A meta-analytic approach to quantifying scientific uncertainty in stock assessments

Stephen Ralston (contact author)[1]

André E. Punt[2]

Owen S. Hamel[3]

John D. DeVore[4]

Ramon J. Conser[5]

Email address for contact author: Steve.Ralston@noaa.gov

[1] National Marine Fisheries Service
Southwest Fisheries Science Center
Fisheries Ecology Division
110 Shaffer Road
Santa Cruz, California 95060

[2] University of Washington
School of Aquatic and Fishery Science
1122 NE Boat Street
Seattle, Washington 98195

[3] National Marine Fisheries Service
Fishery Resource Analysis and Monitoring Division
2725 Montlake Blvd. East
Seattle, Washington 98112

[4] Pacific Fishery Management Council
7700 NE Ambassador Place, Suite 101
Portland, Oregon 97220

[5] National Marine Fisheries Service
Southwest Fisheries Science Center
Fisheries Resource Division
8604 La Jolla Shores Drive
La Jolla, California 92037

Abstract—Quantifying scientific uncertainty when setting total allowable catch limits for fish stocks is a major challenge, but it is a requirement in the United States since changes to national fisheries legislation. Multiple sources of error are readily identifiable, including estimation error, model specification error, forecast error, and errors associated with the definition and estimation of reference points. Our focus here, however, is to quantify the influence of estimation error and model specification error on assessment outcomes. These are fundamental sources of uncertainty in developing scientific advice concerning appropriate catch levels and although a study of these two factors may not be inclusive, it is feasible with available information. For data-rich stock assessments conducted on the U.S. west coast we report approximate coefficients of variation in terminal biomass estimates from assessments based on inversion of the assessment of the model's Hessian matrix (i.e., the asymptotic standard error). To summarize variation "among" stock assessments, as a proxy for model specification error, we characterize variation among multiple historical assessments of the same stock. Results indicate that for 17 groundfish and coastal pelagic species, the mean coefficient of variation of terminal biomass is 18%. In contrast, the coefficient of variation ascribable to model specification error (i.e., pooled among-assessment variation) is 37%. We show that if a precautionary probability of overfishing equal to 0.40 is adopted by managers, and only model specification error is considered, a 9% reduction in the overfishing catch level is indicated.

Manuscript submitted 20 September 2010.
Manuscript accepted 23 February 2011.
Fish. Bull. 109:217–231 (2011).

It has long been recognized that precautionary measures in fisheries management should be related to the amount of uncertainty in the science that is used to evaluate stock status (Caddy and McGarvey, 1996; FAO, 1996). However, few fisheries jurisdictions have adopted precautionary harvest control rules that are designed to reduce "risk-neutral" point estimates of catch based on the amount of uncertainty in the estimates, although at least two examples of this type of precautionary approach exist in the management of marine mammal populations. The International Whaling Commission has adopted a management procedure for baleen whales where, for example, a posterior distribution for the output of a harvest control rule is computed, and the catch limit is set close to the 40th percentile of the distribution (IWC, 1999; Punt and Donovan, 2007). Likewise, with the potential biological removals method (Wade, 1998), the level of marine mammal take at which management action must occur is based on the 20th percentile of the most recent estimate of abundance.

The reauthorization of the Magnuson-Stevens Fishery Conservation and Management Act (MSA) in 2006 changed the requirements for how management actions are developed for U.S. fisheries. The eight Regional Fishery Management Councils are now required to set annual catch limits (ACLs) for all managed stocks that are "in the fishery." National Standard Guidelines have now been developed to assist in the implementation of the reauthorized act (Federal Register, 2009), which defines two sources of uncertainty that must be considered when establishing ACLs: 1) scientific uncertainty, including error pertaining to both the data and to parameter estimation; and 2) management uncertainty, which represents uncertainty in the efficacy of management practices that are designed to ensure that harvest limits are not exceeded. The focus of this study is on the first of these two sources of uncertainty.

Defining "scientific uncertainty" is not trivial. It is therefore not surprising that a variety of approaches have been taken to quantifying uncertainty in fisheries assessments. Methods that aim to quantify the variance of assessment model outputs, given an assumed model structure, include asymptotic statistics, bootstrapping, and the use of Bayesian methods (Hilborn and Walters, 1992; Quinn and Deriso, 1999; Punt and Hilborn, 1997). These techniques are commonly applied in stock assessments, although they all are conditioned on some combination of 1) an assumed model structure; 2) prespecified parameters (e.g., natural mortality); 3) the particular data sets the analyst uses, which may be a subset of those available; and 4) the statistical weights that are assigned to the data elements. Moreover, it is often true that in assessments of data-poor species, more parameters are fixed than in those of data-rich species—a situation that leads to the paradoxical situation where estimates of uncertainty are frequently greater for assessments where more is known (e.g., Pribac et al., 2005). It is also not uncommon that estimated confidence intervals are later shown to have been unrealistically narrow (Stewart and Hamel, 2010).

Uncertainty associated with having selected a particular model from a set of competing models can be assessed by using sensitivity tests, and, in a few cases, model averaging has been used to account for uncertainty due to model structure (e.g., Brandon and Wade, 2006; Brodziak and Piner, 2010). Model averaging is only effective, however, when all selected models are fitted to the same data sets. In principle, Monte Carlo methods can be used to quantify model uncertainty if probabilities can be assigned to the various models and data sets under consideration (e.g., Restrepo et al., 1992). However, these methods are not without their limitations (Poole et al., 1997), and assigning probabilities to, for example, alternative values of a prespecified parameter can be difficult (e.g., Kolody et al., 2008).

The reauthorized MSA and National Standard Guidelines define the overfishing limit (OFL) as the current catch that results from fishing at a rate (F_{MSY}) that is expected to produce the long-term maximum sustainable yield (MSY); catches in excess of the OFL, or fishing mortalities in excess of F_{MSY}, constitute overfishing. Furthermore, the acceptable biological catch (ABC) is the maximum allowable ACL and is defined as a catch which is lower than the OFL to account for scientific uncertainty. On the U.S. west coast, the Pacific Fishery Management Council (PFMC) has adopted a policy of defining the ABC as the product of the OFL and a fractional factor or "buffer" that is based on the probability that the ABC exceeds the true (but unknown) OFL, a value termed P^* (Shertzer et al., 2008; PFMC, 2010). A $P^*=0.5$ is equivalent to fishing at F_{MSY}, with no precautionary reduction to account for scientific uncertainty. Thus, the approach adopted by the PFMC requires the development of an ABC control rule that maps a policy decision ($P^*<0.5$) to a buffer that is used to reduce the OFL to an ABC.

We outline and apply the approach developed by members of the Scientific and Statistical Committee of the PFMC to calculate these factors for groundfish and coastal pelagic species on the basis of results from historical analyses. With a historical analysis, we summarize the results of all the assessments that have been conducted for a particular stock. Importantly, repeat assessments conducted for the PMFC often incorporate a variety of changes that include many of the model specification problems identified above. Although our approach is purely empirical and somewhat *ad hoc*, it is a pragmatic way to address the new legislative requirement to account for scientific uncertainty and to set precautionary catch limits. It was formally adopted by the PFMC for use in setting total allowable groundfish catches for the 2011–12 biennium (PFMC, 2010).

Materials and methods

Sources of uncertainty

Calculation of an OFL typically involves three steps: 1) estimation of current exploitable biomass (B_t); 2) projection of the population biomass into the future for some number of years; and 3) application of an estimate of F_{MSY} to the forecasts of future biomass. Although there are clear uncertainties associated with each step, the Scientific and Statistical Committee elected to focus first and foremost on variation in the estimation of the biomass in the terminal year of groundfish and coastal pelagic species stock assessments. That biomass is a significant source of uncertainty is aptly illustrated in Figure 1, which shows the results of the 15 Pacific whiting (*Merluccius productus*) stock assessments that have been conducted for the PFMC over the last 18 years (Stewart and Hamel, 2010). It is instructive to examine this species because it is one of the most data-rich[1] stocks managed by the PFMC, is of substantial economic importance, and has been assessed largely on an annual basis for many years. However, estimates of biomass have been highly variable from a historical perspective, in spite of considerable scientific resources having been devoted to evaluating the status of this stock. Note, for example, that estimated spawning biomasses in 1985 ranged from 1.2 to 5.9×10^6 metric tons (t) over the 15 stock assessments, representing a 5-fold range in abundance.

