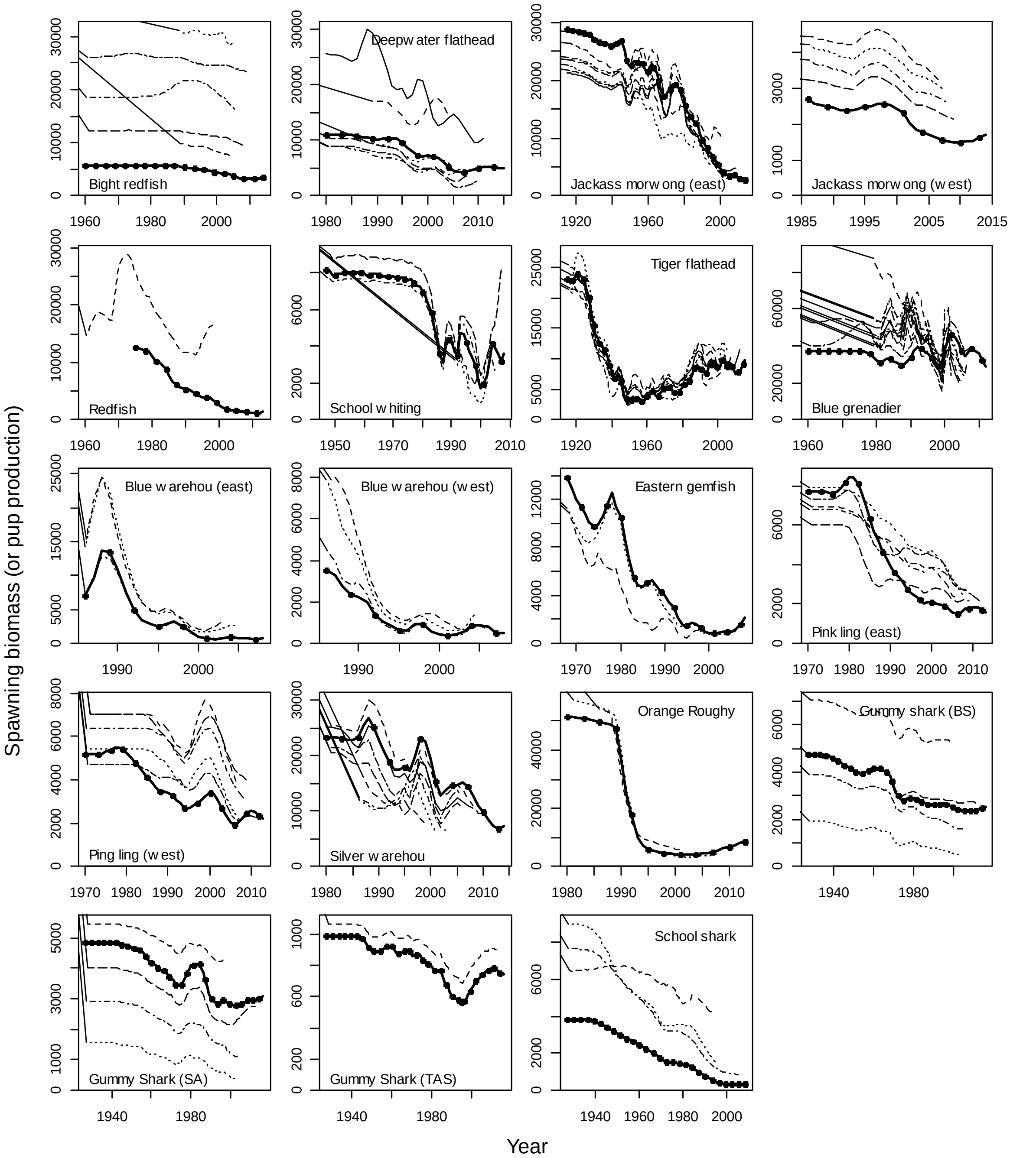

Figure 1. Time trajectories of spawning stock biomass by stock for the assessments considered (Table 2). Solid lines indicate the most recent assessments.

