The Management Strategy Evaluation Approach and the Fishery for Walleye Pollock in the Gulf of Alaska

Z. Teresa A'mar
University of Washington, Quantitative Ecology and Resource Management, Seattle, Washington

André E. Punt
University of Washington, School of Aquatic and Fishery Sciences, Seattle, Washington

Martin W. Dorn
NOAA, NMFS Alaska Fisheries Science Center, Seattle, Washington

Abstract

Management strategy evaluation (MSE) is the process of using simulation testing to examine the robustness of candidate management strategies to error and uncertainty. MSE involves using (1) a model (the "operating model") to represent the true underlying dynamics of the resource and to generate future data, (2) an estimation model to assess the state of the stock relative to agreed target and limit reference points based on the data simulated using the operating model, and (3) a decision rule to determine management actions (e.g., the acceptable biological catch, ABC) given the results of the estimation model. The latter two steps constitute the management strategy. The parameters of the management strategy can be selected to attempt to satisfy desired (but conflicting) management goals and objectives. The results of an MSE are performance measures that quantify the effectiveness of the estimation model and, more generally, the management strategy. MSE is used in this paper to evaluate the extent to which the current management strategy for the fishery for walleye pollock, *Theragra chalcogramma*, in the Gulf of Alaska (GOA) is able to satisfy the management objectives of avoiding low stock size and achieving high, stable catches, given error and uncertainty regarding the data used for assessment purposes, and the form of the stock-recruitment relationship. The results of the analyses indicate that the current management strategy for GOA pollock appears to meet the management sustainability goals when the actual resource dynamics are consistent with the current stock assessment and for two scenarios with alternative stock-recruitment relationships.

Introduction

Management strategy evaluation (MSE) is the process of using simulation testing to determine how robust feedback-control management strategies are to measurement and process error, and model uncertainty (Smith 1994). For the purposes of MSE, a management strategy is the combination of the procedures related to how a resource is monitored, how its status relative to target and limit reference points is determined (the “estimation model”), and how the results of the estimation model are used to determine management actions (the “decision” or “catch control” rule). A management strategy can be complex, involving a stock assessment model coupled with a decision rule (such as that adopted for the management of commercial whaling by the International Whaling Commission; e.g., Cooke 1999, International Whaling Commission 1999), or it can be very simple, such as a decision rule that uses empirical data (e.g., those adopted for anchovy and sardine off South Africa, De Oliveira et al. 1998). In fact, any quantitative method that determines management actions, such as limits on fishing mortality or acceptable biological catch (ABC), gear restrictions, or spatial or temporal limitations, could be evaluated using MSE. The focus of MSE has been on management of single-species fisheries, but there is no reason that management strategies designed primarily to achieve ecosystem objectives could not be evaluated using the MSE approach (Sainsbury et al. 2000, Butterworth and Punt 2003).

The results of an MSE are performance measures that quantify the extent to which a management strategy is able to satisfy the (often conflicting) management goals (Kell et al. 2006). In addition, the results of an MSE can be used to determine how well estimation models are able to estimate quantities (such as current biomass) that are of interest to management. MSE has been used to evaluate current and alternative management strategies for many fisheries worldwide, including those for South African sardine and anchovy (De Oliveira et al. 1998, De Oliveira and Butterworth 2004), prawns off northern Australia (Dichmont et al. 2006), species in Australia’s South East multispecies fishery (Punt et al. 2002), krill off Antarctica (Constable 2005), and flatfish in the Northeast Atlantic (Kell et al. 2005).

MSE has been argued to be an integral part of implementing the UN FAO Code of Conduct for Responsible Fisheries (FAO 1996), because it allows management strategies that are robust to several forms of uncertainty to be identified (Punt 2006). Specifically, MSE analyses can be used to evaluate the extent to which the performance of a management strategy is impacted by uncertainty, in the form of process error (e.g., recruitment variability), observation error (i.e., imprecise or inadequate data collection), and model error (e.g., uncertainty about the true form of the stock-recruitment relationship) (Francis and Shotton 1997). Model error has been identified as a potentially major, but often overlooked, source of uncertainty when considering management decisions, and particularly the selection of management strategies (Butterworth and Punt 1999).

Although the Gulf of Alaska (GOA) walleye pollock fishery has been a fully domestic U.S. fishery since the mid-1980s, it began as a directed foreign fishery in 1964 (Dorn et al. 2005). This fishery is the second largest directed fishery in the Gulf of Alaska, with annual catches between 50,000 and 120,000 metric tons (t) annually since 1986 (Dorn et al. 2005). It is managed by the North Pacific Fishery Management Council (NPFMC) based on scientific advice provided by the National Marine Fisheries Service (NMFS). The management strategy used for the GOA pollock fishery is based upon the “Tier 3 NPFMC decision rule” (NPFMC 2005a). This decision rule determines the target level of fishing mortality each year as a function of the size of the spawning biomass expressed relative to a reference level (Fig. 1). Since 2002, a slightly more conservative decision rule has been used for walleye pollock in the Gulf of Alaska (Dorn et al. 2001). This decision rule increased the buffer between the overfishing level of fishing mortality and the target level of fishing mortality when the spawning biomass is below the reference spawning biomass level to provide a greater protection against assessment uncertainty.

The management strategy for this fishery can therefore be considered as the combination of Dorn et al.’s (2001) decision rule, which produces the ABC (Eq. 1a, with $\alpha = 0.05$), and an estimation model that fits a population dynamics model using fishery, survey, and biological data to produce estimates of the biological reference points $F_{40\%}$ (the fishing mortality that reduces the spawning biomass-per-recruit to 40% of the average unfished spawning biomass-per-recruit), $SB_{40\%}$ (the spawning biomass associated with $F_{40\%}$), $SB_{47\%}$ (the spawning biomass associated with $F_{47\%}$), and the current spawning biomass. The management plan also includes an overfishing level (OFL), defined in terms of fishing mortality (Eq. 1b). If fishing mortality exceeds the overfishing level, “overfishing,” as defined under the Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA), is said to be occurring. To protect endangered Steller sea lions, which consume walleye pollock, a spawning biomass of 20% of the average unfished spawning biomass,

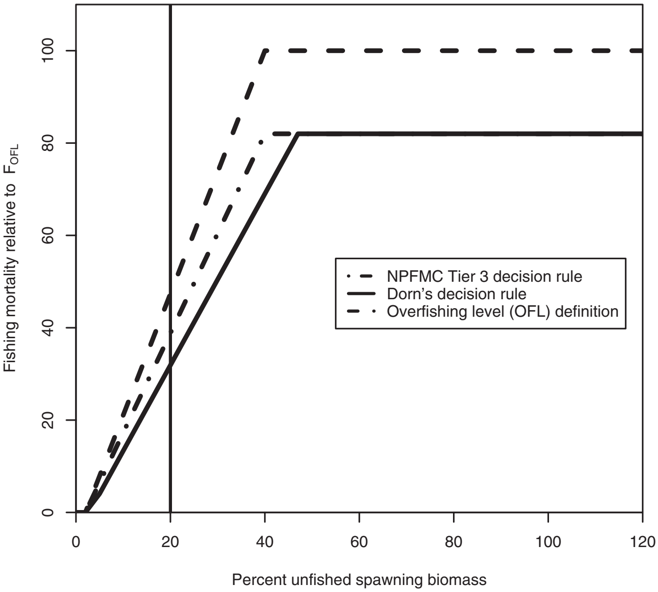

Figure 1. The overfishing level, the North Pacific Fishery Management Council (NPFMC) Tier 3 maximum level of fishing mortality, and the Dorn et al. (2001) decision rule.

$SB_{20\%}$, (the vertical line in Fig. 1), has been established as a level below which no directed fishing would be allowed.

$$F_{ABC} \le \begin{cases} F_{40\%} & \text{if } SB/SB_{47\%} > 1 \\ F_{40\%}\left[\left(SB/SB_{47\%} - \alpha\right)/\left(1-\alpha\right)\right] & \text{if } \alpha < SB/SB_{47\%} \le 1 \\ 0 & \text{if } SB/SB_{47\%} \le \alpha \end{cases} \tag{1a}$$

$$F_{OFL} = \begin{cases} F_{35\%} & \text{if } SB/SB_{40\%} > 1 \\ F_{35\%}\left[\left(SB/SB_{40\%} - \alpha\right)/\left(1-\alpha\right)\right] & \text{if } \alpha < SB/SB_{40\%} \le 1 \\ 0 & \text{if } SB/SB_{40\%} \le \alpha \end{cases} \tag{1b}$$

The spawning biomass reference points (i.e., $SB_{40\%}$ and $SB_{47\%}$) are estimated by multiplying mean recruitment since 1977 by the corresponding spawning biomass–per-recruit (i.e., corresponding to $F_{40\%}$ and $F_{47\%}$). Use of recruitment during the post-1977 time period is intended to represent current stock productivity following a widely recognized cli-

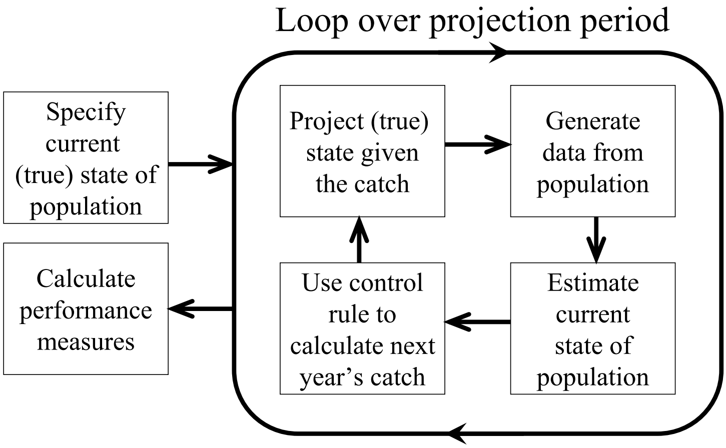

Figure 2. Outline of the key steps of a management strategy evaluation (MSE).

matic regime shift that occurred in 1977. Fishing mortality on GOA pollock since 1977 has varied, but has generally been considerably lower than $F_{40\%}$, so it is unlikely that recruitment during this period provides a suitable estimate of mean recruitment when fishing at $F_{40\%}$, except if recruitment is independent of stock size. An advantage of the MSE approach is that decision rule performance (which includes estimation of biological reference points) can be evaluated both for this idealized situation and for more realistic situations such as when recruitment depends on stock size.