There are many reasons for this type of "among" assessment variability in stock size estimates, including differences in 1) overall model structure; 2) altered fixed values and prior distributions for important parameters; 3) changes in the availability of data; 4) the composition of the review panel; 5) the makeup of the analytical team that conducted the assessment; and 6) the modeling software that was used. Impor-

---

[1] Data-rich stock assessments contain many informative data elements, which typically would include catch (landings+discards), life history information (growth, natural mortality, and reproductive parameters), annual age or length compositions sampled from the fishery, and trend indices.

tantly, these factors contribute to variation in all groundfish and coastal pelagic species stock assessments at the PFMC, collectively exhibit considerable variation among historical assessments. Moreover, it is unsettling to managers when stock size estimates fluctuate greatly from one assessment to the next because this fluctuation undermines confidence for scientific advice. Hence, we assert that quantifying and accounting for this source of uncertainty is the first and most important factor to consider when establishing a buffer between the OFL and the ABC. We recognize, however, that as the quantification of scientific uncertainty develops in the future it will be important to expand consideration to other sources of errors, including forecast uncertainty (Shertzer et al., 2008) and uncertainty in estimating optimal harvest rates (e.g., Dorn, 2002; Prager et al., 2003; Punt et al., 2008). Hence, quantification of variation as revealed here should be considered only a lower bound on total uncertainty. Moreover, even if both forecast and harvest rate uncertainty were incorporated into our analysis, we note that many other factors exist that would be difficult to quantify, including the effects of climate and ecosystem interactions on the estimation of OFLs.

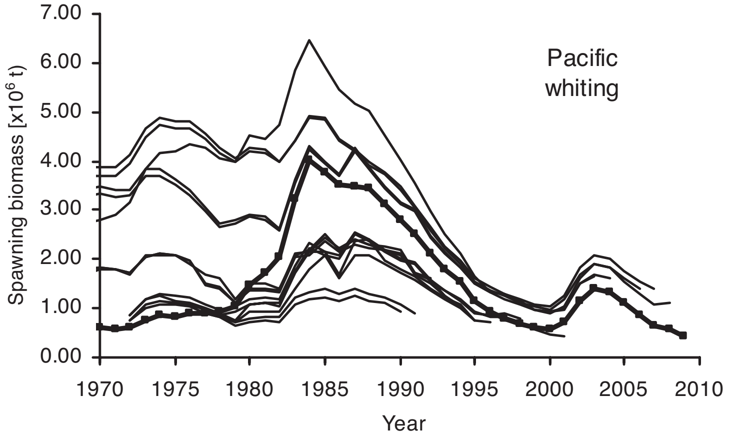

Figure 1

Biomass time series for Pacific whiting (*Merluccius productus*) based on 15 historical stock assessments conducted for the Pacific Fishery Management Council. The bold line with square symbols represents the most recent stock assessment used in the meta-analysis; the other lines represent time series of abundance developed from earlier assessments.

Quantifying biomass uncertainty

We initially consider two types of uncertainty in biomass estimation. The first is due to estimation error, also termed stochastic uncertainty (Pawitan, 2001). We quantify this type of uncertainty using the estimated coefficient of variation (CV) for the terminal-year biomass taken from the most recent stock assessment conducted. In a very limited number of studies (e.g., Pacific ocean perch [*Sebastes alutus*]), full Bayesian integration of uncertainty with Monte Carlo Markov Chain analysis has been achieved. However, on the U.S. west coast such cases are the exception. Hence, we report the asymptotic standard error for the estimate of terminal biomass developed by inverting the model's Hessian matrix as a first-order approximation of variation, i.e., the observed Fisher information statistic (Pawitan, 2001). The accuracy of this approximation depends on how well the log-likelihood surface at its maximum can be approximated by quadratic curvature, and on proper specification of the likelihood components, including appropriate error distributions and variance weightings.

We view this error estimate as a measure of statistical uncertainty within a stock assessment model that is conditioned on all the structural assumptions embedded within the model. We convert the asymptotic standard error to a CV by simple division using the terminal biomass estimate as the denominator. It is important to note that we limit our consideration to terminal-year biomass because under the reauthorized MSA, quantification of scientific uncertainty is used to prevent "overfishing," which occurs when 1) the current year catch exceeds the OFL; or 2) an updated assessment retrospectively indicates that fishing mortality exceeded F_{MSY}. Overfishing *per se* occurs only on an annual basis, although the chronic effect of overfishing results in stocks becoming depleted, and if fishing mortality is substantially greater than F_{MSY}, a stock will eventually become overfished.

The second type of uncertainty can be thought of as among-assessment variation, which is attributable to a wide variety of factors, many of which represent a significant form of model or inductive uncertainty (Pawitan, 2001). Assertion of asymptotic or dome-shaped selectivity patterns is one example, as is incorporation of age-dependent natural mortality. Assumptions regarding such structural issues will often change from one assessment to the next. Likewise, values for biologically important parameters (e.g., natural mortality or spawner-recruit productivity), which are prespecified when using auxiliary information (or expert judgment), may change, or an entire new data series may be incorporated into the assessment as new data become available. Beyond such changes in model specification, among-assessment variation includes other sources of variability due to, for example, differences in the reviewers who evaluated, suggested changes to, and ultimately approved an assessment model.

To quantify among-assessment variability we assembled time series of biomass from historical assessments of groundfish and coastal pelagic species stocks. We excluded updated assessments, where data were simply refreshed and not extensively reviewed, because of strong constraints imposed on how much they could change from the last comprehensive assessment (PFMC[2]). When the definition of biomass changed among the available assessments (e.g., mid-year biomass in one assessment and beginning-year biomass in another), we used ratio estimation (Cochran, 1977) over a common time period to standardize to a common metric across all assessments that were conducted for that stock. We also limited the data points under consideration to no more than those that represent the last 20 years reported in the most recent assessment to focus attention on variation associated with the estimation of terminal year biomass. Finally, we trimmed the time series to include only the most recent 15, 10, and 5 years to evaluate the stability of the estimates of among-assessment uncertainty in relation to time interval selection criteria.

Variation in biomass estimates among a set of stock assessments can be quantified in a number of ways. We evaluated three approaches to calculating variation around a point of central tendency:

1 Consider all biomass estimates for a year as equally plausible representations of reality. Biomass variation between two stock assessments was quantified by forming all possible ratios of estimated biomasses in common years. Specifically, if there was an estimate of biomass (B) for year t from assessments i and j, we calculated: $R_{i|j,t} = B_{i,t}/B_{j,t}$, i.e., the proportional deviation of assessment i using assessment j as a standard. Based on a symmetry argument, we also calculated $R_{j|i,t}$ and all the ratios were $\log_e$-transformed. Note that because $\ln(R_{i|j,t}) = -\ln(R_{j|i,t})$, the distributions were perfectly symmetrical. For each stock under consideration the standard deviation (σ^*) of the ratios was calculated. This statistic is positively biased, however, because it is based on the ratio of two lognormal random variables ($B_{i,t}$ and $B_{j,t}$). The appropriate bias correction term ($\sqrt{2}$) was derived (Mohr[3]) and applied so that the corrected estimator is $\sigma = \sigma^*/\sqrt{2}$. Thus, in the first approach we used the bias-adjusted estimate of the standard deviation of the $\ln(R_i|_{j,t})$ as a quantitative measure of among-assessment variation.

2 Consider the mean of biomass estimates in a year as the best estimate of central tendency. In this approach, variation in biomass was measured as squared deviations from the annual mean in log-space. Specifically, we calculated the mean log-biomass in year t as:

$$\overline{\ln[B_t]} = \frac{1}{n_t}\sum_i \ln[B_{i,t}],$$

where n_t is the number of available assessments in year $t(n_t \geq 2)$. The standard deviation (σ) is then calculated as follows:

$$\sigma = \sqrt{\frac{1}{\sum_t n_t - 1}\sum_t \sum_i \left(\ln[B_{i,t}] - \overline{\ln[B_t]}\right)^2}.$$

3 Consider the most recent stock assessment as the best estimate of central tendency. This approach is the same as the second, except that the mean ($\overline{\ln[B_t]}$) is replaced by the logarithms of the biomass estimates from the most recent stock assessment, and the most recent year is excluded from the summations and the calculation of the n_t. With this approach, the most current information is assumed to represent the best estimate of the population mean.

