An Approach to Quantifying Scientific

Uncertainty in West Coast Stock Assessments

Groundfish & CPS Subcommittees
Scientific and Statistical Committee
Pacific Fishery Management Council
7700 NE Ambassador Place, Suite 101
Portland, Oregon 97220-1384

March 1, 2010
Summary

Quantifying scientific uncertainty in estimating an appropriate catch level for a fish stock is challenging. Multiple sources of error can easily be identified, including measurement error that is conditioned on the adopted model, model specification error, forecast error, and uncertainty about overall stock productivity. In addition, there are without doubt other unknown factors that will negatively influence the precision of scientific advice on catch levels. Notwithstanding these difficulties, the Magnuson-Stevens Act (MSA) as amended in 2007 requires the Scientific and Statistical Committees (SSCs) of the Regional Fishery Management Councils to provide an acceptable biological catch (ABC) recommendation. According to the revised National Standard 1 Guidelines, ABCs must account for scientific uncertainty in estimates of overfishing limits (OFLs). Moreover, the ABC is the catch level that annual catch limits (ACLs) may not exceed.

While many sources of uncertainty exist, the focus here is on quantification of statistical measurement error and model specification error, particularly the latter. While not all inclusive, the study of these two factors is feasible using the information that is currently available. They are also likely to include the dominant sources of scientific uncertainty in the development of scientific advice with respect to catch levels at the Pacific Fishery Management Council for groundfish and coastal pelagic species.

Although full Bayesian integration through MCMC calculations is a preferred method of estimating measurement error “within” a stock assessment, an insufficient number of studies have successfully achieved that type of analysis. Consequently, we report the first order approximate estimates of the standard error on terminal biomass from stock assessments that are calculated by inverting 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 retrospective variation among multiple assessments of the same stock.

Results show that for 17 groundfish and coastal pelagic species the mean of the coefficient of variation on terminal biomass is 18%. This represents the average amount of statistical measurement error within assessments conducted for the PFMC. In contrast, the average coefficient of variation ascribable to model specification error (i.e., among assessment variation) is 37%, which is the greater of the two sources of uncertainty. Given these results, if only this source of variation is considered, and the probability of overfishing is fixed at 0.40, an appropriate buffer on the overfishing catch level is to reduce the harvest by ~9%.
Introduction

The Pacific Fishery Management Council currently manages a wide variety of west coast fish stocks under four different Fishery Management Plans (FMPs), including groundfish, coastal pelagic species (CPS), salmon, and highly migratory species (HMS). In the case of groundfish, the PFMC adopts optimum yields (OYs) for the fishery on a biennial basis following application of a harvest control rule to the results of stock assessments. Functionally, this procedure involves four separate calculations: (1) estimation of exploitable biomass for the current year, (2) projecting the population forward for several years into the near future, (3) applying a harvest rate to the projected population that would be expected to produce Maximum Sustainable Yield (MSY) in the long term, and (4) adjusting the projected catch downwards to account for a variety of factors of particular concern to management if the stock is assessed to be below a specified target level. Application of the MSY harvest rate ($F_{MSY}$ or its proxy) to the projected stock biomass results in an estimated Allowable Biological Catch (ABC)\footnote{The symbol ABC will be used for two quantities in this document “Allowable Biological Catch” and “Acceptable Biological Catch”}, which has been considered an upper bound on annual catches, i.e., catches in excess of the ABC represent “overfishing.” Adjustment of the ABC catch downwards to account for the concerns of management then results in an OY. An example of such an adjustment is the 40:10 groundfish harvest control rule that reduces OY relative to the ABC once the biomass of a stock drops below 40% of its unfished level. Hence, under the Council’s traditional approach to setting groundfish catch levels, the ABC is the absolute upper limit on annual catch, whereas the OY incurs some reduction in catch to account for a variety of conservation concerns to management. A comparable procedure is in place for CPS, except that the OY is termed a harvest guideline (HG).

The Magnuson-Stevens Fishery Conservation and Management Act (MSA) as amended in 2007 requires the establishment of Annual Catch Limits (ACLs)\footnote{MSA § 303(a)(15): Fishery management plans shall “establish a mechanism for specifying annual catch limits in the plan (including a multiyear plan), implementing regulations, or annual specifications, at a level such that overfishing does not occur in the fishery, including measures to ensure accountability.”} to prevent overfishing and measures to ensure accountability. An annual catch limit represents a numerically specified upper limit on total fishing mortality that should not be exceeded, and defines the level of annual catch that serves as the basis for invoking accountability measures\footnote{See NS1 Guidelines § 600.310(f)}. In addition, the MSA stipulates that the Scientific and Statistical Committees (SSCs) of each of the eight regional Fishery Management Councils are now required to recommend an acceptable biological catch (ABC) to their respective Councils, and which the NS1 Guidelines further explain is a reference point that accounts for scientific uncertainty in the estimate of the overfishing limit. This new requirement effectively adds a new step in setting catch levels. In particular, the application of $F_{MSY}$ (or its proxy) to projected biomass values from a stock assessment now results in an Overfishing Limit (OFL), which is functionally identical to the old definition of the ABC in the Groundfish and CPS fishery management plans. As before, annual catches in excess of the OFL constitute overfishing. However, the NS1 Guidelines now define...
ABC as an annual catch amount that is reduced from the OFL in order to account for scientific uncertainty in the development of management advice by SSCs to their Councils. The expectation under the Guidelines is that scientific advice that is relatively uncertain will result in ABCs that are relatively lower, all other things being equal, i.e., a precautionary reduction in catch will occur due purely to scientific uncertainty. The MSA also states that the ACL may not exceed the SSC’s fishing level recommendation, which the NS1 Guidelines equate to the newly required ABC. Moreover, if management is unable to insure that annual catches do not exceed the ACL, possibly due to inadequate in-season monitoring of catches, the NS1 Guidelines recommend the development of an Annual Catch Target (ACT) that is specified at a level sufficiently below the ACL to insure that the ACL is not exceeded more than once in four years due to management uncertainty. The relationships of these various terms are depicted graphically in Figure 1 below.