The impetus for this study comes from the 2002 Goodman report (Goodman et al. 2002), which reviewed the management strategies used by the NPFMC for groundfish stocks, and recommended that "an MSE analysis . . . be undertaken to provide additional assurance that the current NPFMC ABC harvest strategy is a robust one and is likely to continue to meet the objectives of the Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA) and of NPFMC itself." This paper therefore outlines the framework used to evaluate management strategies for the GOA pollock fishery. It then uses this framework to evaluate the current management strategy for the case in which the actual resource dynamics are similar to the assumptions that underlie the current stock assessment and harvest policy. This provides a first test of the current management strategy for GOA pollock.

Methods

Overview and operating model

Fig. 2 outlines the key components of an MSE. The operating model (see Appendix for the technical details of this model) represents the "true" state of the resource, and includes hypotheses regarding how the age-structure of the resource changes through time, and the form of the underlying stock-recruitment relationship. The operating model is age-structured, can relate recruitment to spawning biomass or estimate each independently, and includes process error in both recruitment and fishery selectivity. It is similar in structure to the population dynamics model on which the 2005 GOA pollock stock assessment (Dorn et al. 2005) was based, primarily because the latter represents the best available science regarding the status of this resource. The main difference between the operating model and the original stock assessment model is the age range that each model covers. The operating model covers ages 1 through 15 years, and the stock assessment model covers ages 2 through 10 years, with the oldest ages being the "plus" groups. Some proportion-at-age data are available for ages 1 through 15+ (e.g., the fishery for 1985, 1989-1990, and 1992-2004; the NMFS echo-integration trawl (EIT) survey for 1981, 1983-91, 1994-98, and 2000-05; and the NMFS GOA bottom trawl surveys for years 1984, 1987, 1989, 1990, 1993, 1996, 1999, 2001, and 2003), which is the reason for the age range chosen for the operating model. Both models estimate the parameters of the stock-recruitment relationship, the annual fishing mortalities, the catchability coefficients for each survey, fishery and survey selectivity, recruitment deviations, annual deviations for the fishery selectivity parameters, and the biological reference points needed to apply the decision rule.

There are two main steps to using an operating model to evaluate management strategies: (1) the operating model is fit to historical fishery and survey data using Bayesian methods based on the Markov chain Monte Carlo algorithm (Hastings 1970, Gelman et al. 2004) to determine the best estimates for the values for its parameters, and hence the numbers-at-age at the start of the projection period (2006) and to quantify the uncertainty associated with the estimates of these quantities; and (2) the operating model is used to project the simulated stock forward for a series of draws from the Bayesian posterior distribution when management is based on a simulated management strategy. Step 1, often referred to as conditioning, is undertaken to ensure that the simulations are representative of the actual problem of managing the GOA pollock fishery, as this situation is currently understood through stock assessments.

During the projection period, the operating model generates the survey and fishery data that are used by the estimation model component of the management strategy, and characterizes the impacts, from a single-species perspective, of the ABCs set using the management strategy. In the projections, the recruitment process error, and the observation error applied to the "true" survey indices of abundance, survey catch proportions-at-age, and fishery catch proportions-at-age, are assumed to be temporally uncorrelated.

For the purposes of this study, 100 simulations were conducted for each scenario regarding the specifications of the operating model. This number of simulations was selected because it was sufficient to determine differences among alternative strategies; several previous studies have been based on a similar number of simulations per scenario. Each simulation involved projecting the simulated stock forward for 30 years (2006-2035) by annually applying the estimation model and decision rule, and then updating the "true" population dynamics. Thirty years was selected for the length of the projection period based on the suggestion of Goodman et al. (2002).

Management strategy

The management strategy considered in this paper is the combination of the decision rule in Fig. 1 and the stock assessment model on which the 2005 GOA pollock stock assessment was based (Dorn et al. 2005). This estimation model, which is similar to the operating model when the operating model assumes recruitment to be independent of spawning biomass, except for the age range considered, is designed to provide the input needed to apply the decision rule that produces the F_{ABC} and hence the ABC, i.e., the estimates of $F_{40\%}$, $SB_{47\%}$, and current spawning biomass (see Eq. 1a). With one exception, the catch removed from the simulated stock is set to the ABC calculated using the management strategy, i.e., implementation error is ignored. The exception is that the catch is set to 0.1% of the age 3+ biomass if the spawning biomass is assessed to be below $SB_{20\%}$ (rather than zero). This level of catch reflects the current levels of pollock bycatch in GOA fisheries other than the directed pollock fishery, and represents what is likely to happen in the event that the spawning biomass of pollock is assessed to be below $SB_{20\%}$ and the directed pollock fishery was closed.

Performance measures

The performance measures consist of two types ("estimation" and "management"). The estimation performance measures examine how well the estimation model estimates annual spawning biomass, annual fishing mortality, the fishing mortality needed to apply the decision rule, the biological reference points, and the ABC based on the decision rule. The "estimation" performance measures were selected because they assess the ability of the assessment model to provide the information needed to set the ABC in the following year and to determine the status of the stock relative to biological reference points. The "management" performance measures are selected based on the goals and objectives of the NPFMC (NPFMC 2005b) and the MSFCMA. They are the number of simulations in which the "true" spawning biomass falls below the "true" $SB_{20\%}$, the number of simulations in which the "true" fishing mortality is above the "true" maximum fishing mortality threshold (MFMT), F_{OFL} (Eq. 1b), and the average catch over the projection period. The "management" performance measures were selected because they assess the ability of the management system to prevent the stock from becoming overfished, to prevent overfishing, and to maintain catches at sustainable levels.

The results are shown graphically rather than in tabular form (which would be extremely voluminous). Specifically, the ability of the management strategy to leave the spawning biomass close to the reference level of $SB_{40\%}$ and achieve high catches is summarized using a plot with the following six panels:

1. A summary of the distribution of the time-trajectory of the true spawning biomass;
2. A summary of the distribution of the time-trajectory of annual catches;
3. The time-series of the true spawning biomass for simulations 25, 50, 75, and 100;
4. The time-series of catches for simulations 25, 50, 75, and 100;
5. A summary of the distribution of the time-trajectory of the ratio of the true spawning biomass to the true reference level of $SB_{40\%}$; and
6. A histogram of the average catch during the projection period.

The estimation performance of the management strategy in assessing current stock status is summarized by plots that show the distributions for the percentage relative errors for the time-series of (1) spawning biomass and (2) fishing mortality in the final year of the assessment period; (3) the reference spawning biomass, $SB_{40\%}$ (NPFMC 2005a); (4) the target fishing mortality, $F_{40\%}$ (NPFMC 2005a); (5) the fishing mortality on which the ABC for the following year is based, F_{ABC} (see Eq. 1a); and (6) the ABC under the decision rule.

Scenarios considered

This study considers three scenarios: a base scenario in which recruitment is lognormally distributed about a mean value (Eq. A.3) so there is no impact of a reduction in spawning biomass on expected recruitment, and two scenarios in which recruitment declines with reductions in spawning biomass (Eq. A.2a and A.2b). The form of the relationship between spawning biomass and recruitment forms the focus for this study because the form of this relationship has been identified in the past as a factor to which the performance of management strategies has been found to be sensitive (Butterworth and Punt 1999). The true value of $SB_{40\%}$ used when calculating the performance measures is defined as the product of the spawning biomass-per-recruit corresponding to $F_{40\%}$ and the recruitment expected at $F_{40\%}$.

Results and discussion

Base scenario

The base scenario involves the operating model being based on the same assumption about the form of the stock-recruitment relationship as the management strategy, i.e., that recruitment and spawning biomass are unrelated. It would be expected therefore that the management strategy would perform adequately for this scenario. The resource is maintained at the spawning biomass reference level on average for this scenario (Fig. 3, upper center panel), and there is only a small probability (≤5%) that the spawning biomass is reduced to below $SB_{20\%}$ (Fig. 4b). There is, however, a high probability (80%) that the true fishing mortality is above the overfishing fishing mortality reference level, F_{OFL}, in the fourth and fifth years of the projection period (Fig. 4c), but this probability decreases rapidly and is much lower (<20%) during most of the projection period. North Pacific harvest policies establish a buffer between the reference fishing mortality and F_{OFL}, but some risk of unintentionally exceeding limits is unavoidable given the uncertainties in the assessment. The implications of those risks, in terms of reduction in biomass and loss in yield, could be evaluated using the results of an extension to the outputs of the current MSE.