For lognormally distributed random variables, the CV on the arithmetic scale is equal to

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

where σ^2 is the variance on the logarithmic scale (Johnson and Kotz, 1970). We used this relationship to convert variances on the logarithmic scale to the arithmetic scale for comparison.

Meta-analytic inference for management

The PFMC groundfish fishery management plan includes approximately 90 species and, with the exception of "ecosystem component" species and stock complexes, OFLs, ABCs, and ACLs need to be developed for them all. However, less than 30% of the stocks listed in the fishery management plan have been assessed. Even among stocks that have been assessed, several have been studied only once. Therefore, historical biomass variation among assessments cannot routinely be estimated on a stock-specific basis. Thus, there is some merit in pooling results from well-studied species to develop estimates of meta-analytic proxy variance for all groundfish and coastal pelagic species stocks, and potentially even for those that have been assessed multiple times.

Based on management practices at the PFMC there are four natural groupings of species to consider, i.e., rockfish, roundfish, flatfish, and coastal pelagic species. The first three are groundfish categories that have group-specific proxy F_{MSY} harvest rates (Dorn, 2002; Ralston, 2002), whereas coastal pelagic species are managed in a separate fishery management plan. We considered two methods of pooling stock-specific variances: 1) take the square root of the average of the stock-specific variances; and 2) aggregate all the

---

[2] PFMC. 2008. Terms of reference for the groundfish stock assessment and review process for 2009_2010, 35 p. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, Oregon 97220-1384. [Available at: http://www.pcouncil.org/wp-content/uploads/GF_Stock_Assessment_TOR_2009-102.pdf.]

[3] Mohr, Michael S. 2009. Groundfish ABC accounting for scientific uncertainty derivation of biomass scalar, 4 p. Unpubl. document submitted to Pacific Fishery Management Council Scientific and Statistical Committee. Author's address: NMFS, SWFSC, 110 Shaffer Rd., Santa Cruz, CA 95060.

residuals and calculate the standard deviation of the pooled set. The first method gives each species equal weight and does not overemphasize stocks that have been assessed many times (e.g., Pacific whiting). Conversely, the second method treats each data point as an independent observation. Neither approach is ideal given the lack of independence in the data.

Results

Most of the groundfish and coastal pelagic species stock assessments that have been conducted for the PFMC have employed the stock synthesis framework (Methot, 2000), which provides a very flexible, integrated modeling environment. For our analysis we considered only data-rich stocks that have been assessed more than once (15 groundfish and two coastal pelagic species stocks)—an approach that excluded many species from consideration. Owing to the large number of citations (81) needed to fully document the assessment literature of these stocks, all of which appear in the stock assessment and fishery evaluation documents produced by the PFMC, we present only summary information for each assessment that includes the stock, year, and authorship (Table 1)[4].

There is a preponderance of rockfish among the 17 species analyzed. The number of assessments included in the meta-analysis ranged from two (chilipepper [*Sebastes goodei*]) to a high of fifteen (Pacific whiting) (Table 2). Results presented in Figure 2 (A and B) show biomass trajectories from 1970 through 2009 for the 16 stocks that were not whiting stocks. Note that there is good correspondence among assessments for some species (e.g., darkblotched rockfish [*Sebastes crameri*]) and poor correspondence for others (e.g., shortspine thornyhead [*Sebastolobus alascanus*]). One should be cautious in interpreting this correspondence to indicate the degree of uncertainty in stock biomass because random variation in correspondence between species is to be expected. In the case of shortspine thornyhead, new information indicating dome-shaped selectivity was largely responsible for the large change in the biomass estimates for the 2005 assessment.

Comparisons of methods

When the assessment data were restricted to the last twenty years, the three approaches (all ratio combinations, deviations from the mean, and deviations from the most recent assessment) yielded average estimates of σ over all stocks equal to 0.382, 0.337, and 0.307, respectively. Approach two (i.e., squared deviations from the mean in log-space) was selected as the preferred method for calculating uncertainty by the Scientific and Statistical Committee because it had two desirable features, i.e., deviations were calculated from the best estimate of central tendency and estimated values of σ were unlikely to change markedly with new assessments (unlike approach three, which relies on the most recent assessment as the reference). Coincidently the calculation produced an intermediate result among the three approaches.

Similarly, a sensitivity test of the results to the number of years included in the calculation revealed that estimates of σ were robust to the time period used in the calculation. For example, when only the last 15 years of data were used, σ was 0.338 (compared with 0.337 for 20 years). Likewise, when the final 10 and 5 years of data were used, estimates of σ were 0.371 and 0.344, respectively. Note that in these latter two cases some species were excluded because of sparseness of data for these species. Hence, a standard temporal window equal to the last 20 years of assessment data was adopted as the basis for quantifying variation among stock assessments.

Stock-specific results

Figure 3 shows the distributions of residuals for the 17 stocks based on the selected approach. Note that some species (e.g., chilipepper and shortspine thornyhead) show a strongly bimodal distribution—a pattern that results when few assessments are available and biomass trajectories do not intersect. However, most of the distributions are unimodal, generally symmetric, and centered on or near zero.

Table 2 presents the number of deviations and the estimated log-scale standard deviation (σ) for each of the stocks, which collectively ranged from 0.103 (darkblotched rockfish) to 0.923 (shortspine thornyhead) with an average of 0.337. Also presented in the table are the estimated asymptotic coefficients of variation (CVs) for terminal biomass from the most recently completed stock assessment. These CVs, which approximate within-assessment estimation error, ranged from 9% (shortspine thornyhead and Dover sole [*Microstomus pacificus*]) to 41% (Pacific sardine [*Sardinops sagax*]), with a mean of 18%. This is undoubtedly an underestimate, however, because of the presence of key fixed parameters (e.g., natural mortality) in almost all the assessments we reviewed.

To compare among-assessment variation to within-assessment variation, the log-scale standard deviation estimates were expressed as CVs on the arithmetic scale (Johnson and Kotz, 1970), and the two statistics were plotted against one another (Figure 4). It is evident that shortspine thornyhead is an outlier, with the lowest “within” CV (stochastic uncertainty) and the highest “among” CV (inductive uncertainty). As a rule, among-assessment CVs (mean=36%) were greater than within-assessment CVs, as evidenced by the preponderance of points falling to the right side of the line of equality.

---

[4] Individuals should contact the PFMC (John.DeVore@noaa.gov) or visit http://www.pcouncil.org/groundfish/stock-assessments/ or http://www.pcouncil.org/coastal-pelagic-species/stock-assessment-and-fishery-evaluation-safe-documents/ for copies of specific assessment documents.

Table 1

Groundfish and coastal pelagic species stock assessments used in quantifying historical retrospective biomass variation.