Figure 1. Relationships among the various definitions of terms under the MSA and NS1 Guidelines, and the Council’s old system.

Given the new requirement that each SSC is now responsible for characterizing scientific uncertainty in a manner that allows their Council to establish a risk policy that provides a precautionary “buffer” between the OFL and the ABC, this document summarizes the Pacific Fishery Management Council SSC’s preliminary approach to addressing this problem for groundfish and CPS stocks.

Sources of Uncertainty

As described previously, estimation of the OFL (formerly ABC) involves three basic steps: (1) estimation of current exploitable biomass ($B_t$), (2) projecting the current exploitable biomass into the future for several years ($B_{t+1}$, $B_{t+2}$, etc.), and (3) applying an estimate of $F_{MSY}$ to predictions of future biomass. While there are clear uncertainties associated with each step, the PFMC SSC elected to focus its attention first and foremost

---

4 MSA§ 302(h)(6)
on variation in the estimation of current biomass in the terminal year of groundfish and CPS stock assessments. Our reason for doing so is aptly illustrated in Figure 2, which shows the results of 15 different Pacific whiting stock assessments that have been conducted for the PFMC over the last 18 years. It is instructive to consider this species because it is one of the most data-rich stocks managed by the Council, it is of great economic importance, and it has been assessed on an annual basis for many years. However, in spite of considerable resources having been devoted to evaluating the status of this stock, from an assessment retrospective perspective, estimates of biomass have been highly variable. Note, for example, that estimated spawning biomass in 1985 has ranged from $1.2-5.9 \times 10^6$ mt; approximately a 5-fold range in abundance.

![Figure 2](image)

**Figure 2.** Results of 15 separate stock assessments of Pacific whiting conducted for the PFMC since 1991. The legend refers to the year the stock assessment was conducted.

There are many reasons for these variations in stock size estimation, including differences in: (1) the modeling software that was used, (2) the composition of the analytical team doing the assessment, (3) the composition of the review panel, (4) changes in the availability of data, (5) altered priors for the parameters, and (6) changes in overall model structure. Importantly, these issues contribute to variation in all groundfish and CPS stock assessments, which collectively demonstrate considerable “among” assessment variance. Hence, it is currently the view of the SSC that quantifying and accounting for this source of uncertainty is the first and most important to consider when establishing a buffer between the OFL and the ABC. However, as this process develops into the next biennial management cycle the SSC intends to consider other types of errors, including forecast uncertainty (Shertzer et al. 2008) and uncertainty in the optimal harvest rate (e.g., Dorn 2002; Punt et al. 2008). Hence, quantification of variation as revealed in this exercise should be considered a lower bound on total uncertainty at this time. However, even if forecast and harvest rate uncertainty were incorporated explicitly in this analysis, numerous other unaccounted for factors exist that may never be fully evaluated, including for example the effects of climate and/or ecosystem interactions on the estimation of an OFL.
For our analysis we initially consider two types of uncertainty in biomass estimation. The first we term “within” assessment variability and is represented by the coefficient of variation (CV) for the terminal year biomass taken from the most recent stock assessment that has been conducted, whether it was a full or update assessment. In a very limited number of studies (e.g., Pacific Ocean perch) full Bayesian integration of uncertainty via Monte Carlo Markov Chain (MCMC) analysis has been achieved. Such cases are unusual, however. Consequently, we report the asymptotic standard error estimate on terminal biomass developed by inverting of the model’s Hessian matrix as a first order approximation of variation. This error estimate can be considered 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 statistic as the denominator.

However, as previously noted, “among” assessment variation is attributable to a wide variety of factors, many of which represent a significant form of model or structural uncertainty. Assertion of asymptotic or dome-shaped selectivity patterns is one example, as is incorporation of age-dependent natural mortality. Such structural issues will often change from one assessment to the next. Likewise, biologically important fixed parameters often change from one assessment to the next (e.g., natural mortality or spawner-recruit productivity) and whole new data time series can be incorporated into the assessment model (e.g., the NWFSC combined trawl survey). Beyond such changes in model specification, among assessment variation includes other sources of variability due to, for example, differences in the reviewers who evaluated and approved an assessment.

To quantify among assessment variability we assembled time series of biomass from historical assessments of a stock. We excluded update assessments unless they were the most recent assessment conducted (in which case we excluded the original full assessment upon which the update was based) because of strong constraints imposed by the TOR for groundfish stock assessments on how much update assessments could change from the last full assessment. In situations where a change in biomass metric across assessments occurred (e.g., mid-year biomass in one assessment and beginning year biomass in another) we used ratio estimation (Cochran 1977) over a common time frame to standardize to a common metric across all assessments that were conducted on a stock. Lastly, we limited the number of data points under consideration to no more than the last 20 years from each assessment in order to focus our attention on variation associated with the estimation of terminal year biomass. We also trimmed the time series to include only the most recent 20, 15, 10, and 5 years to evaluate the stability of the estimates 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 fundamentally different approaches to calculating variation around a point of central tendency, that is the population mean: (1) consider all biomass estimates in a year as equally plausible representations of reality, (2) consider the mean of biomass estimates in a year as the best estimate of central tendency, and (3) consider the most recent stock assessment as the best estimate of the population mean.
In the first approach, biomass variation between two stock assessments was quantified by forming all possible ratios of estimated abundances in common years. Specifically, if there existed an estimate of biomass \( B \) in common year \( t \) from assessments \( i \) and \( j \), we calculated: 
\[
R_{ij,t} = \frac{B_{i,t}}{B_{j,t}},
\]

i.e., the proportional deviation of assessment \( i \) using assessment \( j \) as a standard. Based on a symmetry argument we also calculated \( R_{ji,t} \) and all the ratios were log\(_e\)-transformed. Note that because \( \ln(R_{ij,t}) = -\ln(R_{ji,t}) \) the distributions were perfectly symmetrical. For each stock under consideration the standard deviation (\( \sigma^* \)) of the data was calculated. This statistic is positively biased, however, because it is based on the ratio of two random variables \( (B_{i,t} \text{ and } B_{j,t}) \). The appropriate bias correction term (\( \sqrt{2} \)) can be derived\(^5\) and applied so that the corrected estimate of variation is:

\[
\sigma = \frac{\sigma^*}{\sqrt{2}}
\]

Thus, in method one, we used the bias-adjusted estimate of the standard deviation of the \( \ln(R_{ij,t}) \) as a quantitative measure of among assessment variation.