The high level of fishing mortality in the earliest years of the projection period appears to be due to the transition from the actual data for GOA pollock to data simulated using the operating model, which leads to positive biases in the estimates of F_{ABC} and ABC for the earliest years of the projection period (Fig. 5). Bias in the estimates of spawning biomass, and particularly fishing mortality, occurs in the first few years of the projection period because, for example, the signs of the differences between the estimated biomass and an index of abundance for a year before 2006 will tend to be the same because the data for the years prior to 2006 are set to the actual values, while the differences will be centered around zero for the future (see, for example, Fig. 6). Transient behavior in the first years of the projection period for MSE studies has been observed in the past for other species (e.g., Mapstone et al. 2004). The bias of approximately 5% in $F_{40\%}$ (see the lower center panel of Fig. 5) occurs because of the difference in the plus-group ages in the operating and estimation models.

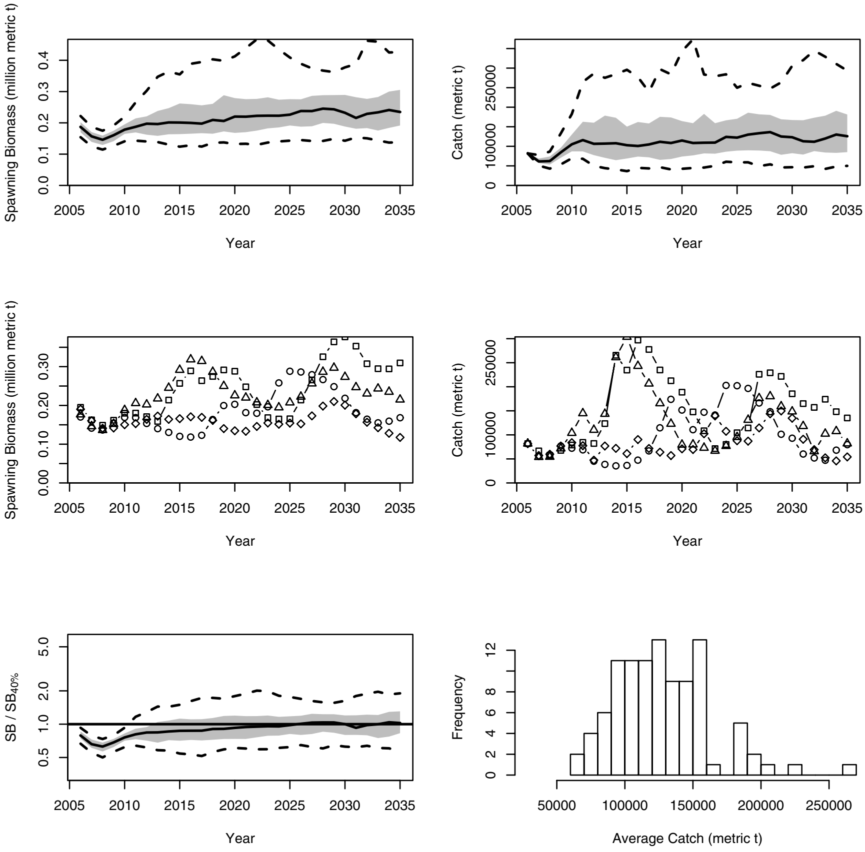

Figure 3. **Results for the base operating model for the "true" spawning biomass, catch, four realizations of spawning biomass trajectories, four realizations of catch trajectories, the relative level of "true" spawning biomass (on a logarithmic scale), and the distribution of average catch over the 100 simulations. For the envelopes, the solid black line is the median, the shaded area covers the 25th through the 75th percentiles, and the dashed lines are the 5th and 95th percentiles.**

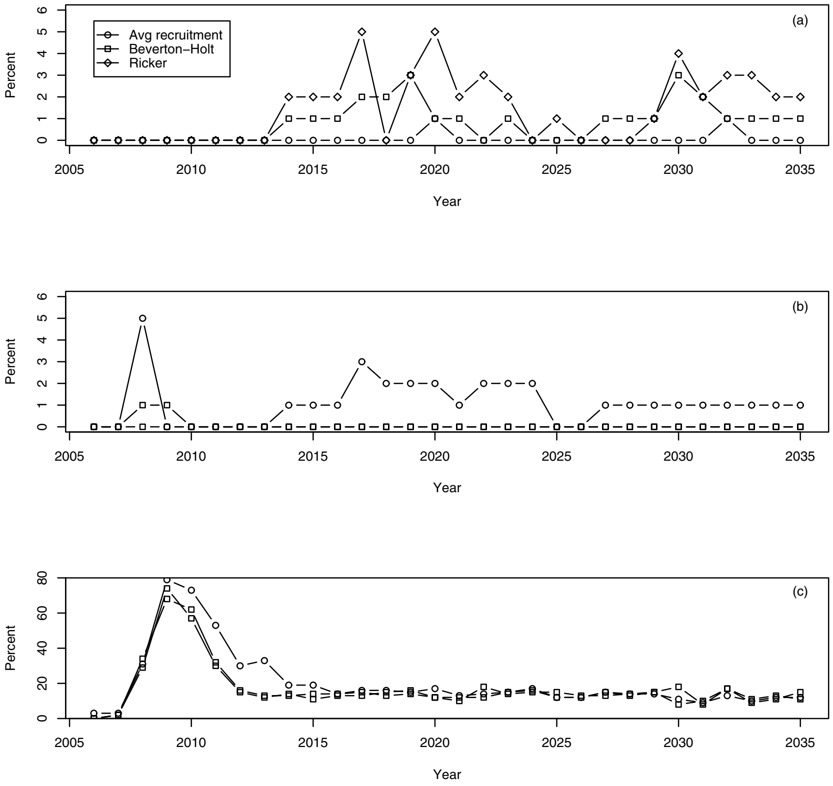

Figure 4. Number of simulations in each year where the estimated spawning biomass was below the estimated SB_{20} (a); number of simulations in each year where the "true" spawning biomass was below the "true" $SB_{20\%}$ (b); and number of simulations in each year where the "true" fishing mortality was above the "true" maximum fishing mortality threshold (MFMT), F_{OFL} (c). The results in this figure are based on the three operating model scenarios.

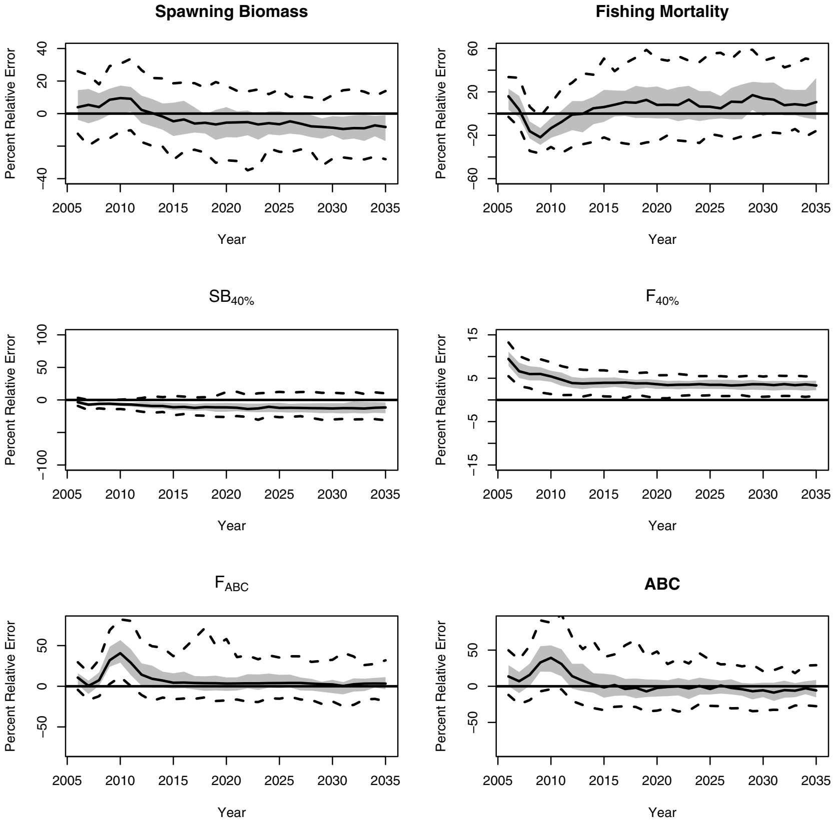

Figure 5. Results for the base operating model showing the percent relative error between the estimated and "true" values for spawning biomass, fishing mortality, $SB_{40\%}$, $F_{40\%}$, F_{ABC}, and ABC. The solid black line is the median, the shaded area covers the 25th through the 75th percentiles, and the dashed lines are the 5th and 95th percentiles.