Stock group	Species	Year	Author(s)[1]
Rockfish	bocaccio (*Sebastes paucisipinis)*	1996	Ralston et al.
		1999	MacCall et al.
		2002	MacCall
		2003	MacCall
		2009	Field et al.
	canary rockfish (*Sebastes pinniger)*	1994	Sampson and Stewart
		1996	Sampson
		1999	Crone et al.
		1999	Williams et al.
		2002	Methot and Piner
		2005	Methot and Stewart
		2008	Stewart
		2009	Stewart
	chilipepper (*Sebastes goodei)*	1998	Ralston et al.
		2008	Field
	darkblotched rockfish (*Sebastes crameri*)	2003	Rogers
		2005	Rogers
		2009	Wallace and Hamel
	Pacific ocean perch (*Sebastes alutus*)	1992	Ianelli et al.
		1998	Ianelli and Zimmerman
		2009	Hamel
	widow rockfish (*Sebastes entomelas*)	1997	Ralston and Pearson
		2000	Williams et al.
		2003	He et al.
		2006	He et al.
		2009	He et al.
	yelloweye rockfish (*Sebastes ruberrimus*)	2001	Wallace
		2002	Methot et al.
		2006	Wallace et al.
		2009	Stewart et al.
	yellowtail rockfish (*Sebastes flavidus*)	1991	Tagart
		1993	Tagart
		1996	Tagart and Wallace
		1997	Tagart et al.
		2000	Tagart et al.
		2005	Wallace and Lai
	shortspine thornyhead (*Sebastolobus alascanus*)	1994	Ianelli et al.
		2001	Piner and Methot
		2005	Hamel
Roundfish	cabezon (*Scorpaenichthys marmoratus*)	2004	Cope et al.
		2006	Cope and Punt
		2009	Cope and Key
	Pacific whiting (*Merluccius productus*)	1991	Dorn and Methot
		1992	Dorn and Methot
		1993	Dorn et al.
		1994	Dorn
		1995	Dorn
		1996	Dorn
		1997	Dorn and Saunders
		1999	Dorn et al.
		2002	Helser et al.

continued

Table 1 (continued)

Stock group	Species	Year	Author(s)[1]
Roundfish *(continued)*	Pacific whiting (*Merluccius productus*)	2004	Helser et al.
		2005	Helser et al.
		2006	Helser et al.
		2007	Helser and Martell
		2008	Helser et al.
		2009	Hamel and Stewart
	lingcod (*Ophiodon elongatus*)	2000	Jagielo et al.
		2003	Jagielo et al.
		2005	Jagielo and Wallace
		2009	Hamel et al.
	sablefish (*Anoplopoma fimbria*)	1992	Methot
		1994	Methot et al.
		1997	Crone et al.
		1998	Methot et al.
		2001	Schirripa and Methot
		2005	Schirripa and Colbert
		2007	Schirripa
Flatfish	Dover sole (*Microstomus pacificus*)	1997	Brodziak
		2001	Sampson and Wood
		2005	Sampson
	petrale sole (*Eopsetta jordani*)	1999	Sampson and Lee
		2005	Lai et al.
		2009	Haltuch and Hicks
Coastal pelagic	Pacific mackerel (*Scomber japonicus*)	2004	Hill and Crone
		2005	Hill and Crone
		2007	Dorval et al.
		2009	Crone et al.
	Pacific sardine (*Sardinops sagax*)	2004	Conser et al.
		2007	Hill et al.
		2009	Hill et al.

[1] Individuals should contact the PFMC (John.DeVore@noaa.gov) for copies of specific assessment documents or visit "http://www.pcouncil.org/groundfish/stock-assessments/" or "http://www.pcouncil.org/coastal-pelagic-species/stock-assessment-and-fishery-evaluation-safe-documents/".

Pooled results

Figure 5 shows the unweighted, pooled distributions of residuals for the four groupings of stocks. The distribution for rockfish is close to normal, whereas those for roundfish and flatfish exhibit some non-normal features. For example, the distribution for coastal pelagic species exhibits a tail to the left. However, the pooled estimates of σ are all rather similar, regardless of the method used to aggregate the data (Table 3). Although to some degree the point estimates of σ differ among groups, the sample sizes are also rather small. To explore whether the data provide support for treating each group separately, estimates of σ^2 were fitted by using a linear mixed model with group as a random effect. That analysis provided no support for stratification by group; the point estimate of the between-group variance was essentially zero ($<10^{-5}$).

Given the lack of support for among-group variation in σ, we combined the data over all stocks. In this instance, because the need to treat each species as a replicate is not required, method 2 (simple pooling of all residuals) is most justified. Aggregating the deviations over all stocks (Fig. 6) led to a point estimate of $\sigma=0.358$. If the residuals are assumed to be independent, an approximate 95% confidence interval for the statistic is $0.342 \leq \sigma \leq 0.374$.

Discussion

We evaluated three approaches for quantifying scientific uncertainty in the groundfish and coastal pelagic stock assessments that have been conducted over the last 20 years for the PFMC. We conclude that measurement of log-scale variability as deviations from the mean is a suitable analytical approach for measuring uncertainty in biomass estimates. Moreover, a comparison of stock- and group-specific estimates indicated that a single value of $\sigma=0.36$, which is equivalent to a CV of 37% on the arithmetic scale, is a reasonable lower-bound

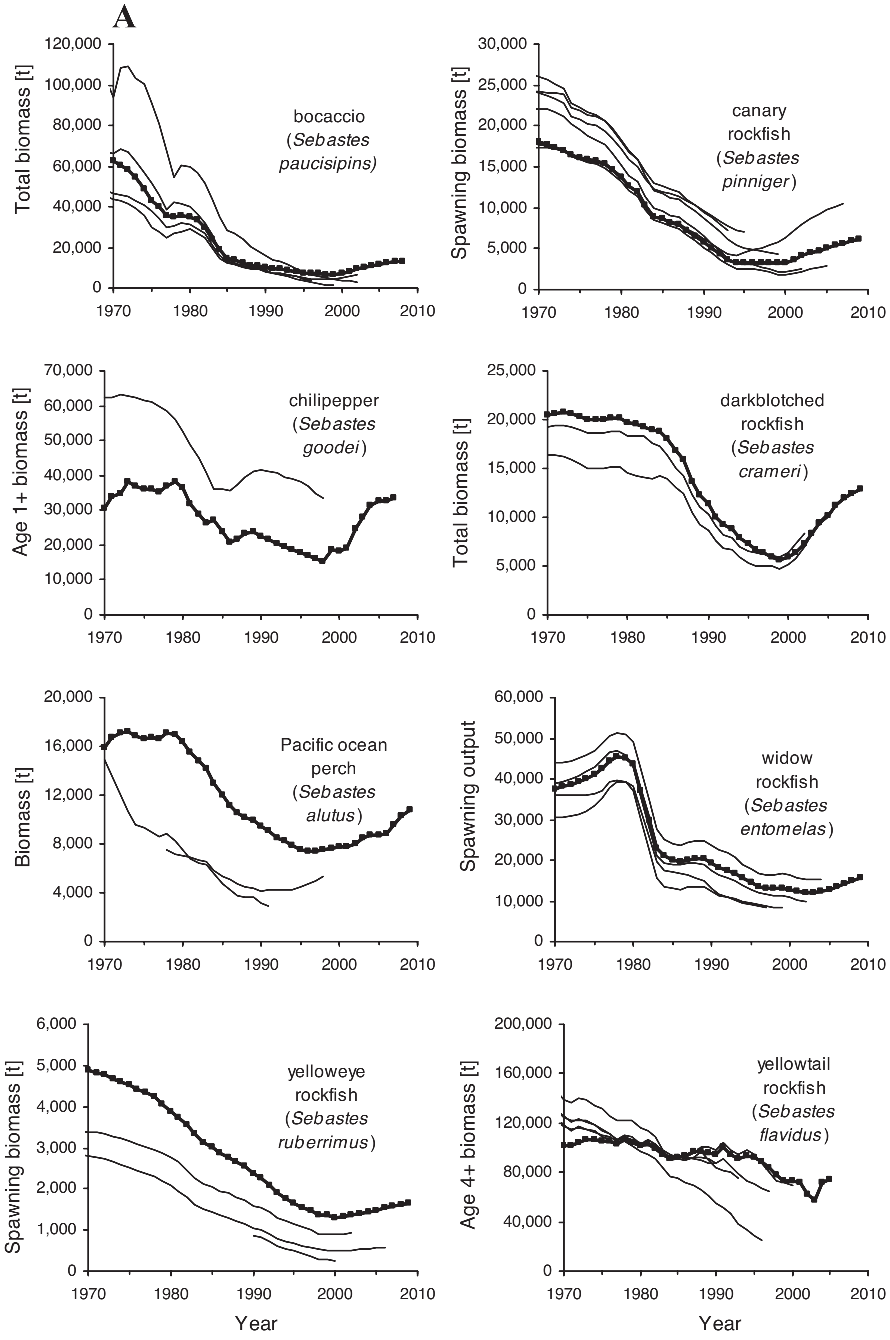

Figure 2

(**A** and **B**) Biomass time series for 16 selected groundfish and coastal pelagic species from stock assessments conducted for the Pacific Fishery Management Council on the west coast of the United States. In each panel the bold line with square symbols represents the most recent stock assessment that was completed, whereas the lighter lines represent biomass time series from earlier assessments.

proxy for all groundfish and coastal pelagic species stocks. Among all the 17 stocks listed in Table 2, only Pacific sardine yielded a Hessian-based "within" CV that is greater (41%). On average, variation among stock assessments was about twice that of the estimation error within assessments (18%).