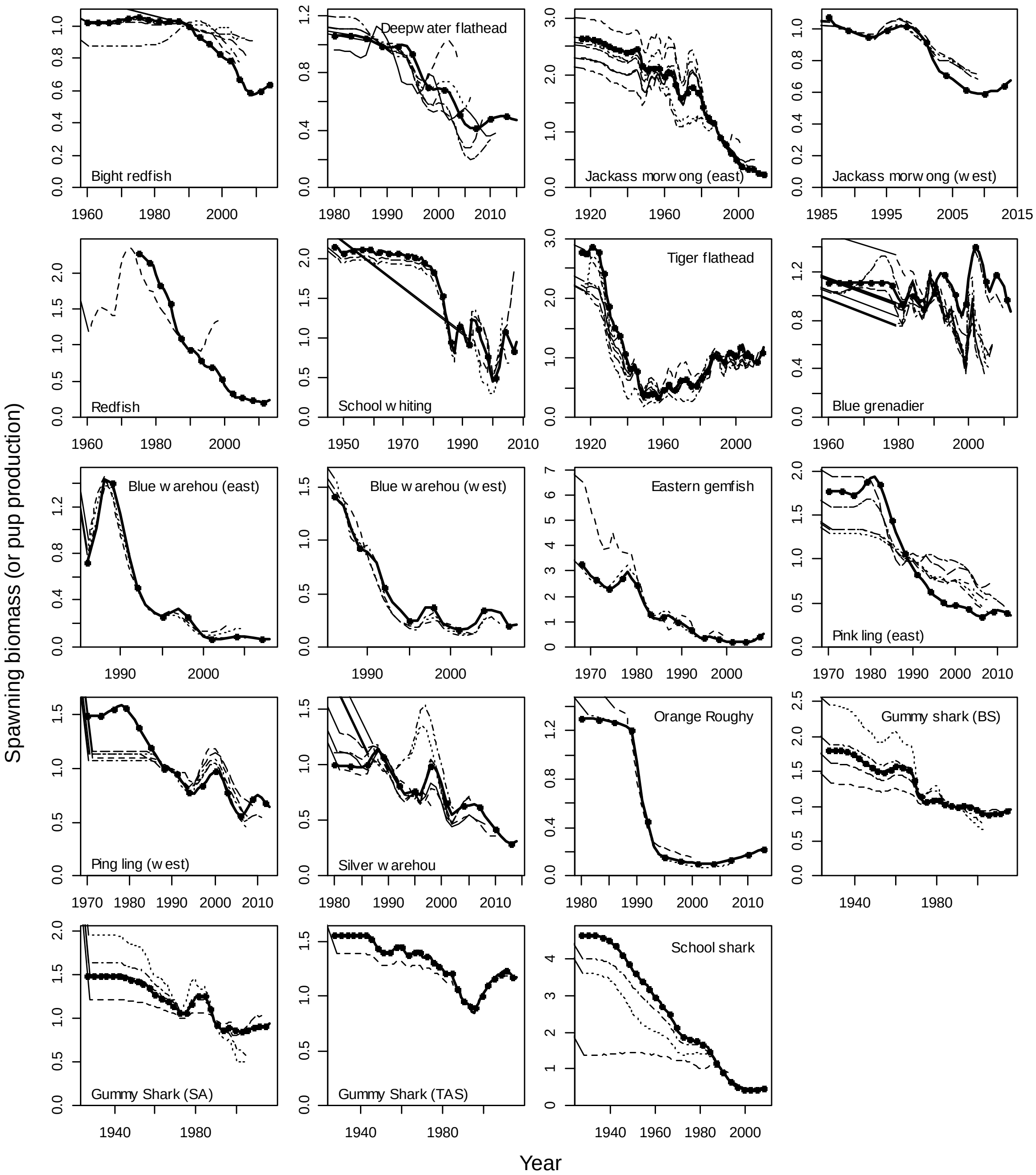

Figure 2. Time trajectories of spawning stock biomass relative to the average spawning stock biomass over 1986-1992 by stock for the assesments considered (Table 2). Solid lines indicate the most recent assessments.

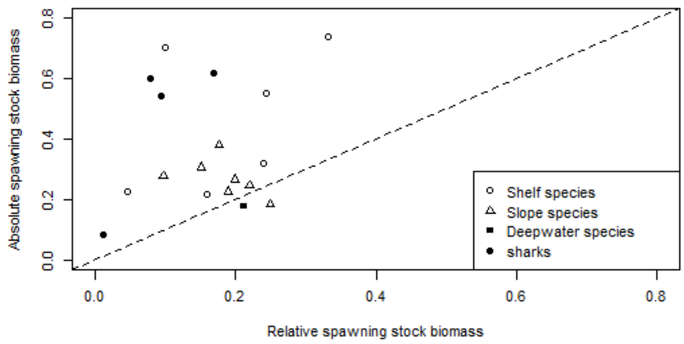

Figure 3. Relationship between the base-case meta-analysis-based values for σ.