The average level of catch during the projection period is 127,000 t, which is greater than the average level of catch during 1964-2005 (84,000 t), and greater than the average level of catch since 1986 (82,000 t). The catch variability (standard deviation of catch) over the projection period is 80,000 t, which is greater than the catch variability over the years 1964-2005 (67,000 t), and much greater than that since 1986 (20,000 t). Differences between the average catch during the projection period and during the historical period are due to several factors. First, the current control rule has been used only since 2001. Prior to 2001, a variety of rationales were used to arrive at an ABC, most of which

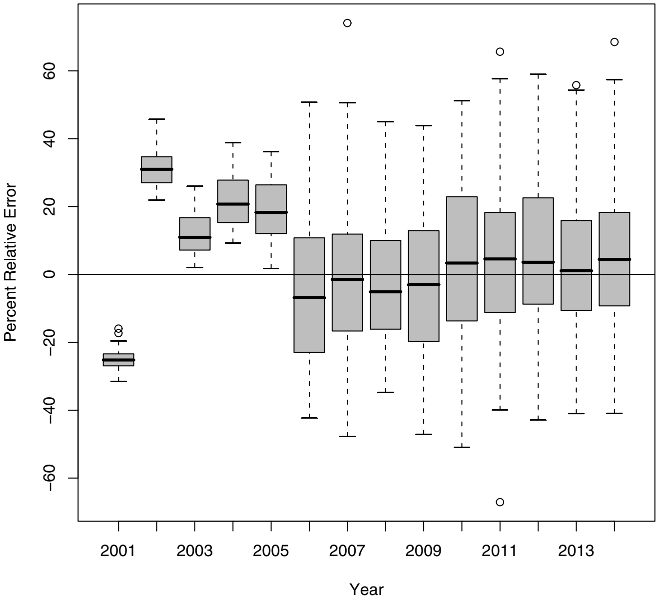

Figure 6. Box and whisker plots of the percent relative errors for the NMFS winter Gulf of Alaska (GOA) Shelikof Strait echo-integration trawl (EIT) survey indices (2001-2014) based on simulated assessments conducted in 2015. The indices for the years prior to 2006 are the actual data, while those from 2006 onward are generated by the operating model.

resulted in harvest rates lower and more stable than would be recommended under the current harvest policy. Second, during the early period of the fishery, catches were constrained more by capacity and interest than by management actions. In fact, average catches during the projection period correspond well to equilibrium catches under constant $F_{40\%}$ and $F_{50\%}$ fishing mortality rates reported by Dorn et al. (2005), indicating that the projection model is performing as expected.

The biases in Fig. 5 appear to be due primarily to the transition between actual and simulated data and differences in the fishery and survey selectivity functions between the operating model and the estimation model, as the operating model covers ages 1 through 15 and the estimation model considers only ages 2 through 10. In general, the biases of the estimation model are modest and do not compromise the implementation of appropriate harvest limits.

The results in Figs. 3-5 confirm the expectation that the management strategy appears to perform adequately, and as expected, when

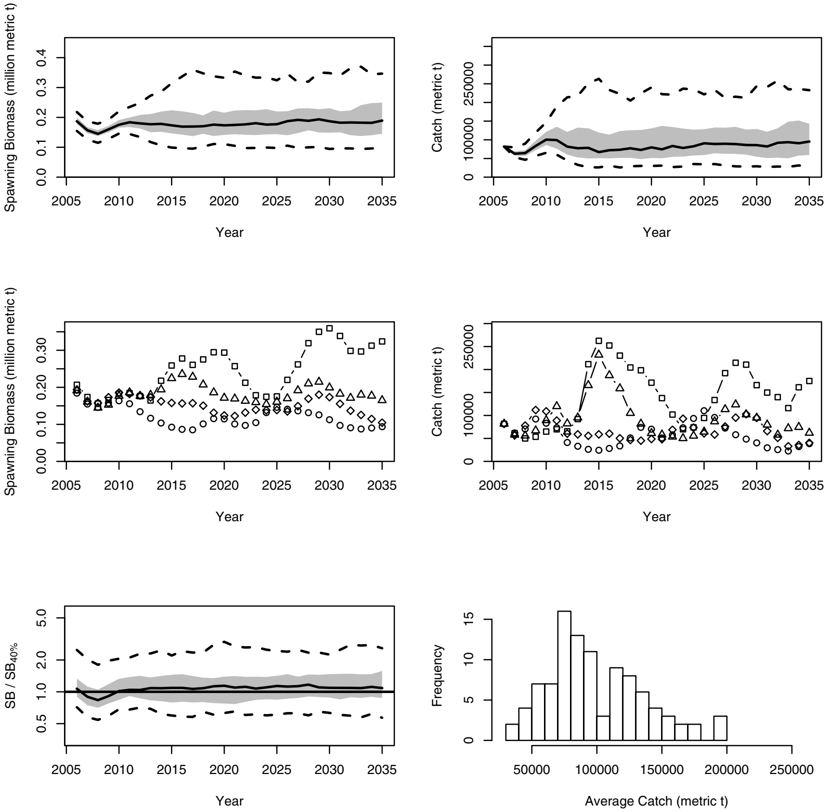

Figure 7. **As for Fig. 3, except that the results pertain to the operating model in which recruitment is governed by the Beverton-Holt stock-recruitment relationship.**

the dynamics of the resource nearly exactly correspond to the assessment model assumptions and mean recruitment does not decline with declining spawning biomass.

Sensitivity to the form of the stock-recruitment relationship

Figs. 7 and 8 summarize the management performance of the management strategy when a stock-recruitment relationship is included in the operating model. Estimated stock recruit relationships (Fig. 9) indicate that recruitment declines with decreasing stock size (the posterior modal estimates of steepness are respectively 0.647 and 0.465 for the Beverton-Holt and Ricker stock-recruitment relationships).

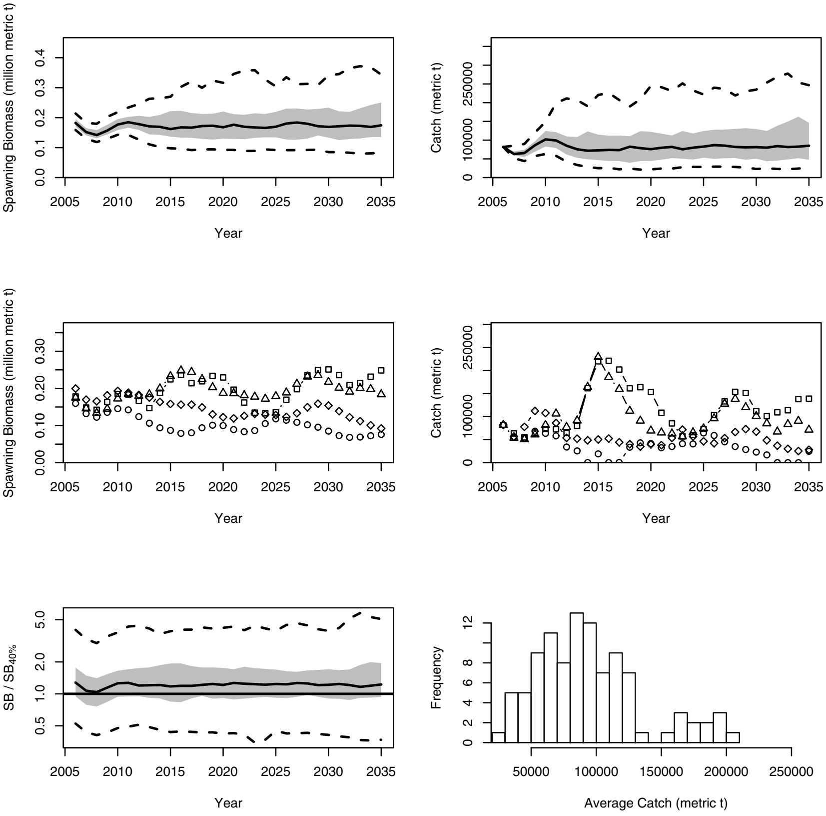

Figure 8. **As for Fig. 3, except that the results pertain to the operating model in which recruitment is governed by the Ricker stock-recruitment relationship.**

The stock is left above the spawning biomass reference level on average for both scenarios in which recruitment declines with reductions in spawning biomass. However, $SB_{40\%}$ for these scenarios is lower than $SB_{40\%}$ for the base scenario, as the level of recruitment expected under $F_{40\%}$ is lower for these scenarios than the average level of recruitment used to calculate $SB_{40\%}$ for the base scenario (Fig. 9).