In method two, the data were log-transformed and the mean and standard deviation calculated (\( \sigma \)). Variation in this approach was measured as squared deviations from the annual mean in log-space. Specifically, we calculate the mean log-biomass in year \( t \) as:

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

where, as before, \( B_{i,t} \) is the estimated biomass from the \( i^{th} \) assessment in year \( t \) and \( n_t \) is the number of available assessments in year \( t \) (\( n_t \geq 2 \)). The standard deviation (\( \sigma \)) is then calculated as:

\[
\sigma = \sqrt{\frac{\sum_i \sum_j (\ln[B_{i,t}] - \ln[B_t])^2}{\sum i n_t - 1}}
\]

Method three is exactly the same as method two except that the mean (\( \ln[B_t] \)) is replaced with the biomass estimate from the most recent stock assessment, and the most recent year is excluded from the summations and the calculation of the \( n_t \). This approach assumes that the most current information represents the best estimate of the population mean.

It is understood that these statistics are not valid measures of “among” assessment variance, at least in the traditional ANOVA context, because one would expect some variation in biomass estimates due to incorporation of new data, even if the structural characteristics of the model remained completely unchanged. In this regard the measured variance might be thought of as “total” variance. However, the data used from one assessment to the next are not independent of one another. The same age compositions and trend indices, for example, are used over and over again. Hence, to estimate the degree to which the “total” variance statistic was contaminated by serially correlated “within” variance, retrospective analyses of the most recent stock assessments were conducted and the variances calculated as described above. This effectively removed any model specification differences in the calculation of $\sigma$, which we term $\sigma_{\text{retro}}$. True “among” model variance could then be calculated as $\sigma^2_{\text{among}} = \sigma^2_{\text{total}} - \sigma^2_{\text{retro}}$. Given a valid estimate of $\sigma^2_{\text{among}}$, total variance could be calculated by summation with $\sigma^2_{\text{within}}$ based on Hessian approximations.

In order to combine “within” and “among” sources of variation we note that for lognormally distributed random variables, the CV on the arithmetic scale is equal to:

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

where $\sigma^2$ is the variance on the logarithmic scale (Johnson and Kotz 1970). When necessary we used this relationship to convert a within assessment CV to a variance term on the logarithmic scale, added the square of the among assessment log-scale standard deviation, and back-transformed the total variance to a coefficient of variation on the arithmetic scale.

**Methodological Comparisons**

When the data were restricted to include only the most recent 20 years (1990-2009), methods one, two, and three yielded mean estimates of $\sigma$ equal to 0.382, 0.335, and 0.307, respectively. The SSC’s groundfish and CPS subcommittees elected to express uncertainty using method two (squared deviations from the mean in log-space) upon consideration of these results and discussion of the relative merits of the different approaches. Similarly, a sensitivity of the results to the choice of the number of years included in the calculation, i.e., the most recent 5, 10, 15, or 20 years, showed the estimates were robust to this choice, except that $\sigma$ was not estimable for some species as the time series of data was truncated to $\leq 10$ years. The subcommittees recommended a standard window of time extending from 1990-2009 as the basis for quantifying variation among stock assessments based on that finding.

The evaluation of $\sigma_{\text{retro}}$, as a basis for estimating $\sigma_{\text{among}}$, showed that $\sigma_{\text{retro}}$ was sometimes greater than $\sigma_{\text{total}}$. Considering that a “within” model retrospective analysis will lead, on occasion, to deletion of an entire data component, this result is perhaps not surprising, particularly if structural changes to the model were made to partially offset the effect of a
new data source. Calculation of $\sigma_{\text{retro}}$ as an approach to estimating a corrected $\sigma_{\text{among}}$ assessments was abandoned because of this type of erratic behavior. Rather, it was concluded that whichever variance statistic was greater, i.e., $\sigma^2_{\text{total}}$ or $\sigma^2_{\text{within}}$ (based on the Hessian approximation), would be used as the basis for calculating the adjustment to the OFL to account for scientific uncertainty.

Following a similar line of inquiry, subcommittee members considered estimation of variation in biomass estimates from stock assessments using the information contained in the decision tables contained in stock assessments. The Terms of Reference for Groundfish Stock Assessments requires development of decision tables for use in characterizing stock assessment uncertainty. The guidance states:

“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 biomass uncertainty in a quantitative manner by appropriately weighting the results for different states of nature represented in the decision table. A preliminary analysis of this approach used alternative forecasts for the biomass in 2011 from decision tables to calculate log-scale $\sigma_{\text{decision}}$. $\sigma_{\text{decision}}$ would be used to calculate the ABC buffer if it were larger than both $\sigma_{\text{total}}$ and or $\sigma_{\text{within}}$.

The table below shows estimates of log-normal $\sigma_{\text{decision}}$ calculated by assuming: (1) the base model is the true mean and (2) weights of 25%, 50% and 25% on either the Low, Base and High alternatives (approach A), or the Low and Base alternatives, and a high alternative constructed to be equidistant in log space from the Base alternative as the Low alternative is (this is equivalent to giving the Low and Base alternative each a weight of 0.5) (approach B). The latter approach is preferred because it represents uncertainty around the base model in the lower direction, which is the direction of uncertainty we are concerned about when defining buffers to avoid exceeding OFLs.

Table 1: Uncertainty estimation based on decision table log-scale standard deviations.