Our approach is empirical and we simply strive to quantify variation in biomass estimates from repeats

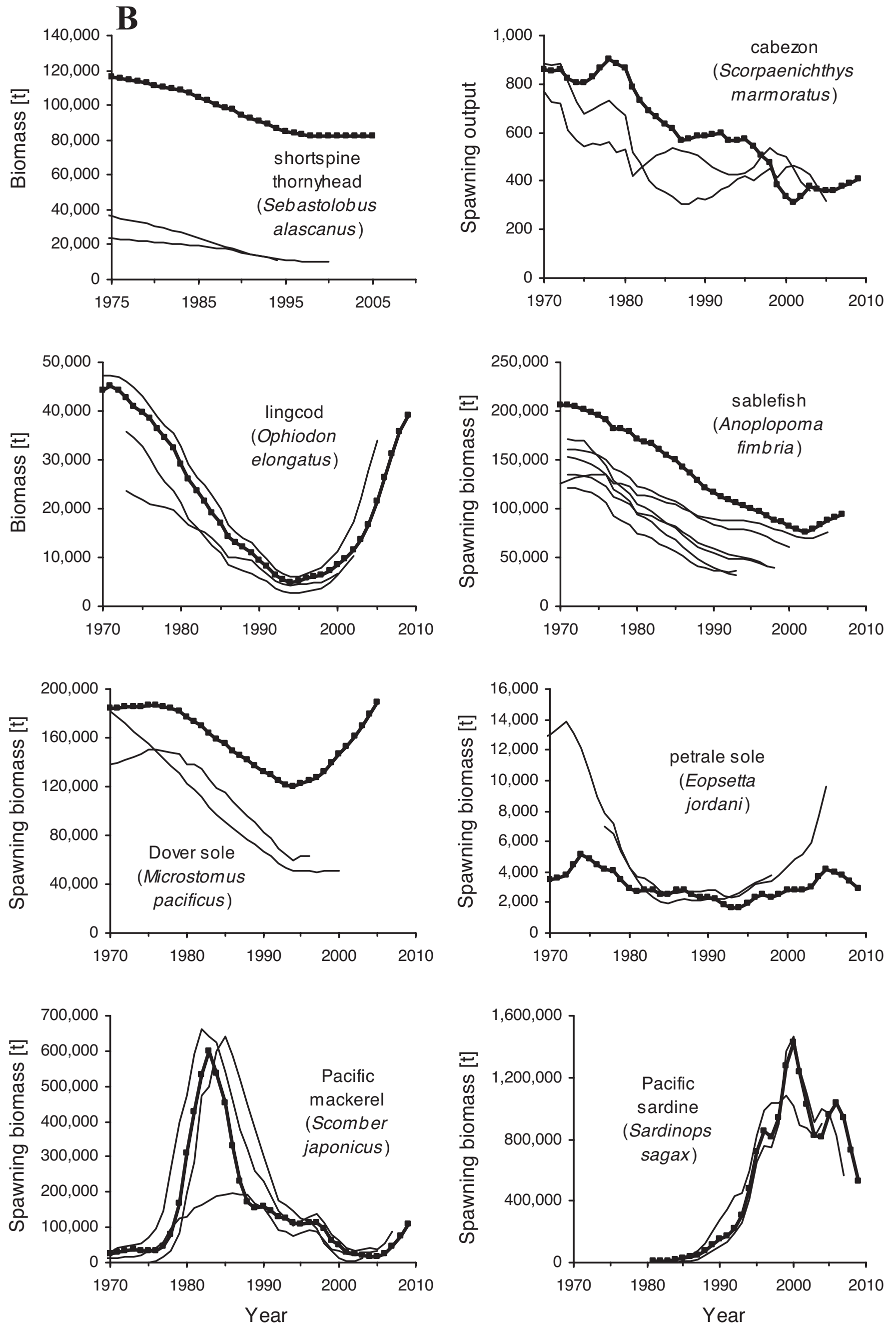

Figure 2 (continued)

of data-rich stock assessments that have been conducted for the PFMC. Although the approach lacks a theoretical basis, the method incorporates many of the factors that lead to model uncertainty, which we have shown is much greater than within-model estimation errors. One concern with the analysis is that calculation of uncertainty as squared deviations from the mean (approach 2) includes the assumption of the independence of the residuals, which is surely violated given that repeats of an assessment provide much the same data. Likewise, our findings pertain strictly to groundfish and coastal pelagic species found off the U.S. west coast. To the extent that the availability of data and the assessment "culture" in that region is distinctive (e.g., the use of the Stock Synthesis modeling platform and a willingness to adopt meta-analytic results as proxy metrics), our specific findings may not be of general use elsewhere. Even so, our general

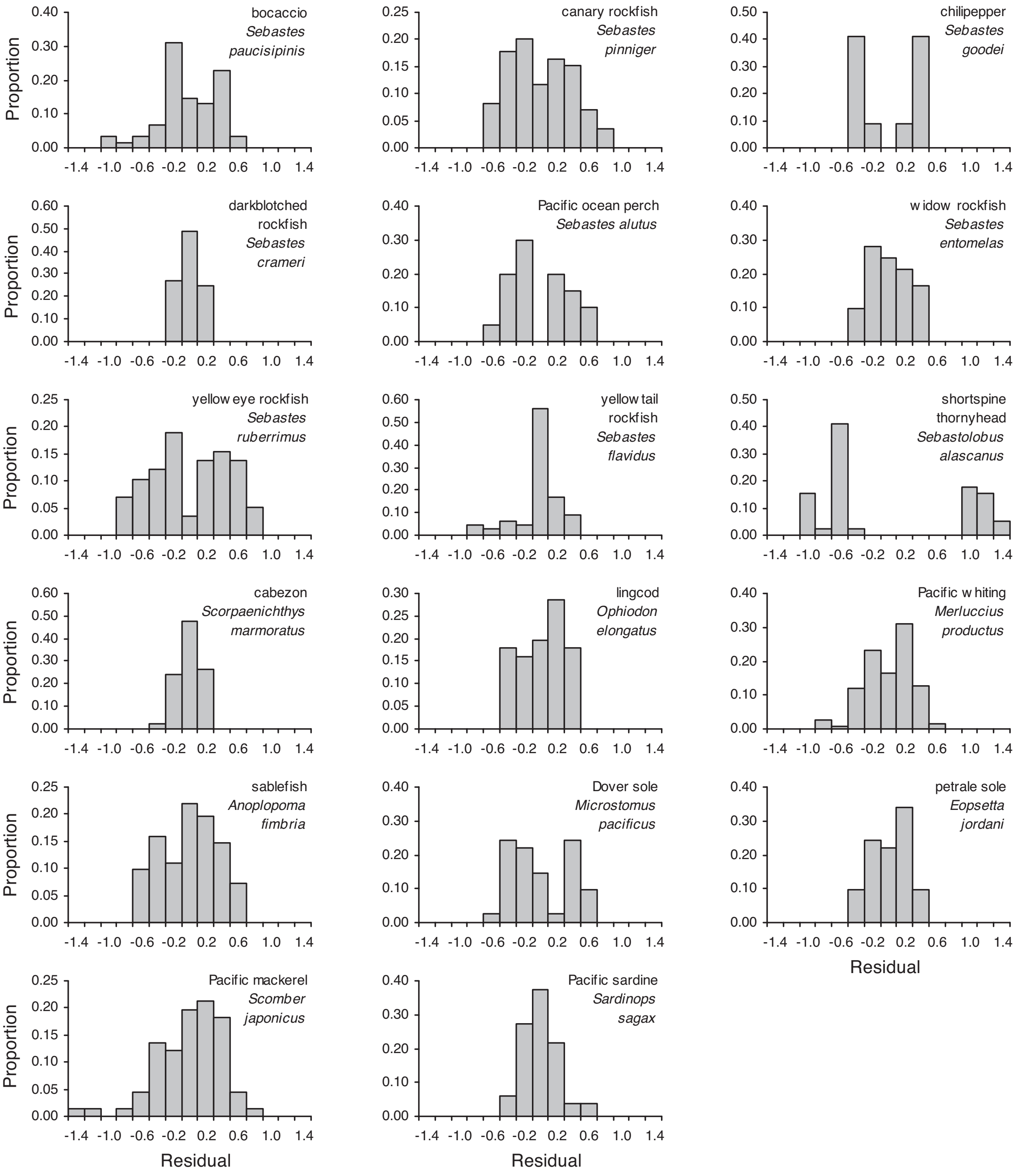

Figure 3

Frequency distributions of log-scale biomass deviations for selected groundfish and coastal pelagic species in stock assessments conducted for the Pacific Fishery Management Council. Residual deviations were calculated from annual means taken from the biomass time series presented in Figure 2 (**A** and **B**).