The probability that the stock really dropped below the true $SB_{20\%}$ is lower for both scenarios in which recruitment declines with reductions in spawning biomass than for the base scenario. The converse is true for the estimated values, although the probability across all scenarios is ≤ 5% (Fig. 4). The reason for the lower probability of being below the

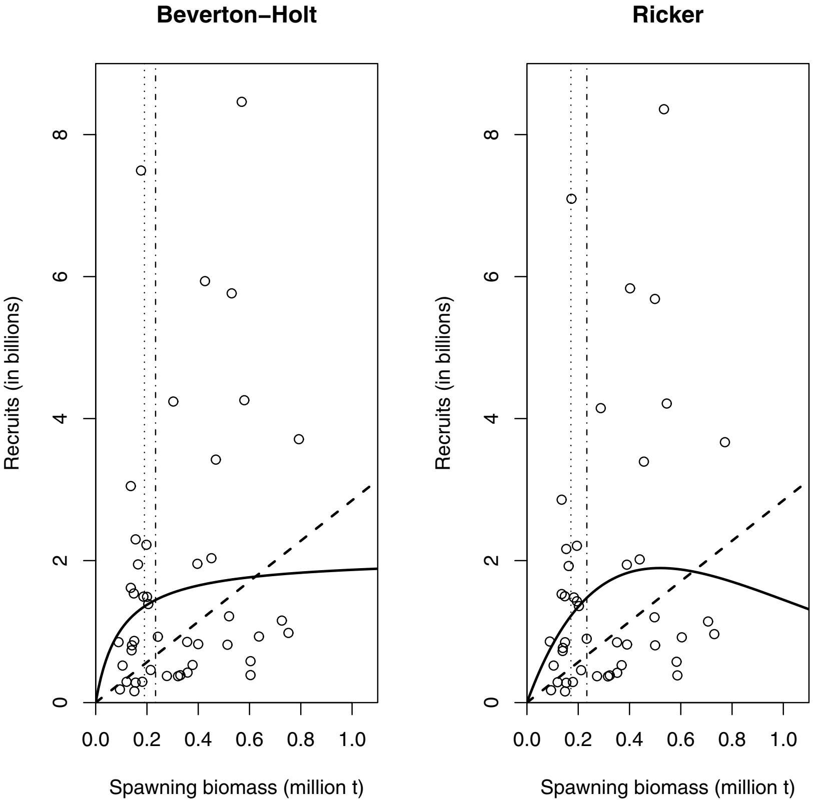

Figure 9. The stock-recruitment relationship along with spawning biomass and recruitment over the historical period based on the values for the parameters corresponding to the mode of the posterior distribution. Results are shown for the Beverton-Holt and Ricker scenarios. The dashed lines are the replacement lines corresponding to $F = 0$, the dotted vertical lines indicate $SB_{40\%}$ for each scenario, and the dot-dash vertical lines indicate $SB_{40\%}$ for the base operating model.

true $SB_{20\%}$ is that catches are adjusted much lower for the scenarios in which recruitment is related to spawning biomass according to a stock-recruitment relationship (contrast Figs. 3, 7, and 8) which is in turn related to positive bias in the estimates of $SB_{40\%}$ (see Figs. 10 and 11). One consequence of this bias is that there is an increased probability that the estimation model will assess the stock to be below $SB_{20\%}$ (see the upper panel of Fig. 4). The probability that the true fishing mortality exceeds the overfishing level is similar to that for the base scenario.

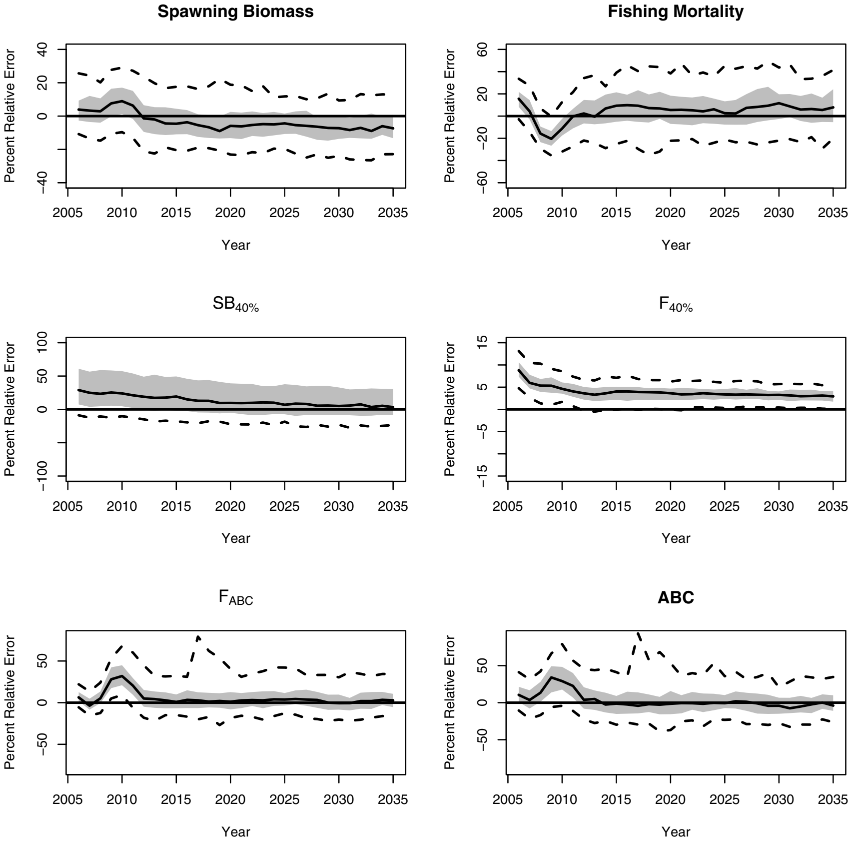

Figure 10. As for Fig. 5, except that the results pertain to the operating model in which recruitment is governed by the Beverton-Holt stock-recruitment relationship.

This is because the overfishing level (Eq. 1b) is defined in terms of fishing mortality rates that lead to particular reductions in spawning biomass-per-recruit rather than spawning biomass. The average level of catch during the projection period is 99,000 t (Beverton-Holt) and 95,000 t (Ricker), i.e., notably lower than for the base scenario. This result is not unexpected given that the population is less resilient when recruitment declines with reductions in spawning biomass. Catch variability, expressed as the standard deviation in catch, is less for the Beverton-Holt and Ricker relationships (64,000 and 65,000 t, respectively) than for the base scenario.

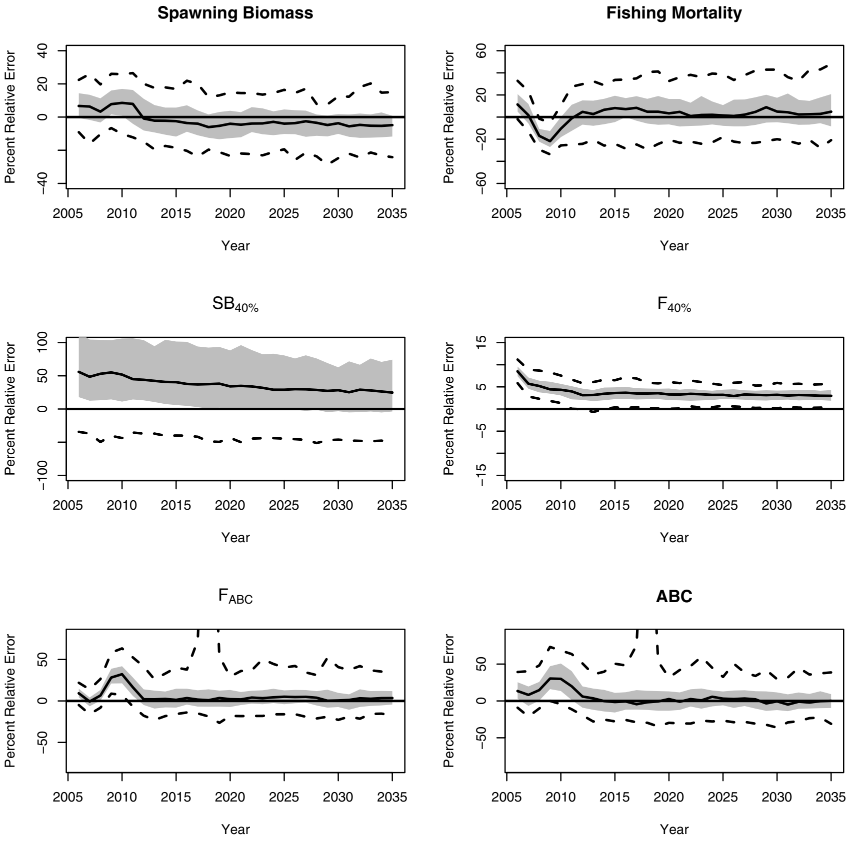

Figure 11. As for Fig. 5, except that the results pertain to the operating model in which recruitment is governed by the Ricker stock-recruitment relationship.

Somewhat surprisingly, the patterns of relative error for the scenarios in which recruitment declines with reductions in spawning biomass (Figs. 10 and 11) are remarkably similar to those for the base scenario. Specifically, spawning biomass is negatively biased and fishing mortality is positively biased by about 5%, $F_{40\%}$ is positively biased by 5%, and F_{ABC} and ABC are virtually unbiased. However, in contrast to the base scenario, $SB_{40\%}$ is positively biased, although this is not unexpected because the estimator of $SB_{40\%}$ assumes that recruitment is independent of spawning biomass.