<table>
<thead>
<tr>
<th>Species</th>
<th>Abundance Measure</th>
<th>Approach A (Low, Base, High)</th>
<th>Preferred Approach B (Low, Base)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canary rockfish</td>
<td>Sp. biomass</td>
<td>0.546</td>
<td>0.678</td>
</tr>
<tr>
<td>Darkblotted rockfish</td>
<td>Sp. biomass</td>
<td>0.671</td>
<td>0.792</td>
</tr>
<tr>
<td>Lingcod N. of 42º</td>
<td>Sp. output</td>
<td>0.052</td>
<td>0.062</td>
</tr>
<tr>
<td>Lingcod S. of 42º</td>
<td>Sp. biomass</td>
<td>0.442</td>
<td>0.551</td>
</tr>
<tr>
<td>Pacific ocean perch</td>
<td>Sp. biomass</td>
<td>0.129</td>
<td>0.125</td>
</tr>
<tr>
<td>Petrale sole</td>
<td>Sp. biomass</td>
<td>0.301</td>
<td>0.351</td>
</tr>
<tr>
<td>Splitnose rockfish</td>
<td>Sp. output</td>
<td>0.082</td>
<td>0.115</td>
</tr>
<tr>
<td>Widow rockfish</td>
<td>Sp. output</td>
<td>0.234</td>
<td>0.290</td>
</tr>
<tr>
<td>Yelloweye rockfish</td>
<td>Sp. output</td>
<td>0.371</td>
<td>0.332</td>
</tr>
</tbody>
</table>
Stock-Specific Accounts

In the accounts that follow information for 15 groundfish and 2 CPS stocks is summarized. Specifically, we include the following well-studied, relatively data-rich species: bocaccio, canary rockfish, chilipepper, darkblotched rockfish, Pacific Ocean perch (POP), shortspine thornyhead, widow rockfish, yelloweye rockfish, yellowtail rockfish, cabezon, lingcod, Pacific whiting, sablefish, Dover sole, petrale sole, Pacific mackerel, and Pacific sardine. All have been assessed using some version of the Stock Synthesis modeling program in a fully dynamic context, except for POP which uses a stand-alone forward-projection age-structured model programmed in ADMB.

The summary for each stock includes a brief description of the species, references to what assessments were included in the analysis, and whether any ratio estimation was required to standardize biomass metrics. In addition, graphs are provided depicting: (1) time series of population biomass trajectories from 1970-2009, summarized from previously completed full stock assessments [Figure 3a, b] and (2) frequency distributions of log-deviations from the mean [Figure 4a, b]. The latter form the basis of stock-specific estimates of $\sigma_{total}$.

**Bocaccio (Sebastes paucispinis):**

Bocaccio is an overfished rockfish that is currently under rebuilding (Figs. 3a & 4a). It is principally distributed in the State of California. We identified five stock assessments that could be incorporated into the meta-analysis (Ralston et al. 1996; MacCall et al. 1999; MacCall 2002; MacCall 2003; Field et al. In press). While earlier assessments of this species have been conducted, they did not identify a base model and instead presented a range of alternatives predicated on a predefined array of possibilities. Results from Field et al. (In press) were presented as mid-year total biomass, whereas the four earlier studies referenced biomass at the beginning of the year. However, results from MacCall (2003) included time series in both biomass metrics and ratio estimation over the period 1951-2002 from that assessment was used to standardize to biomass at the beginning of the year ($\Sigma$ begin-year biomass / $\Sigma$ mid-year biomass = 1.044). For bocaccio the standard deviation ($\sigma$) calculated using method two (squared deviations from the mean in log-space) is 0.367 (n = 61).

**Canary rockfish (Sebastes pinniger):**

Canary rockfish is also an overfished stock under a highly restrictive rebuilding plan (Figs. 3a & 4a). It is distributed along the entire U.S. west coast and is largely responsible for the implementation of the Rockfish Conservation Area (RCA) spatial trawl area closures. For the meta-analysis we report the results of eight stock assessments (Sampson and Stewart 1994; Sampson 1996; Williams et al. 1999; Crone et al. 1999; Methot and Piner 2002; Methot and Stewart 2005; Stewart 2008; Stewart 2009). All report their results in terms of spawning biomass [mt]. However, we made the following adjustments to the abundance time series from these assessments: (1) we averaged the “base-1” and “base-2” models from Sampson and Stewart (1994), Sampson
Figure 3a. Time series of biomass estimates for selected groundfish and CPS stocks that have been assessed for the PFMC. The bold line in each graph represents the most recent analysis, whereas the lighter gray lines are time series of abundance from older stock assessments. Uncertainty is measured based on the variability of estimates within years. Units of biomass are provided in the individual species-specific accounts.
Figure 3b. Time series of biomass estimates for selected groundfish and CPS stocks that have been assessed for the PFMC. The bold line in each graph represents the most recent analysis, whereas the lighter gray lines are time series of abundance from older stock assessments. Uncertainty is measured based on the variability of estimates within years. Units of biomass are provided in the individual species-specific accounts.
Figure 4a. Frequency distributions of log-scale biomass deviations from the mean for selected groundfish and CPS stocks that have been assessed for the PFMC.
Figure 4b. Frequency distributions of log-scale biomass deviations from the mean for selected groundfish and CPS stocks that have been assessed for the PFMC.
(1996) and Crone et al. (1999), (2) we added the southern results of Williams et al. (1999) to the northern results of Crone et al. (1999) to obtain a coastwide estimate, (3) we used a ratio estimate developed from the period 1967-93 based on the coastwide biomass from the combined 1999 assessments relative to the north (×1.16) to expand the northern results of Sampson and Stewart (1994) and Sampson (1996) to coastwide values, and (4) we averaged the “diff” and “no-diff” models from Methot and Stewart (2005). Following these adjustments we calculate that for canary rockfish $\sigma = 0.375$ based on 85 deviations.