approach may prove useful for quantifying biomass uncertainty if it is applied in other geographical or management regions.

To illustrate how an estimate of log-scale variance can be used to form the basis of an ABC control rule, we exponentiate a lognormal distribution with a mean equal to zero and σ=0.36. Half of the probability density is then below a value of 1.00, which represents the median of the distribution. One can then select a cumulative probability less than 0.50 that maps onto a multiplier (=buffer) that can be interpreted as a reduction from the median of the distribution (Fig. 7). For example, 40% of the probability density is found at values ≤0.913. If one assumes that the median of

Table 2

Summary of stock-specific analyses of variation for estimates of terminal stock size from assessments of groundfish and coastal pelagic species. CV=coefficient of variation.

Stock group	Common name	Scientific name	No. of stock assessments	Squared deviations (*n*)	Log-scale standard deviation	Statistical uncertainty CV
Rockfish	bocaccio	*Sebastes paucispinis*	5	61	0.367	15%
	canary rockfish	*Sebastes pinniger*	8	85	0.375	15%
	chilipepper	*Sebastes goodei*	2	22	0.354	14%
	darkblotched rockfish	*Sebastes crameri*	3	45	0.103	13%
	Pacific ocean perch	*Sebastes alutus*	3	20	0.352	15%
	widow rockfish	*Sebastes entomelas*	5	61	0.241	31%
	yelloweye rockfish	*Sebastes ruberrimus*	4	58	0.492	14%
	yellowtail rockfish	*Sebastes flavidus*	6	66	0.269	24%
	shortspine thornyhead	*Sebastolobus alascanus*	3	39	0.923	9%
Roundfish	cabezon	*Scorpaenichthys marmoratus*	3	46	0.154	21%
	lingcod	*Ophiodon elongatus*	4	56	0.263	10%
	Pacific whiting	*Merluccius productus*	15	151	0.286	28%
	sablefish	*Anoplopoma fimbria*	7	82	0.340	10%
Flatfish	Dover sole	*Microstomus pacificus*	3	41	0.360	9%
	petrale sole	*Eopsetta jordani*	3	41	0.227	15%
Coastal pelagic	Pacific mackerel	*Scomber japonicus*	4	66	0.415	25%
	Pacific sardine	*Sardinops sagax*	3	51	0.206	41%

Table 3

Comparison of different methods of pooling stock-specific variance estimates. Method 1 weights each species equally, whereas method 2 weights each data point equally. In the table, σ is the standard deviation of log-scale anomalies from the mean.

Group	Number of stocks	σ Method 1	σ Method 2
rockfish	9	0.442	0.418
roundfish	4	0.269	0.281
flatfish	2	0.301	0.299
coastal pelagic	2	0.328	0.339
All stocks	17	0.337	0.358

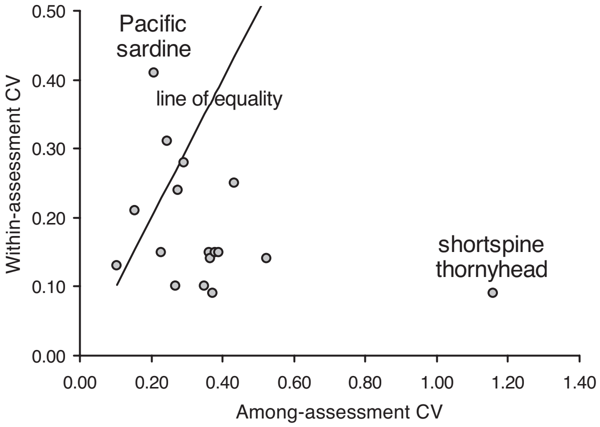

Figure 4

Relationship between the coefficients of variation (CV) calculated from biomass variation over multiple full stock assessments (x axis) and the CV based on the measurement error of the most recent analysis (y axis).

the lognormal distribution (1.00) is indicative of the best risk-neutral point estimate of catch (=OFL), 91% of that amount would be associated with a 0.40 probability of exceeding the true OFL.

Other approaches and future work

The approach outlined in this study is a pragmatic way to address the legislative requirement to calculate ABCs from OFLs, accounting for scientific uncertainty. Although the approach has been adopted and implemented as an ABC control rule for decision-making at the PFMC,[5] quantification of scientific uncertainty

---

[5] PFMC and NMFS. 2010. Proposed harvest specifications and management measures for the 2011–2012 Pacific Coast groundfish fishery and Amendment 16-5 to the Pacific Coast Groundfish Fishery Management Plan to update existing rebuilding plans and adopt a rebuilding plan for petrale sole: Draft environmental impact statement including Regulatory Impact Review and Initial Regulatory Flexibility Analysis. Pacific Fishery Management Council, Portland, OR (submitted to NOAA Fisheries Service), June 2010.

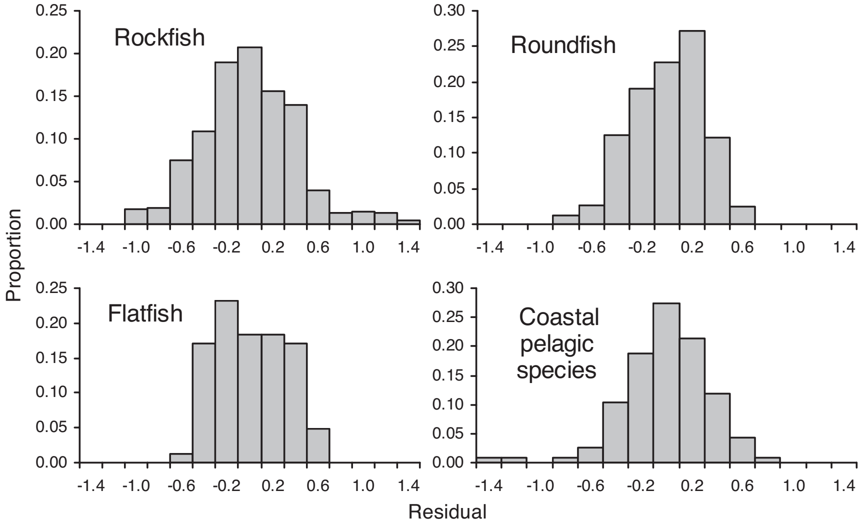

Figure 5

Composite distributions of log-deviations from the mean, pooled for four meta-analytic groupings (rockfish, roundfish, flatfish, and coastal pelagic species).

remains an active field of research. This fact is reflected by the range and changes over time in the ways that risk and uncertainty have been represented in fisheries assessments (see reviews by Francis and Shotton, 1997, and Patterson et al., 2001).