Overall, performance for these sensitivity tests is generally fairly satisfactory because the spawning biomass is stabilized above the refer-

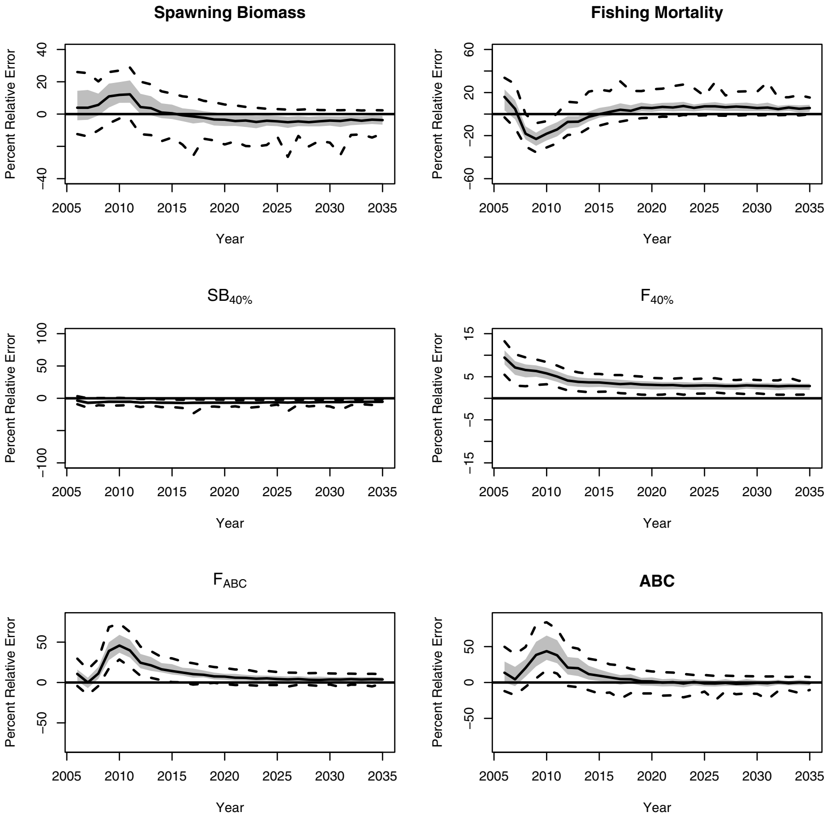

Figure 12. As for Fig. 5, except that the results pertain to the base operating model when there is no observation or process error during the projection period.

ence level and the probability of being below *SB*$_{20\%}$ is lower than for the base scenario. Given the differences in average catch levels, it seems clear that the decision rule is adjusting catch limits in the correct direction even though it is making incorrect assumptions about productivity when recruitment is really related to spawning biomass.

Impact of future observation and process error

An additional series of simulations was conducted in which there was no future process or observation error to more fully understand the behavior of the management strategy (Fig. 12). As expected, there is considerably less inter-annual and inter-simulation variation when there is no error in the future (but the assessment model is not aware of this). The relative errors are somewhat smaller than was the case for the original base scenario. However, the patterns of relative error are almost identical to those for the base operating model with observation and process error. This confirms that the reasons for the biases observed in Figs. 5, 10, and 11 are likely to be due to the structural differences between the operating model and the estimation model and the nature of how the operating model is parameterized rather than to how the future data are generated.

General discussion

The current management strategy for GOA pollock appears to meet the management sustainability goals that can be inferred from the goals and objectives of the NPFMC and the MSFCMA and the estimates of management-related quantities are not unreasonably inaccurate and imprecise. Although this result was anticipated (hoped) when the analyses initiated, it is not always guaranteed that a management strategy will perform as anticipated even under fairly ideal conditions. For example, Kell et al. (2005) performed an MSE on eight North Atlantic fish stocks based on the decision rule used by the International Council for the Exploration of the Sea (ICES). The results of this MSE indicated that the ICES management strategy did not perform well. The poor performance was exacerbated by the time lags in the assessment and monitoring processes. Similarly, using an MSE, Punt and Ralston (2007) examined the possible decision rules that could be inferred from the decisions by the Pacific Fishery Management Council regarding how groundfish resources are managed, and found that the rule that most closely mimics recent decision-making performed appreciably poorer than alternative approaches in terms of inter-annual variation in catch and the need for frequent revisions to rebuilding plans.

This management strategy evaluation dealt only with the current management strategy for GOA pollock, and evaluated whether it is robust to a range of uncertainties. Additional work is planned to evaluate whether other management strategies would improve performance in achieving management goals and objectives.

The results of the simulation exercise of this paper are based on an operating model that is almost identical to the estimation model except that alternative stock-recruitment relationships can be used in the operating model. The analyses of this paper are based on fitting the operating model to the actual data for the resource in question and quantifying parameter uncertainty using samples from a Bayesian posterior distribution. While this ensures that the scenarios considered are constrained by what is understood for GOA pollock, it is questionable whether even the full range of parameter uncertainty is captured. For example, weight-at-age is necessarily assumed to be known exactly for the purposes of sampling from the posterior distribution. Furthermore, basing the projections on samples from the posterior distribution raises the possibility that it will prove impossible for the Markov chain Monte Carlo algorithm to converge. This problem was not encountered in this study. However, given the paucity of data, the values for the parameters of the stock-recruitment relationship on which the analyses of this paper were based were very imprecise.

Future work will explore the robustness of the management strategy to the effects of (plausible) violations of the assumptions of the operating model on which this paper is based. Butterworth and Punt (1999) reviewed the uncertainties considered in past MSE analyses and found that there were no uncertainties that were important in all cases; rather the most important uncertainties were case-specific. In the context of the GOA pollock fishery, the major uncertainties that need to be examined relate to ecosystem changes, multispecies interactions, climate variability, and/or regime shifts. This is because it is unknown how changes in the species composition of the GOA have led to changes in predation on pollock and hence pollock natural mortality (Hollowed et al. 2000, Gaichas 2006). In addition, it has been postulated that climate variability and regime shifts in the GOA may account directly or indirectly for changes in pollock productivity and recruitment (Anderson and Piatt 1999, Bailey 2000). Finally, account could be taken in future work of the difference between TACs and subsequent landings.

Acknowledgments

ZTA would like to acknowledge the financial support of the NMFS Stock Assessment Improvement Program. This publication is partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA17RJ1232, Contribution no. 1423. The comments of two anonymous reviewers are gratefully acknowledged.

References

Anderson, P.J., and J.F. Piatt. 1999. Community reorganization in the Gulf of Alaska following ocean climate regime shift. Mar. Ecol. Prog. Ser. 189:117-123.

Bailey, K.M. 2000. Shifting control of recruitment of walleye pollock *Theragra chalcogramma* after a major climatic and ecosystem change. Mar. Ecol. Prog. Ser 198:215-224.

Blackburn, J., and D. Pengilly. 1994. A summary of estimated population trends of seven most abundant groundfish species in trawl surveys conducted by Alaska Department of Fish and Game in the Kodiak and Alaska Peninsula areas, 1988 through 1993. Alaska Department of Fish and Game, Regional Information Report 4K94-31. 19 pp.

Britt, L.L., and M.H. Martin. 2000. Data report: 1999 Gulf of Alaska bottom trawl survey, NOAA Tech. Memo. NMFS-AFSC-121. 249 pp.

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

Butterworth, D.S., and A.E. Punt. 2003. The role of harvest control laws, risk and uncertainty and the precautionary approach in ecosystem-based management. In: M. Sinclair and G. Valdimarsson (eds.), Responsible fisheries in the marine ecosystem. CAB International, Wallingford, pp. 331-319.

Constable, A.J. 2005. A possible framework in which to consider plausible models of the Antarctic marine ecosystem for evaluating krill management procedures. CCAMLR Science 12:99-117.

Cooke, J.G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES J. Mar. Sci. 56:797-810.

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

De Oliveira, J.A.A., D.S. Butterworth, B.A. Roel, K.L. Cochrane, and J.P. Brown. 1998. The application of a management procedure to regulate the directed and bycatch fishery of South African sardine *Sardinops sagax*. S. Afr. J. Mar. Sci. 19:449-469.

Dichmont, C.M., R. Deng, A.E. Punt, W. Venables, and M. Haddon. 2006. Management strategies for short lived species: The case of Australia's northern prawn fishery. 3. Management strategy scale and complexity. Fish. Res. 82:235-245.

Dorn, M.W., and R.D. Methot. 1990. Status of the coastal Pacific whiting resource in 1989 and recommendation to management in 1990. NOAA Tech. Memo. NMFS-F/NWC-182. 84 pp.

Dorn, M.W., A.B. Hollowed, E. Brown, B.A. Megrey, C. Wilson, and J. Blackburn. 1999. Walleye pollock. In: Stock assessment and fishery evaluation report for groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council, Anchorage, Alaska.

Dorn, M.W., A.B. Hollowed, E. Brown, B.A. Megrey, C.D. Wilson, and J. Blackburn. 2001. Assessment of the walleye pollock stock in the Gulf of Alaska for 2002. In: Stock assessment and fishery evaluation report for groundfish resources in the Gulf of Alaska. North Pacific Fishery Management Council, Anchorage, Alaska.

Dorn, M.W., K. Aydin, S. Barbeaux, B.M. Guttormsen, B.A. Megrey, K. Spalinger, and M. Wilkins. 2005. Assessment of the walleye pollock stock in the Gulf of Alaska for 2006. In: Stock assessment and fishery evaluation report for groundfish resources in the Gulf of Alaska. North Pacific Fishery Management Council, Anchorage, Alaska.

FAO. 1996. Precautionary approach to capture fisheries and species introductions. Elaborated by the technical consultation on the precautionary approach to capture fisheries (including species introductions). Lysekil, Sweden, 6-13 June 1995. FAO Technical Guidelines for Responsible Fisheries No. 2. 62 pp.