**Chilipepper (Sebastes goodei)**

Only two stock assessments of chilipepper were included into this study (Ralston et al. 1998; Field 2008). This species is currently underutilized because landings have been constrained by restrictions that have been placed on the bocaccio fishery, a species with which it regularly co-occurs. Chilipepper is predominately found only in California. No adjustments to the abundance time series were required because estimates of total age 1+ biomass were available from both assessments (Figs. 3b & 4b). We calculate that $\sigma = 0.354$ based on the variation of 22 deviations from the mean.

**Darkblotched rockfish (Sebastes crameri)**

This species is primarily distributed off the State of Oregon and is one of several overfished rockfish stocks that are currently under rebuilding (Figs. 3a & 4a). A review of past assessments indicates that full stock assessments were conducted in 2003 and 2005 and an update assessment was completed in 2009 (Rogers 2003; Rogers 2005; Wallace and Hamel In press). All three assessment report time series of total age-1+ biomass and summarize the stock over the same geographical area. Consequently, no standardization of biomass metrics was required. Analysis of the 45 data points yields a standard deviation of 0.103.

**Pacific Ocean perch (Sebastes alutus)**

Like the preceding species, Pacific Ocean perch (a.k.a. POP) is a northerly distributed overfished rockfish stock (Figs. 3a & 4a). Large removals occurred due to distant water foreign fishing fleets in the 1960s and POP was one of the first stocks of conservation concern to the PFMC. Based on an examination of material in previous PFMC Stock Assessment and Fishery Evaluation (SAFE) documents it was determined that only the assessments conducted in 1992, 1998, and 2009 (an update) could be included in this analysis (Ianelli et al. 1992; Ianelli and Zimmerman 1998; Hamel 2009). All of these studies provided time series of stock size in terms of total biomass. A summarization of the data yielded 20 deviations and resulted in $\sigma = 0.352$.

**Shortspine thornyhead (Sebastolobus alascanus)**

Shortspine thornyhead is a member of the “DTS” complex (Dover sole, thornyhead, and sablefish) and is harvested primarily in the continental slope trawl fishery (Figs. 3a & 4a). Like rockfishes, it is a member of the scorpionfish family, although it has quite different
life history characteristics (e.g., oviparity). For this study three stock assessments were identified for detailed analysis (Ianelli et al. 1994; Piner and Methot 2001; Hamel 2005); no standardization of time series was required. Results show that the standard deviation of the 39 anomalies was 0.923. The estimate of among assessment variances was highest for this stock because of the markedly different biomass time series due to a change from an assumption of asymptotic to dome-shaped selectivity in the 2005 assessment. (Fig. 3a).

**Widow rockfish (Sebastes entomelas)**

This species is another overfished rockfish that is under rebuilding (Figs. 3a & 4a). Five assessments met the necessary criteria for inclusion in the meta-analysis: Ralston et al. (1997), Williams et al. (2000), He et al. (2003), He et al. (2006), and He et al. (In press). All studies reported total spawning output, although the data presented in Ralston et al. (1997) scaled differently relative to the other assessments. Hence, a ratio estimate was developed to convert spawning output from that study to be equivalent to the others. To accomplish the standardization the ratio of the sums of spawning output (SO) over the period 1970-97 was utilized, i.e., \( \frac{\sum \text{SO } 2000 \text{ model}}{\sum \text{SO } 1997 \text{ model}} = 0.083 \). Following standardization the 61 data points resulted in \( \sigma = 0.241 \).

**Yelloweye rockfish (Sebastes ruberrimus)**

This is yet another overfished rockfish stock that is found along the entire US west coast, typically in rockfish shelf habitats. Six assessments have been completed since 2001 (Figs. 3a & 4a), although two were updates and could not be utilized. The remaining four assessment were used here, i.e., Wallace 2001; Methot et al. 2002; Wallace et al. 2006; Stewart et al. In press. Results obtained from the most recent assessments (Stewart et al. In press) presented stock size in terms of larval production, whereas the three earlier assessments presented spawning biomass. The former statistic was therefore converted to units of spawning biomass using ratio estimation over the period 1924-2006. Specifically, results in Wallace et al. (2006) showed that \( \frac{\sum \text{spawning biomass}}{\sum \text{total biomass}} = 0.429 \), which when multiplied by the total biomass estimates provided in Stewart et al. (In press) to calculate estimates of spawning biomass. The standard deviation of the 58 values resulted in \( \sigma = 0.492 \) after standardization of the time series.

**Yellowtail rockfish (Sebastes flavidus)**

This more northerly species, like chilipepper, has been underutilized in recent years due to constraints placed on it by concerns over other overfished rockfish, including canary and widow rockfish (Figs. 3b & 4b). Six assessments were identified that could inform the estimation of scientific uncertainty in stock size estimation (Tagart 1991; Tagart 1993; Tagart and Wallace 1996; Tagart et al. 1997; Tagart et al. 2000; Wallace and Lai 2005). All stocks reported the abundance of yellowtail rockfish in terms of total age-4+ biomass, which was summed over three sub-regions. However, the first four assessments each presented two alternative models, which were averaged for this analysis. This produced 66 deviations, resulting in a standard deviation of 0.269.
Cabezon (*Scorpaenichthys marmoratus*)

Cabezon is a member of the sculpin family (Cottidae) that inhabits shallow, high relief reef systems in California and Oregon (Figs. 3b & 4b). Three stock assessments have been completed for this stock and were evaluated as part of the meta-analysis (Cope *et al.* 2004; Cope and Punt 2006; Cope and Key In press). The first assessment of cabezon did not report spawning output in the same units as the last two studies. A ratio estimate of the 1970-2003 summed spawning outputs from the 2006 and 2004 assessments was therefore used to standardize the 2004 data (2006 units = 1.06×10^{-3} · 2004 units). After standardizing the data, 46 anomalies were calculated, yielding $\sigma = 0.154$.

Lingcod (*Ophiodon elongatus*)

Lingcod is a large hexagrammid west coast species of considerable importance to both commercial and recreational fisheries. While once overfished, it recovered rapidly and is currently forms the basis for a productive fishery. Four stock assessments (Figs. 3a & 4a) were included in the meta-analysis: Jagielo *et al.* (2000), Jagielo *et al.* (2003), Jägielo and Wallace (2005), and Hamel *et al.* (In press). All assessments reported biomass time series in equivalent units and no standardization was required. Analysis of the 56 values led to a standard deviation of 0.263.