Our characterization of uncertainty does not include variability attributable to sources other than terminal-year biomass, which would tend to lead to negatively biased estimates of scientific uncertainty. Procedures for incorporating forecast uncertainty have been developed (Shertzer et al., 2008) and could be blended with our approach. Likewise, there is fertile ground to be explored with respect to uncertainty in F_{MSY}. Dorn (2002), for example, has developed a Bayesian prior for rockfish spawner-recruit steepness (h) that expresses uncertainty in estimates of stock productivity (see also Brooks et al., 2010). Because steepness maps almost directly onto F_{MSY} over a diverse range of groundfish life history patterns (Punt et al., 2008), a distribution of fishing mortality rates could be developed by mathematical composition of these functions, conditioned on the form of the stock-recruitment relationship. We assumed that estimates of F_{MSY} have negligible error and, as a consequence, uncertainty in OFL arises only from uncertainty in biomass estimates—an obvious simplification.

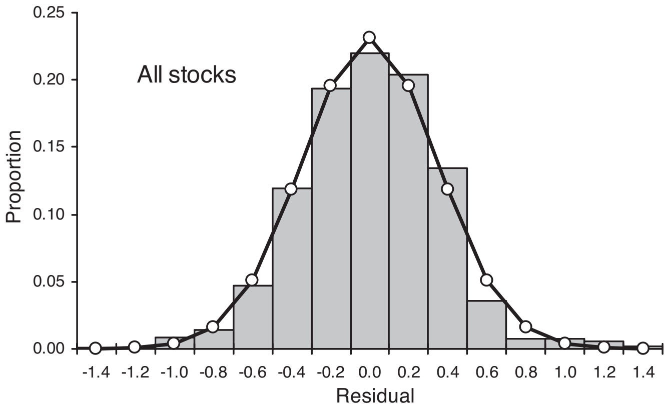

Figure 6

Aggregate distribution of log-deviations pooled over all 17 stocks with the fit of a normal distribution shown as the line with symbols ($\sigma = 0.36$).

Although we elected to characterize uncertainty by analyzing variability in biomass estimates from historical stock assessments, an alternative approach might be to use decision table results, which are a required element in groundfish stock assessments. Specifically, the PFMC terms of reference for groundfish assessments[2] require the development of a decision table for use in characterizing uncertainity in stock assessments. The guidance states the following:

> Once a base model has been bracketed on either side by alternative model scenarios, which capture the overall degree of uncertainty within the assessment, a 2-way decision table analysis (states-of-nature versus management action) is the preferred way to present the repercussions of uncertainty to management. An attempt should be made to develop alternative model scenarios such that the base model is considered twice as likely as the alternative models, i.e., the ratio of probabilities should be 25:50:25 for the low stock size alternative, the base model, and the high stock size alternative.

It is therefore possible, in theory, to express uncertainty regarding biomass in a quantitative manner by appropriately weighting different states of nature presented in groundfish decision tables, which are derived through the collective expert opinion of the analytical team, the review panel, and the Scientific and Statistical Committee. A preliminary analysis of this approach has been completed, although a comprehensive analysis was not possible because of incomplete data in stock assessment documents. In particular, statistical weights for all three states of nature that are defined in the decision analysis (low, base, and high) have not always been explicitly expressed. When a characterization of the relative probabilities of the various states of nature under consideration is lacking, decision tables provide a type of risk assessment, but they are inadequate for risk management (*sensu* Francis and Shotton, 1997). Still, in three of the nine cases examined, variances from decision tables were greater than a CV=37%. We view these preliminary findings as promising and recommend that a thorough analysis of statistically weighted states of nature be considered as an alternative approach to characterization of scientific uncertainty in groundfish stock assessments.

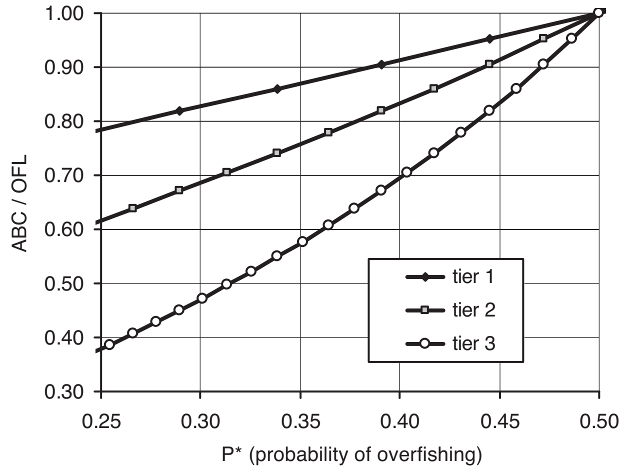

Figure 7

Relationship between the probability of overfishing (P^*) and an appropriate buffer between the allowable biological catch (ABC) and the overfishing level (OFL), based on varying amounts of uncertainty (σ=0.36, 0.72, and 1.44) assigned to different stock assessment tiers (1=data-rich, 2=data-moderate, and 3=data-poor), respectively.

Conclusions

Present and future management approaches for setting catch limits

This analysis was prepared in response to a pressing management need that arose from the requirements of the reauthorized MSA to implement ACLs by 2011. It is revealing to consider the ultimate impact of accounting for scientific uncertainty when setting catch limits at the PFMC. For all assessed groundfish species Table 4 provides the ABC and ACL as a percentage of the estimated F_{MSY} harvest level (OFL) as they were adopted by the PFMC in June 2010 (see footnote 5). Note that groundfish stocks are classified into three tiers based on the amount and quality of the information that is available for assessment modeling: tier-1 stocks are those for which there is data-rich information; tier-2 stocks are those for which there is data-moderate information; and tier-3 stocks are those for which there is data-poor information. Moreover, there was a consensus that scientific uncertainty cannot be lower for stocks that are more data limited. Hence, because σ=0.36 was derived from 81 tier-1 stock assessments, the Scientific and Statistical Committee recommended, and the PFMC elected to adopt, proxy estimates of uncertainty equal to double (0.72) and quadruple (1.44) σ for tier-2 and tier-3 stocks, respectively. This framework then provided a basis for separate ABC control rules for each tier. The PFMC then elected to adopt a P^*=0.45 for all tier-1 stocks and, with certain exceptions, P^*=0.40 for tier-2 and tier-3 stocks. Hence the scientific uncertainty buffers for the Council's data-rich stocks amounted to setting the ABC 4% below the OFL. Similarly, the adjustment for tier-2 stocks (σ=0.72 and P^*=0.40) was a 17% reduction (ABC=83% of the OFL) (Fig. 7). The differences between ACLs and ABCs shown in Table 4 reflect a variety of other factors, including 1) requirements for rebuilding overfished stocks; 2) harvest control rules for preventing stocks from becoming overfished; 3) socioeconomic considerations; 4) bycatch concerns for depleted species; 5) ecological considerations, and other factors.

Parsing scientific uncertainty in estimates of OFL from these other considerations was a particular challenge for the PFMC. Before the implementation of the new harvest specification framework recommended in the revised National Standard Guidelines, which were compelled by the reauthorized MSA, scientific and management uncertainties were considered jointly in setting optimum yields below the MSY harvest level. We have shown that quantifying scientific uncertainty in estimating exploitable biomass across multiple assessments through meta-analysis is a reasonable first approximation for explicitly accounting for uncertainty in preventing overfishing. Although all sources of error may not be have been considered with this approach, it was a helpful first step in the PFMC process. Importantly, with this approach the role of the Scientific and Statistical Committee in quantifying scientific uncertainty (by determining σ, a purely technical issue) and the role of the PFMC in deciding a preferred level of risk aversion to overfishing (by choosing P^*, which is a policy decision), are both duly respected. Coupling these two independent actions will help determine the ABC harvest level in a manner that is responsive to the mandates of the reauthorized MSA.

Table 4

Relative reductions from the overfishing limit (OFL) due to accounting for scientific and management uncertainty in setting 2011 groundfish allowable biological catches (ABCs) and annual catch limits (ACLs) at the Pacific Fishery Management Council (stocks in bold are overfished and their ACLs are based on rebuilding analyses). Tier 1=data rich; tier 2=data moderate; tier 3=data poor.