Fournier, D. 2006. An introduction to AD Model Builder Version 7.1.1 for use in nonlinear modeling and statistics. Otter Research Ltd., Sidney, BC, Canada. 193 pp.

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

Gaichas, S.K. 2006. Development and application of ecosystem models to support fishery sustainability: A case study for the Gulf of Alaska. Ph.D. dissertation, Univ. Washington, Seattle. 370 pp.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2004. Bayesian data analysis, 2nd edn. Chapman and Hall, London, UK.

Goodman, D., M. Mangel, G., Parkes, T., Quinn, V., Restrepo, T., Smith, and K. Stokes. 2002. Scientific review of the harvest strategy currently used in the BSAI and GOA groundfish fishery management plans. North Pacific Fishery Management Council, Anchorage, Alaska. 153 pp.

Guttormsen, M.A., and P.T. Yasenak. 2006. Results of the February-April 2005 echo integration-trawl surveys of walleye pollock (*Theragra chalcogramma*) conducted in the Gulf of Alaska, cruises MF2005-01 and MF2005-05. NOAA AFSC Processed Rep. 2006-03. 52 pp.

Hastings, W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97-109.

Helser, T.E., M.W. Dorn, M.W. Saunders, and R.D. Methot. 2001. Pacific whiting assessment update for 2000. In: Status of the Pacific Coast groundfish fishery through 2001 and recommended acceptable biological catches in 2002. Pacific Fishery Management Council, Portland, Oregon.

Hollowed, A.B., J.N. Ianelli, and P.A. Livingston. 2000. Including predation mortality in stock assessments: A case study for Gulf of Alaska walleye pollock. ICES J. Mar. Sci. 57:279-293.

International Whaling Commission (IWC). 1999. The revised management procedure (RMP) for baleen whales. J. Cetacean Res. Manag. 1(Suppl.):251-258.

Kell, L.T., J.A.A. De Oliveira, A.E. Punt, M.K. McAllister, and S. Kuikka. 2006. Operational management procedures: an introduction to an evaluation frameworks. In: L. Motos and D. Wilson (eds.). The knowledge base for fisheries management. Elsevier Limited, pp. 379-407.

Kell, L.T., M.A. Pastoors, R.D. Scott, M.T. Smith, F.A. Van Beek, C.M. O'Brien, and G.M. Pilling. 2005. Evaluation of multiple management objectives for Northeast Atlantic flatfish stocks: Sustainability vs. stability of yield. ICES J. Mar. Sci. 62:1104-1117.

Mapstone, B.D., C.R. Davies, L.R. Little, A.E. Punt, A.D.M. Smith, F. Pantus, D.C. Lou, A.J. Williams, A. Jones, G.R. Russ, and A.D. MacDonald. 2004. The effects of line fishing on the Great Barrier Reef and evaluations of alternative potential management strategies. CRC Reef Research Centre Technical Report No. 52, Townsville, Australia. 205 pp.

Martin, M.H. 1997. Data report: 1996 Gulf of Alaska bottom trawl survey. NOAA Tech. Memo. NMFS-AFSC-82. 235 pp.

McKelvey, D.R. 1996. Juvenile walleye pollock, *Theragra chalcogramma*, distribution and abundance in Shelikof Strait: What can we learn from acoustic survey results? In: R.D. Brodeur, P.A. Livingston, T.R. Loughlin, and A.B. Hollowed (eds.), Ecology of juvenile walleye pollock, *Theragra chalcogramma*. NOAA Tech. Rep. 126, pp. 25-34.

NPFMC. 2005a. Introduction to the stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. In: NPFMC Gulf of Alaska stock assessment and fishery evaluation report. North Pacific Fishery Management Council, Plan Team for the Groundfish Fisheries of the Gulf of Alaska, Anchorage, Alaska. 40 pp.

NPFMC. 2005b. Fishery management plan for groundfish of the Gulf of Alaska. North Pacific Fishery Management Council, Anchorage, Alaska. 282 pp.

Picquelle, S.J., and B.A. Megrey. 1993. A preliminary spawning biomass estimate of walleye pollock, *Theragra chalcogramma*, in Shelikof Strait, Gulf of Alaska, based on the annual egg production method. Bull. Mar. Sci. 53:728-749.

Punt, A.E. 2006. The FAO Precautionary Approach after almost 10 years: Have we progressed towards implementing simulation-tested feedback-control management systems for fisheries management? Nat. Resour. Model. 19:441-464.

Punt, A.E., and S.V. Ralston. 2007. A management strategy evaluation of rebuilding revision rules for overfished rockfish species. In: J. Heifetz, J. DiCosimo, A.J. Gharrett, M.S. Love, V.M. O'Connell, and R.D. Stanley (eds.), Biology, assessment, and management of North Pacific rockfishes. Alaska Sea Grant, University of Alaska Fairbanks, pp. 329-351.

Punt, A.E., A.D.M. Smith, and G. Cui. 2002. Evaluation of management tools for Australia's south east fishery. 3. Towards selecting appropriate harvest strategies. Mar. Freshw. Res. 53:645-660.

Sainsbury, K.J., A.E. Punt, and A.D.M. Smith. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES J. Mar. Sci. 57:731-741.

Smith, A.D.M. 1994. Management strategy evaluation: The light on the hill. In: D.A. Hancock (ed.), Population dynamics for fisheries management. Australian Society for Fish Biology, Perth, Western Australia, pp. 249-253.

Sullivan, P.J., A.M. Parma, and W.G. Clark. 1997. Pacific halibut assessment: Data and methods. Int. Pac. Halibut Comm. Sci. Rep. 97. 84 pp.

Appendix: The operating model

Population dynamics

The number of animals aged one and older is governed by the equation:

$$N_{y+1,a} = \begin{cases} N_{y+1,1} & \text{if } a = 1 \\ N_{y,a-1}\, e^{-(M_{a-1}+\tilde{S}_{y,a-1}F_y)} & \text{if } 1 < a < x \\ N_{y,x-1}\, e^{-(M_{x-1}+\tilde{S}_{y,x-1}F_y)} + N_{y,x}\, e^{-(M_x+\tilde{S}_{y,x}F_y)} & \text{if } a = x \end{cases} \tag{A.1}$$

where $N_{y,a}$ is the number of fish of age *a* at the start of year *y*;

M_a is the (time-invariant) instantaneous rate of natural mortality for fish of age *a*;

$\tilde{S}_{y,a}$ is the selectivity of harvesting on fish of age *a* during year *y*;

F_y is the fishing mortality on fully-selected ($\tilde{S}_{y,a} \to 1$) animals during year *y*; and

x is the plus-group (all fish in this age-class are mature and recruited to the fishery—assumed to be age 15).

Three alternative relationships between stock size and the number of subsequent recruits (at age 1) are considered. Recruitment is lower than expected at unfished equilibrium when the spawning biomass is a small fraction of its unfished size for two of these relationships (Beverton-Holt and Ricker):

$$N_{y+1,1} = \frac{4R_1 h SB_y}{\psi_0 R_1 (1-h) + (5h-1) SB_y} e^{\varphi_{y+1}-\sigma_R^2/2} \qquad \varphi_y \sim N(0,\sigma_R^2) \tag{A.2a}$$

$$N_{y+1,1} = \frac{SB_y}{\psi_0} \exp\left[A_r \left(1 - \frac{SB_y}{\psi_0 R_1} \right) \right] e^{\varphi_{y+1}-\sigma_R^2/2} \qquad \varphi_y \sim N(0,\sigma_R^2) \tag{A.2b}$$

but not for the third:

$$N_{y+1,1} = \bar{R}_1\, e^{\varphi_{y+1}-\sigma_R^2/2} \qquad \varphi_y \sim N(0,\sigma_R^2) \tag{A.3}$$

where SB_y is the female spawning biomass during year *y* (corresponding to 1 April):

$$SB_y = \sum_{a=2}^{x} w_{y,a}^{spawn} \phi_{y,a} N_{y,a} e^{-frac(M_a+\tilde{S}_{y,a}F_y)} \tag{A.4}$$

$\phi_{y,a}$	is the fraction of fish of age *a* that are mature/spawning during year *y*;
frac	is the fraction of the year at which spawning takes place (set to 3/12);
R_1	is the number of age-1 animals at unfished equilibrium;
ψ_0	is spawning biomass-per-recruit in the absence of exploitation;
h	is the steepness of the stock-recruitment relationship ($h = 1/(1+4e^{-A_r})$ for the Ricker model);
$\bar{R}_1$	is average age-1 recruitment;
σ_R	is the log-scale standard deviation of the random fluctuations in recruitment about the underlying deterministic stock-recruitment relationship (set to 1.0; Dorn et al. 2005); and
$w_{y,a}^{spawn}$	is the average mass of a spawning fish of age *a* during year *y*.

Fishery selectivity

Following Dorn and Methot (1990), Helser et al. (2001), and Sullivan et al. (1997), Dorn et al. (2005) modeled fishery selectivity for the historical period using a double-logistic function with time-varying parameters:

$$S_{y,a} = \left[\frac{1}{1+\exp[-\beta_1 e^{\delta_{\beta_1,y}}(a-\alpha_1-\delta_{\alpha_1,y})]}\right]\left[1-\frac{1}{1+\exp[-\beta_2 e^{\delta_{\beta_2,y}}(a-\alpha_2-\delta_{\alpha_2,y})]}\right] \tag{A.5}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are the parameters that determine the shape of the selectivity curve;

$\delta_{\alpha_1,y}, \delta_{\alpha_2,y}$	are the deviations in the fishery selectivity α parameters for year *y*; and
$\delta_{\beta_1,y}, \delta_{\beta_2,y}$	are the deviations in the fishery selectivity β parameters for year *y*.