Pacific whiting (*Merluccius productus*)

Pacific whiting, also known as Pacific hake, has been assessed more times than any other groundfish stock. It is a gadoid species that undertakes annual migrations along the entire U.S. west coast to summer feeding grounds off Oregon, Washington, and British Columbia. Time series of spawning biomass from 15 different stock assessments (Figs. 3a & 4a) were summarized for the meta-analysis (Dorn and Methot 1991, 1992; Dorn *et al.* 1993; Dorn 1994, 1995, 1996; Dorn and Saunders 1997; Dorn *et al.* 1999; Helser *et al.* 2002, 2004, 2005, 2006; Helser and Martell 2007; Helser *et al.* 2008; Hamel and Stewart In press). The four assessments conducted from 2004-2007 each presented two separate models that differed due to assumptions about the acoustic survey $q$; these were averaged within each assessment to produce a single assessment-specific time series for this analysis. Analysis of the 151 log-differences yielded $\sigma = 0.286$.

Sablefish (*Anoplopoma fimbria*)

Sablefish is a very important commercial species that is harvested in fixed gear and trawl fisheries operating on the continental shelf and slope. It is found along the entire U.S. west coast. Seven stock assessments (Figs. 3b & 4b) were incorporated into the meta-analysis (Methot 1992; Methot *et al.* 1994; Crone *et al.* 1997; Methot *et al.* 1998; Schirripa and Methot 2001; Schirripa and Colbert 2005; Schirripa 2007). All analyses reported stock size in terms of spawning biomass. However, the 1997 and 1998 assessments presented two and three, respectively, different model scenarios that were blended (averaged) into a single representation for each assessment. From these data a
total of 82 deviations were calculated, which yielded an estimated standard deviation of 0.340.

Dover sole (*Microstomus pacificus*)

This flatfish species is a member of the continental slope DTS complex that is harvested by trawl fisheries along the whole west coast. Although the stock has been assessed for many years, only three assessments (Figs 3b & 4b) were utilized in the meta-analysis because of changing geographic stock definitions. For this study we summarized spawning biomass estimates from Brodziak *et al.* (1997), Sampson and Wood (2001), and Sampson (2005). Even then a ratio estimate (1967-96) was required to expand the 1997 assessment results (Monterey to US Vancouver INPFC areas) to a coastwide estimate (×1.42). Following standardization a total of 41 log-deviations were calculated, with $\sigma = 0.360$.

Petrale sole (*Eopsetta jordani*)

This is a high-value flatfish species that is taken in trawl fisheries along the entire west coast. It has been fished intensively for decades. We analyzed results from three petrale sole stock assessments (Figs. 3b & 4b), including Sampson and Lee (1999), Lai *et al.* (2005), and Haltuch and Hicks (In press). Results in all documents are presented as time series of spawning biomass and consequently no standardization to a common abundance metric was required. From the three reports 41 anomalies were calculated, resulting in a standard deviation of 0.227.

Pacific mackerel (*Scomber japonicus*)

Pacific mackerel is a CPS species that is fished primarily off the State of California and Mexico in “wetfish” purse seine fisheries. Two update stock assessments were excluded from the meta-analysis, but four other full assessments were included (Hill and Crone 2004, 2005; Dorval *et al.* 2007; Crone *et al.* 2009) (Figs. 3b & 4b). All assessments report population abundance in terms of spawning biomass [mt] and no ratio-based standardization was needed. From those four citations 66 deviations were calculated, resulting in $\sigma = 0.415$.

Pacific sardine (*Sardinops sagax*)

The last species considered in this analysis is Pacific sardine, which is a very important CPS species that is harvested from Mexico to Canada in purse seine fisheries. We considered three full sardine stock assessments (Figs 3b & 4b) in the analysis, including Conser *et al.* (2004), Hill *et al.* (2007), and Hill *et al.* (2009). All three assessment documents reported population abundance in terms of spawning biomass over a common geographical area and no standardization of metrics was required. A total of 51 log-anomalies were obtained, with a standard deviation of 0.206.
Synopsis: Seventeen species were considered in this analysis and individual stock-specific results for all are summarized in Table 2. Also included are the “within” assessment estimates of statistical uncertainty as measured by the asymptotic standard deviation derived by inverting the Hessian matrix. In order to directly compare the two measures of uncertainty, the log-scale variation among assessments was expressed as a CV on the arithmetic scale (Johnson and Kotz 1970). When plotted against one another (Fig. 5) it is evident that the total variation among stock assessments is typically greater than that within assessments. The most obvious exception to that generalization is Pacific sardine. In contrast, the total assessment CV for shortspine thornyhead (SST) is far in excess of that measured for any other stock. We note that there is not a significant correlation between the two measures of variation ($r = -0.36, P = 0.23$). Finally, the mean and standard deviation of the 17 total CVs is 0.359 and 0.231, whereas for within assessment CVs these statistics are 0.181 and 0.090, respectively.