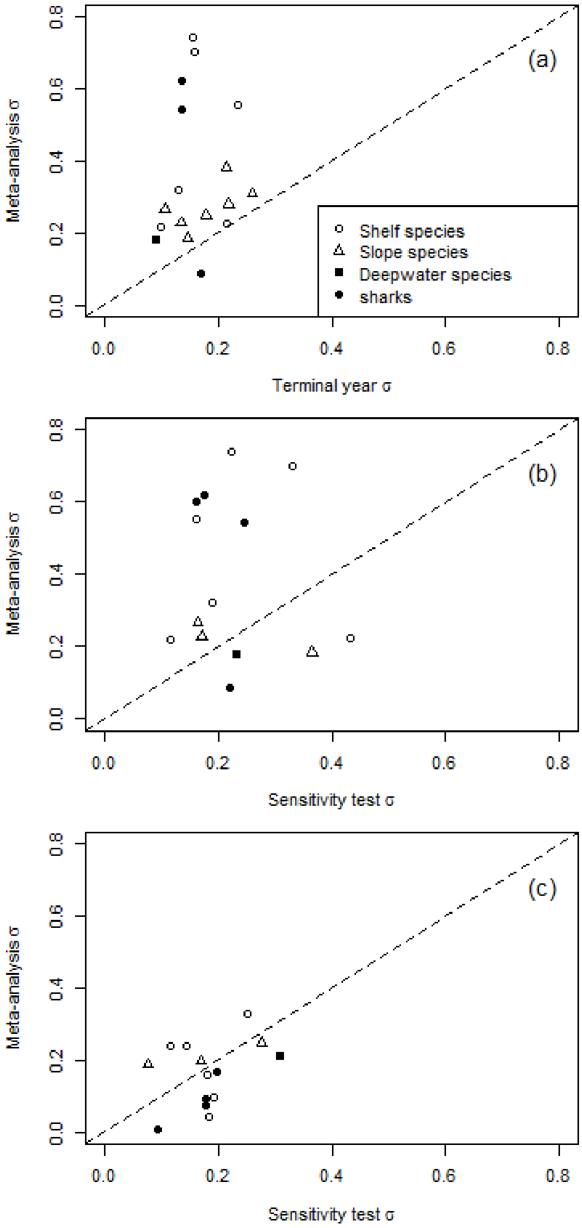

Figure 4. Relationship between estimates of σ between approaches (a,b: absolute spawning stock biomass; c: relative spawning stock biomass).

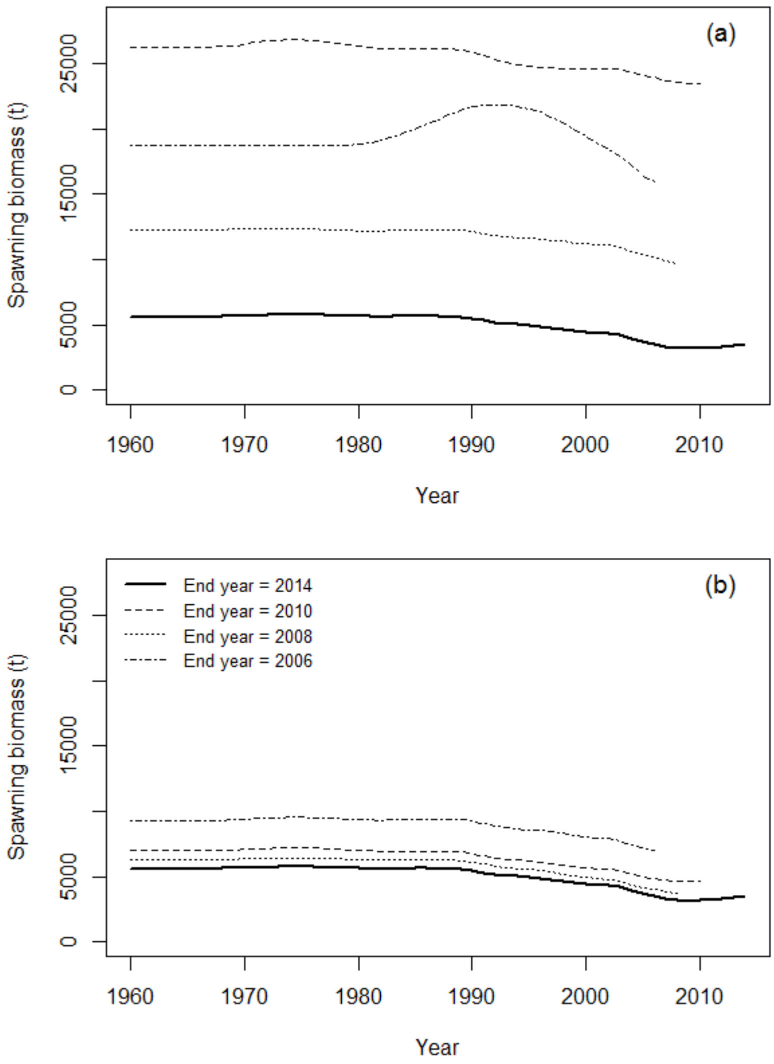

Figure 5. (a) Time-trajectories of spawning stock biomass for Bight redfish from historical assessments, and (b) the equivalent spawning biomass trajectories for a retrospective reduction in data starting with data to the end of the 2014 season. Solid lines indicate the most recent assessments.