Catches

Under the assumption of continuous fishing throughout the year, the fully selected fishing mortality is calculated by solving the equation:

$$\hat{C}_y = \sum_{a=1}^{x} w_{y,a}\hat{C}_{y,a}\ ;\ \hat{C}_{y,a} = \frac{\tilde{S}_{y,a}F_y}{M_a + \tilde{S}_{y,a}F_y} N_{y,a}(1 - e^{-(M_a + \tilde{S}_{y,a}F_y)}) \qquad \text{(A.6)}$$

where $\hat{C}_y$ is the estimated catch (in mass) during year *y*;

$w_{y,a}$ is the average mass of a fish of age *a* during year *y* (Dorn et al. 1999); and

$\hat{C}_{y,a}$ is the estimated catch of fish of age *a* (in numbers) during year *y*.

Surveys

The data used to estimate the values for the parameters of the operating model are available from six fishery-independent sources: (1) the NMFS winter GOA Shelikof Strait echo-integration trawl (EIT) survey (Guttormsen and Yasenak 2006); (2) the NMFS GOA bottom trawl survey (Martin 1997, Britt and Martin 2000); (3) the egg production estimates of female spawning biomass (Picquelle and Megrey 1993); (4) the ADFG crab/groundfish nearshore bottom trawl survey (Blackburn and Pengilly 1994); (5) the McKelvey age 1 survey (McKelvey 1996); and (6) an historical ADFG 400-mesh eastern trawl survey (Dorn et al. 2005). The data for each survey include indices of abundance, and survey size- and age-composition (see Dorn et al. (2005) for full details of the data available for each survey type).

The model estimates of the survey biomass indices are calculated using the equation:

$$\hat{I}_{d,y} = \hat{q}_d \hat{B}^e_{d,y} \qquad \text{(A.7)}$$

where

$\hat{I}_d$ is the model estimate of the biomass index for survey *d*;

$\hat{q}_d$ is the model estimate of catchability for survey *d*;

$\hat{B}^e_{d,y}$ is the model estimate of the total biomass available to survey *d* during year *y*:

$$\hat{B}^e_{d,y} = \sum_{a=1}^{x} w^d_{y,a}\hat{N}^e_{d,y,a} = \sum_{a=1}^{x} w^d_{y,a}\tilde{S}^d_a N_{y,a} e^{-frac_d(M_a + \tilde{S}_{y,a}F_y)} \qquad \text{(A.8)}$$

$\hat{N}^e_{d,y,a}$ is the model estimate of total number of age *a* animals available to survey *d* during year *y*;

$\tilde{S}^d_a$ is the age-specific selectivity pattern for the survey type *d* (assumed to be descending logistic function for the EIT survey, an ascending logistic function for the ADFG coastal survey and the historical trawl survey, and a double logistic function for the NMFS bottom trawl survey; the egg production estimates are assumed to be indices of female spawning biomass; while the McKelvey data are indices of age 1 abundance);

$w^{d}_{y,a}$ is the average mass of a fish of age *a* in year *y* during survey *d* (Dorn et al. 1999); and

$frac_d$ is the fraction of the year at which survey *d* takes place.

Initial conditions

The numbers-at-age at the start of 1961 are assumed to be lognormally distributed about the equilibrium numbers-at-age.

$$N_{1961,a} = N_{1961,1} e^{-\sum_{a'=1}^{a-1} M_{a'}} e^{v_a} \quad ; \quad v_a \sim N(0, \sigma_R^2) \qquad \text{(A.9)}$$

Parameter estimation and Bayesian analysis

The estimable parameters of the operating model are the parameters of the stock-recruitment relationship, the deviations in recruitment (initial conditions and about the stock-recruitment relationship), the annual fishing mortalities, the parameters that define fishery selectivity and how it changes over time, and the parameters that determine survey catchability and survey selectivity. The sample of parameter vectors used for projection purposes is generated using the Markov chain Monte Carlo (MCMC) algorithm (AD Model Builder, Fournier 2006). The application of the MCMC algorithm is based on 11 million iterations, of which the results of every 10,000th iteration are saved. The objective function that forms the basis for the application of the MCMC algorithm follows that used by the actual assessment (Dorn et al. 2005) and includes components for fishery catch, fishery catch proportions-at-age, fishery catch proportions-at-length (for years for which fishery catch proportions-at-age data are not available), survey abundance indices, survey catch proportions-at-age, and survey catch proportions-at-length (for years for which survey catch proportions-at-age data are not available). This objective function also includes priors on the deviations about the deterministic stock-recruitment relationship (Eq. A.2 and A.3), on the deviations about the deterministic 1961 age-structure (Eq. A.9), and on the extent to which fishery selectivity may vary among years (Eq. A.5).

When recruitment is assumed to be related to spawning stock size according to the Beverton-Holt or Ricker relationships, a normal distribution with mean given by the average age 1 recruitment since 1977

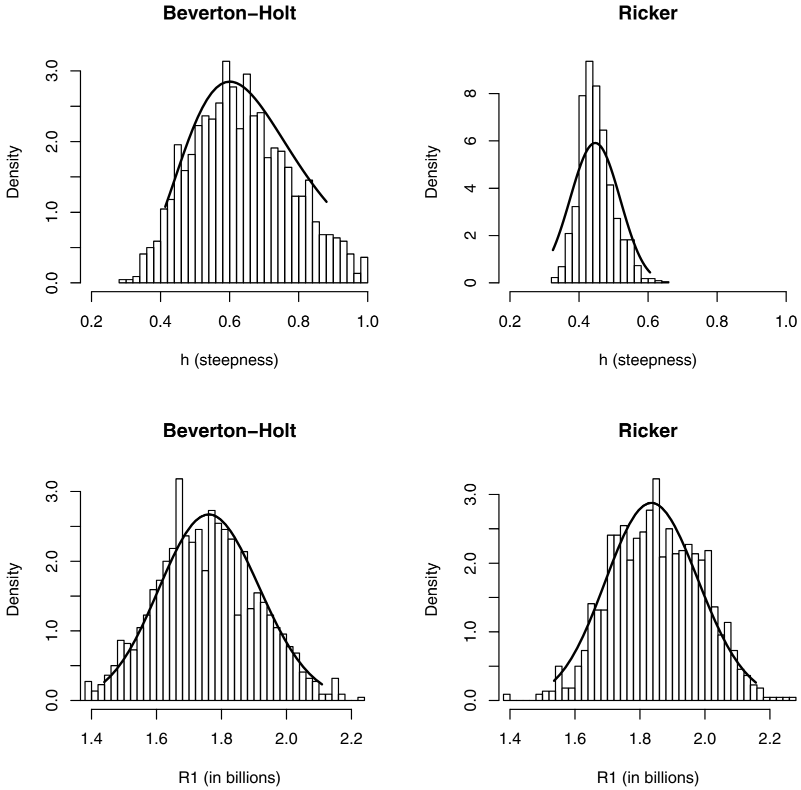

Figure A.1. Likelihood profiles (solid lines) and the marginal posteriors (histograms) for R_1 and *h* for the Beverton-Holt and Ricker scenarios.

from the scenario in which recruitment is independent of spawning stock size and a CV of 0.1, is used as a prior for the recruitment in the unfished state, R_1. Fig. A.1 contrasts the posterior distributions for R_1 and h (steepness) with likelihood profiles for these quantities.

Data generation

The data available to the estimation model in the future are the data currently available (and on which the parameters of the operating model are based) along with future data generated by the operating model. The latter data are generated in a manner consistent with how they are treated in the actual GOA pollock assessment. Data are generated for fishery catch proportions-at-age (using multinomial sampling, with an effective sample size of 400) and for the EIT, NMFS bottom trawl and ADFG coastal surveys. Future data are not generated for the egg production and historical trawl surveys because these surveys no longer take place, while the age-1 indices are sufficiently imprecise that they provide effectively no information and are consequently not simulated. The selectivity for the fishery and the surveys are constant over the projection period in both the operating model and the stock assessment model. In the operating model, the projected fishery selectivity is the average fishery selectivity for 1992-2004.

The simulated indices of abundance from the EIT and ADFG surveys are lognormally distributed with CVs of 0.2 and 0.25 respectively. The survey proportions-at-age for these two surveys are multinomially distributed with effective sample sizes of 60 and 10 respectively. The lognormal CV assumed for the NMFS bottom trawl surveys in each future year is selected at random from the historical past CVs for this survey. Similarly, the annual effective sample sizes for the proportions-at-age for this survey are selected at random from the historical effective sample sizes. The EIT and ADFG surveys are assumed to be conducted annually in the future while, consistent with actual practice, data from the NMFS bottom trawl survey are only generated for every second year in the future.

The annual weights-at-age for the fishery, all surveys, and the average population weights and weights at spawning are based on the data used in the 2005 GOA pollock stock assessment.