![Figure 5. Relationship between the “total” coefficient of variation (CV), calculated from biomass variation over multiple full stock assessments, and an estimate of the CV based on statistical uncertainty measured “within” the most recent analysis. CVs are presented in arithmetic-scale.](image-url)
<table>
<thead>
<tr>
<th>Group</th>
<th>Common Name</th>
<th>Scientific Name</th>
<th>Number of Assessments</th>
<th>Total Variation</th>
<th>Within Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Squared Deviations</td>
<td>log-scale Standard Deviation</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Bocaccio</td>
<td><em>Sebastes paucispinis</em></td>
<td>5</td>
<td>61</td>
<td>0.367</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Canary rockfish</td>
<td><em>Sebastes pinniger</em></td>
<td>7</td>
<td>85</td>
<td>0.375</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Chilipepper</td>
<td><em>Sebastes goodei</em></td>
<td>2</td>
<td>22</td>
<td>0.354</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Darkblotched rockfish</td>
<td><em>Sebastes crameri</em></td>
<td>3</td>
<td>45</td>
<td>0.103</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Pacific Ocean Perch</td>
<td><em>Sebastes alutus</em></td>
<td>3</td>
<td>20</td>
<td>0.352</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Shortspine thornyhead</td>
<td><em>Sebastolobus alascanus</em></td>
<td>3</td>
<td>39</td>
<td>0.923</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Widow rockfish</td>
<td><em>Sebastes entomelas</em></td>
<td>5</td>
<td>61</td>
<td>0.241</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Yelloweye rockfish</td>
<td><em>Sebastes ruberrimus</em></td>
<td>4</td>
<td>58</td>
<td>0.492</td>
</tr>
<tr>
<td>Rockfish</td>
<td>Yellowtail rockfish</td>
<td><em>Sebastes flavidus</em></td>
<td>6</td>
<td>66</td>
<td>0.269</td>
</tr>
<tr>
<td>Roundfish</td>
<td>Cabezon</td>
<td><em>Scorpaenichthys marmoratus</em></td>
<td>3</td>
<td>46</td>
<td>0.154</td>
</tr>
<tr>
<td>Roundfish</td>
<td>Lingcod</td>
<td><em>Ophiodon elongatus</em></td>
<td>4</td>
<td>56</td>
<td>0.263</td>
</tr>
<tr>
<td>Roundfish</td>
<td>Pacific whiting</td>
<td><em>Merluccius productus</em></td>
<td>15</td>
<td>151</td>
<td>0.286</td>
</tr>
<tr>
<td>Roundfish</td>
<td>Sablefish</td>
<td><em>Anoplopoma fimbria</em></td>
<td>7</td>
<td>82</td>
<td>0.340</td>
</tr>
<tr>
<td>Flatfish</td>
<td>Dover sole</td>
<td><em>Microstomus pacificus</em></td>
<td>3</td>
<td>41</td>
<td>0.360</td>
</tr>
<tr>
<td>Flatfish</td>
<td>Petrale sole</td>
<td><em>Eopsetta jordani</em></td>
<td>3</td>
<td>41</td>
<td>0.227</td>
</tr>
<tr>
<td>CPS</td>
<td>Pacific sardine</td>
<td><em>Sardinops sagax</em></td>
<td>3</td>
<td>51</td>
<td>0.206</td>
</tr>
<tr>
<td>CPS</td>
<td>Pacific mackerel</td>
<td><em>Scomber japonicus</em></td>
<td>4</td>
<td>65</td>
<td>0.415</td>
</tr>
</tbody>
</table>
Pooled Results

The PFMC’s groundfish FMP includes approximately 80 species and ACLs will need to be established for all species that are “in the fishery”. Of the stocks listed in the FMP, only about 25-30% have been assessed using population dynamics models, e.g., Stock Synthesis (Methot 2000). Importantly, a number of species have only been assessed once, so that total assessment biomass variation cannot always be estimated using meta-analysis, even when an assessment has been conducted. Therefore, some merit in pooling results from the well-studied species described here in order to develop proxy relationships for all groundfish and CPS stocks, even those that have been assessed multiple times.

Three natural groupings exist for the groundfish species we have summarized, which are classified in the FMP as rockfish, roundfish, and flatfish. In Table 2 each of the 15 groundfish stocks we considered is assigned to one of these three species groupings. In a similar manner, the two CPS species were grouped together. We considered two approaches for pooling the results for the 17 stocks: (a) taking the square root of the average of the variances, and (b) pooling all of the residuals and finding the standard deviation of the pooled set. The first approach gives each stock equal weight and hence does not overemphasize stocks for which there are many assessments, while the second treats each year data point as independent of all others. Neither approach is ideal given that lack of independence among the data. However, the outcomes of applying the two methods are very similar $\sigma = 0.379$ and 0.358 for the two approaches respectively.

Figure 5 shows the distributions of residuals for the four groups. The distribution for rockfish is closest to normal, while those for roundfish and flatfish exhibit less kurtosis than the distribution for rockfish; the distribution for CPS species exhibits a tail to the left. The pooled estimates of $\sigma$ are, however, similar (Table 3).

Table 3. Comparison of different methods of pooling stock-specific variance estimates. Approach $a$ weights each species equally, whereas approach $b$ weights each data point equally. Reported values are estimates of $\sigma$.

<table>
<thead>
<tr>
<th>Group</th>
<th>Approach $a$</th>
<th>Approach $b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rockfish (n=8)</td>
<td>0.442</td>
<td>0.418</td>
</tr>
<tr>
<td>Roundfish (n=4)</td>
<td>0.269</td>
<td>0.281</td>
</tr>
<tr>
<td>Flatfish (n=2)</td>
<td>0.301</td>
<td>0.299</td>
</tr>
<tr>
<td>CPS (n=2)</td>
<td>0.328</td>
<td>0.339</td>
</tr>
</tbody>
</table>

While the point estimates of $\sigma$ differ among groups to some degree, the sample sizes are also very small. To explore whether the data provide support for treating each group separately, the estimates of $\sigma^2$ were fitted using a linear mixed model with group as a random effect. This analysis provided no evidence in support of stratifying by group, i.e., the point estimate of the between-group variance in $\sigma^2$ was essentially zero ($< 10^{-5}$).
Given the lack of support for between-group variability in $\sigma$, the need to treat each species as a replicate is not great and approach $b$ (pooling all of the residuals and finding the standard deviation of the pooled set) seems most justified. This leads to an estimate of 0.36 for $\sigma$ (to two significant places). Assuming the residuals are independent, an approximate 95% confidence interval for $\sigma$ is [0.342, 0.374].