Stock/Complex	Tier	ABC ÷ OFL	ACL ÷ OFL
Bocaccio (*Sebastes paucispinis*)	1	96%	36%
Canary rockfish (*Sebastes pinniger*)	1	96%	17%
Cowcod (*Sebastes levis*)	2/3	76%	30%
Darkblotched rockfish (*Sebastes crameri*)	1	96%	59%
Pacific ocean perch (*Sebastes alutus*)	1	96%	18%
Widow rockfish (*Sebastes entomelas*)	1	96%	12%
Yelloweye rockfish (*Sebastes ruberrimus*)	1	96%	42%
Petrale sole (*Eopsetta jordani*)	1	96%	96%
Lingcod (OR & WA) (*Ophiodon elongatus*)	1	96%	96%
Lingcod (CA)	2	83%	83%
Pacific cod (*Gadus macrocephalus*)	3	69%	50%
Sablefish (*Anoplopoma fimbria*)	1	96%	77%
Shortbelly rockfish (*Sebastes jordani*)	2	83%	1%
Chilipepper (S 40°10′) (*Sebastes goodei*)	1	96%	96%
Splitnose rockfish (S 40°10′) (*Sebastes diploproa*)	1	96%	96%
Yellowtail rockfish (N 40°10′) (*Sebastes flavidus*)	1	96%	96%
Shortspine thornyhead (*Sebastolobus alascanus*)	1	96%	83%
Longspine thornyhead (*Sebastolobus altivelis*)	2	83%	70%
Black rockfish (WA) (*Sebastes melanops*)	1	96%	96%
Black rockfish (OR-CA)	1	96%	82%
California scorpionfish (*Scorpaena guttata*)	1	96%	96%
Cabezon (CA) (*Scorpaenichthys marmoratus*)	1	96%	96%
Cabezon (OR)	1	96%	96%
Dover sole (*Microstomus pacificus*)	1	96%	56%
English sole (*Parophrys vetulus*)	1	96%	96%
Arrowtooth flounder (*Atheresthes stomias*)	2	83%	83%
Starry flounder (*Platichthys stellatus*)	2	83%	75%
Longnose skate (*Raja rhina*)	1	96%	43%
Minor Nearshore rockfish North (species complex)	3	85%	85%
Minor Shelf rockfish North (species complex)	3	88%	44%
Minor Slope rockfish North (species complex)	3	91%	79%
Minor Nearshore rockfish South (species complex)	3	87%	87%
Minor Shelf rockfish South (species complex)	3	84%	32%
Minor Slope rockfish South (species complex)	3	92%	69%
Other Flatfish (species complex)	3	69%	48%
Other Fish (various) (species complex)	3	69%	50%

Acknowledgments

This study was generated, reviewed, and endorsed by members of the Pacific Fishery Management Council's Scientific and Statistical Committee, whom we would like to thank and acknowledge for their critical suggestions. In particular, members of the groundfish and coastal pelagic species subcommittees were most engaged in developing the analytical approach, including T. Barnes, L. Botsford, M. Dorn, S. Heppell, T. Jagielo, T. Tsou, and V. Wespestad. In addition, a number of people mined the primary PFMC stock assessment literature and provided us with a stock-specific time series of population abundance for use in the meta-analysis. In particular, we would like to offer our thanks to J. Cope (cabezon), P. Crone (Pacific mackerel), J. Field (bocaccio), M. Haltuch (petrale sole), X. He (widow rockfish), K. Hill (Pacific sardine), I. Stewart (canary rockfish, Pacific whiting, and yelloweye rockfish), and J. Wallace (yellowtail rockfish). Lastly, we appreciate the constructive reviews of this work that were provided by J. Cope, E.J. Dick, and A. MacCall. We also acknowledge the helpful comments of two anonymous reviewers and E. Williams who evaluated the manuscript for the journal.

Literature cited

Brandon, J., and P. R. Wade.
2006. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using Bayesian model averaging. J. Cet. Res. Manag. 8:225–239.

Brodziak, J., and K. Piner.
2010. Model averaging and probable status of North Pacific striped marlin, *Tetrapturus audax*. Can. J. Fish. Aquat. Sci. 67:793–805.

Brooks, E. N., J. E. Powers, and E. Cortés.
2010. Analytical reference points for age-structured models: application to data-poor fisheries. ICES J. Mar. Sci. 67:165–175.

Caddy, J. F., and R. McGarvey.
1996. Targets or limits for management of fisheries? N. Am. J. Fish. Manag. 16:479–487.

Cochran, W. G.
1977. Sampling techniques, 3rd ed., 429 p. John Wiley & Sons, New York.

Dorn, M. W.
2002. Advice on west coast rockfish harvest rates from Bayesian meta-analysis of stock-recruit relationships. N. Am. J. Fish. Manag. 22:280–300.

FAO (United Nations Food and Agriculture Organization).
1996. Precautionary approach to capture fisheries and species introductions. Technical guidelines for responsible fisheries, no. 2, 54 p. FAO, Rome.

Federal Register.
2009. Magnuson-Stevens Act Provisions; Annual Catch Limits; National Standard Guidelines; final rule. vol. 74, no. 11, January 16, p. 3178–3213. GPO, Washington, D.C.

Francis, R. I. C. C., and R. Shotton.
1997. "Risk" in fisheries management: a review. Can. J. Fish. Aquat. Sci. 54:1699–1715.

Hilborn, R., and C. J. Walters.
1992. Quantitative fisheries stock assessment—choice, dynamics, and uncertainty, 570 p. Kluwer Academic Publs., Boston.

IWC (International Whaling Commission).
1999. The report of the Scientific Committee, annex N. The revised management procedure (RMP) for baleen whales. J. Cet. Res. Manag. 1 (suppl.):251–258.

Kolody, D., T. Polacheck, M. Basson, and C. Davies.
2008. Salvaged pearls: lessons learned from a floundering attempt to develop a management procedure for southern bluefin tuna. Fish. Res. 94:339–350.

Johnson, N. L., and S. Kotz.
1970. Continuous univariate distributions, part 1, 300 p. Distributions in statistics. John Wiley & Sons, New York.

Methot, R. D.
2000. Technical description of the stock synthesis assessment program. NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.

Patterson, K., J. Cook, C. Darby, S. Gavaris, L. Kell, P. Lewy, B. Mesnil, A. Punt, V. Restrepo, D. W. Skagen, and G. Stefansson.
2001. Estimating uncertainty in fish stock assessment and forecasting. Fish Fish. 2:125–157.

Pawitan, Y.
2001. In all likelihood—statistical modelling and inference using likelihood, 528 p. Oxford Univ. Press, Oxford.

PFMC (Pacific Fishery Management Council).
2010. FMP amendment addresses National Standard 1—annual catch limits, accountability measures. Pacific Council News 34(1):5. [Available at: http://www.pcouncil.org/wp-content/uploads/Spring_2010_Newsletter.pdf.]

Poole, D., G. H. Givens, and A. E. Raftery.
1997. A proposed stock assessment method and its application to bowhead whales, *Balaena mysticetus*. Fish. Bull. 97:144–152.

Prager, M. H., C. E. Porch, K. W. Shertzer, and J. F. Caddy.
2003. Targets and limits for management of fisheries: a simple probability-based approach. N. Am. J. Fish. Manag. 23:349–361.

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

Punt, A. E., and. G. Donovan.
2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES J. Mar. Sci. 64:603–612.

Punt, A. E., M. W. Dorn, and M. A. Haltuch.
2008. Evaluation of threshold management strategies for groundfish off the U.S. west coast. Fish. Res. 94:251–266.

Punt, A. E., and R. Hilborn.
1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. 7:35–63.

Quinn, T. J., II, and R. B. Deriso.
1999. Quantitative fish dynamics, 524 p. Oxford Univ. Press, New York.

Ralston, S.
2002. West coast groundfish harvest policy. N. Am. J. Fish. Manag. 22:249–250.

Restrepo, V. R., J. M. Hoenig, J. E. Powers, J. W. Baird, and S. C. Turner.
1992. A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. Fish. Bull. 90:736–748.

Shertzer, K. W., M. H. Prager, and E. H. Williams.
2008. A probability-based approach to setting annual catch levels. Fish. Bull. 106:225–232.

Stewart, I. J., and O. S. Hamel.
2010. Stock assessment of Pacific hake, *Merluccius productus*, (a.k.a. whiting) in U. S. and Canadian waters in 2010, 290 p. [Available from Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR, 97220-1384.]

Wade. P. R.
1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. Mar. Mamm. Sci. 14:1–37.