Figure 5. Composite distributions of log-deviations from the mean for rockfish (n=8), roundfish (n=4), flatfish (n=2), and CPS (n=2).
Discussion

We evaluated three methods of quantifying the scientific uncertainty in groundfish and coastal pelagic species stock assessments that have been conducted over the last 20 years for the Pacific Fishery Management Council. We concluded that measurement of log-scale variability as deviations from the mean of biomass estimates in common years from past assessments is an appropriate analytical approach to measuring variation in biomass. Moreover, a comparison of stock- and group-specific estimates of $\sigma$ indicated that a single value of $\sigma_{total} = 0.36$ is a reasonable proxy for all groundfish and CPS stocks. That value translates into a CV on the arithmetic scale of 37$. Of the 17 stocks listed in Table 2, only Pacific sardine yielded a Hessian-based “within” CV that is greater (41%). On first principles variance within cannot be greater than total variance. Therefore, for sardine a stock-specific relationship with $\sigma = 0.39$ may better represent biomass uncertainty.

A preliminary evaluation of uncertainty based on an analysis of states of nature contained in groundfish decision tables was also conducted, although a complete analysis could not be accomplished in time for this assessment cycle. Still, in three of the nine cases examined, the estimate of $\sigma_{decision}$ exceeded $\sigma_{total} = 0.36$, whether calculated using approach A (low, base, and high) or approach B (low and base only). For a fourth species (yelloweye rockfish), only the non-preferred approach A produced a greater estimate of variation. In all but two cases (one being Pacific ocean perch, for which the decision table is based upon a Bayesian posterior), approach A yields a smaller variance estimate than approach B, reflecting the tendency for decision tables to express greater uncertainty in the direction of lower biomass. We view these preliminary findings as promising and recommend that a thorough analysis of statistically weighted states of nature produced for stock assessments be considered as an alternative approach to characterization of scientific uncertainty. Therefore, future decision tables should: (1) clearly define the exact probabilities of all specified states of nature and (2) should include a measure of summary or exploitable biomass in the calculations.

To illustrate how an estimate of $\sigma$ can be used to quantify scientific uncertainty, and thus form the basis of an ABC control rule, we back-transform to the arithmetic scale a lognormal distribution with $\sigma = 0.36$ (Fig. 6). Note that half of the probability density is below a value of 1.00, which represents the mode of the distribution. One can then select a cumulative probability less than 0.50 that maps onto a multiplier that can be interpreted as a reduction from the point estimate of the mean of the distribution (Fig. 7). For example, 40% of the probability density is found at values $\leq 0.913$. If one assumes that the mode of the lognormal distribution (1.00) is indicative of the best point estimate of catch (= OFL), 91% of that amount would associated with a 0.40 probability of exceeding the OFL. Of course an actual policy decision will need to be made as to an appropriate level of $P^*$ (the probability of overfishing), whether it be 0.40 or some other value. Likewise, this example does not include scientific uncertainty attributable to sources other than current year biomass. The groundfish and CPS subcommittees recommend that a concerted effort be made to better characterize forecast and harvest rate uncertainty so that those components of variance can be accounted for in the next management cycle.
Figure 6. Lognormal probability distribution with $\sigma = 0.36$ exponentiated to the arithmetic scale.

Figure 7. Relationship between the probability of overfishing ($P^*$) and an appropriate buffer between the ABC and the OFL, based on $\sigma_{total} = 0.36$. 
Acknowledgments

This study was developed by the Groundfish and Coastal Pelagic Species subcommittees of the Pacific Fishery Management Council’s Scientific and Statistical Committee. Stephen Ralston conducted the analyses and prepared the document, which was then reviewed by subcommittee members, including Tom Barnes, Loo Botsford, Ray Conser, Martin Dorn, Owen Hamel, Selina Heppell, Tom Jagielo, Andre Punt, Theresa Tsou, and Vidar Wespéstad. In addition, a number of people mined the primary PFMC stock assessment literature and provided the SSC with the various time series of abundance for use in the meta-analysis. In particular, we would like to offer our sincerest gratitude and thanks to Jason Cope (cabezon), Paul Crone (Pacific mackerel), John DeVore (sablefish and Dover sole), John Field (bocaccio), Melissa Haltuch (petrale sole), Owen Hamel (darkblotched rockfish, lingcod, POP, and shortspine thornyhead), Xi He (widow rockfish), Kevin Hill (Pacific sardine), Ian Stewart (canary rockfish, Pacific whiting, and yelloweye rockfish), and John Wallace (yellowtail rockfish).

References


Crone, P.R., etc. 1999. canary


Dorn, M.W., etc. 1993. hake

Dorn, M.W., etc. 1999. hake


Dorn, M.W. 1994. hake

Dorn, M.W. 1995. hake

Dorn, M.W. 1996. hake


Hamel, O. 2005. SST

Hamel, O. In press. POP

Hamel, O., and I.J. Stewart. In press. hake

Hamel, O., etc. In press. lingcod


Helser, T., etc. 2002. hake

Helser, T., etc. 2004. hake

Helser, T., etc. 2005. hake

Helser, T., etc. 2006. hake

Helser, T., etc. 2008. hake

Helser, T., and S. Martell. 2007. hake


Ianelli, J., etc. 1992. POP

Ianelli, J., etc. 1994. SST

Ianelli, J, and M. Zimmerman. 1998. POP

Jagielo, T., etc. 2000. lingcod

Jagielo, T., etc. 2003. lingcod

Jagielo, T., and J. Wallace. 2005. lingcod


evaluation. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, OR.


Methot, R., and K. Piner. 2002. canary


Piner, K., and R.D. Methot. 2001. SST


Rogers, J.B. 2003. darkblotched
Rogers, J.B. 2005. darkblotched

Sampson, D.B. 1996. canary

Sampson, D.B. 2005. The status of Dover sole off the U.S. west coast in 2005


Sampson, D.B., and E. Stewart. 1994. canary


Tagart. 1991. yellowtail
Tagart. 1993. yellowtail

Tagart, and Wallace. 1996. yellowtail

Tagart, Ianelli, Hoffmann, and Wallace. 1997. yellowtail

Tagart, Wallace, and Ianelli. 2000. yellowtail


Wallace, J., and O. Hamel. In press. darkblotched
