REPORT 3.1

Analysis of the Sea of Okhotsk Pollock stock assessment model and its effectiveness in addressing all major sources of uncertainty

Prepared by Alexei Sharov
Baltimore Maryland USA

Report to

Pollock Catchers Association

690091 Russia, Primorsky Krai, Vladivostock, Svetlanskaya Street 51A

April 30 2016

Summary

This report was produced for the Pollock Catchers Association (PCA) to address MSC certification condition, requiring an evaluation of the stock assessment model of the Sea of Okhotsk Pollock with respect to its effectiveness in addressing all major sources of uncertainty. This review is based on two reports prepared by KamchatNIRO (Varkentin and Ilyin, 2014; 2015) and provided by PCA and the original MSC certification document (Intertek Moody Marine, 2013).

The “Synthesis” model used for the assessment of the North Sea of Okhotsk Pollock is a version of the statistical catch at age class of models which are employed extensively throughout the world. The “Synthesis” model accounts for measurement errors in catch, fishing effort and indices of abundance, as well as uncertainty in the stock recruitment relationship. Weight coefficients of the objective function components are adjusted inversely proportional to the variance. Thus the contribution of various sources of measurements error is adjusted according to their reliability. Both measurement and process errors are further accounted for when the bootstrap method is used to generate a probability distribution for population parameters in order to quantify the uncertainty and evaluate the risk of exceeding the target or threshold reference points.

Uncertainties in model parameters and natural variability are also accounted for in the harvest control rule. Bootstrap generated estimates of error in model parameters are used as input for the Monte Carlo random re-sampling simulation to generate distributions of population abundance, spawning stock biomass and fishing mortality for a selected level of Total Acceptable Catch (TAC) and estimate the probability of exceeding the target or limit reference points for biomass and fishing mortality.

Overall, the Northern Sea of Okhotsk Pollock stock assessment does characterize major sources of uncertainty, such as uncertainty caused by measurement errors in input data, uncertainty in the model approximation of population dynamics, and uncertainty in the natural variability of the Northern Sea of Okhotsk ecosystem. The methods used for uncertainty characterization are similar to those used by stock assessment teams throughout the world. However, not all available information is fully utilized in the uncertainty characterization. In particular, information on age composition is not explicitly included in the objective function. Because the age composition is currently not included in the objective function, the uncertainty level is likely to be underestimated. However, the model is flexible and can easily accommodate the addition of age composition data. Accounting for age composition could be easily achieved by modeling catch at age composition assuming a multinomial distribution, developing age specific indices the indices of abundance and including the age composition deviations into the objective function. Other changes to model structure or additional analyses that could improve uncertainty characterization in the model include incorporation of ageing error, different weighting schemes for the objective function, model sensitivity analysis, risk analysis being applied relative to the target rather than threshold reference point, consideration of uncertainty in reference points estimates.

Introduction

This report was commissioned by the PCA with an objective to provide an analysis of the assessment model of the Sea of Okhotsk Pollock stock and its effectiveness in addressing all major sources of uncertainty. According to the terms of reference, the scope of work was defined as follows:

- To evaluate the efficiency of the current harvest strategy of the Sea of Okhotsk Pollock (based on KamchatNIRO materials);
- To evaluate the consistency in uncertainty consideration in stock assessment modeling and TAC forecast for Pollock in the northern part of the Sea of Okhotsk;
- To evaluate the consistency of the methods used for uncertainty consideration in stock assessment and TAC forecast of Pollock in the northern part of the Sea of Okhotsk.

Background

The Russian Sea of Okhotsk Midwater Trawl Walleye Pollock Fishery was originally certified by MSC on September 24, 2013 (Intertek, 2013). The certification report included several conditions, including the following:

By the third surveillance audit, provide a report which details how the assessment appropriately evaluates major sources of uncertainty and takes them into account.

With respect to this condition, Client Action Plan included the following action:

By the second surveillance audit PCA will commission a written review provided by fishery science institutions (both from those within and external to the SOO fishery) of the assessment model and its effectiveness in addressing all major sources of uncertainty. PCA will require the review to consider whether there are alternative assessment models which should be evaluated. Any recommendations from the review will be considered for incorporation into the assessment process where appropriate. By the third annual surveillance audit, the client must provide the completed written review detailing how the assessment appropriately evaluates major sources of uncertainty and takes them into account. This report addresses the aforementioned request.

Stock assessment model.

Prior to 2009, the status of the North Sea of Okhotsk Pollock was assessed using both direct enumeration (a combination of trawl, acoustic and egg surveys) and analytical models (VPA, ISVPA) (Intertek Moody Marine, 2013). In 2009 an age structured forward-projecting statistical catch at age model was introduced for stock assessment of the North Sea of Okhotsk Pollock. With some modifications, this model was used as the principal assessment tool during a seven year period (2009-2014), thus ensuring consistent assessment methodology (Varkentin and Ilyin, 2014; 2015). The model is a variation of a widely used class of statistical age structured models initially introduced by Fournier and Archibald (1982) and Deriso et al. (1985) in their CAGEAN model and by Methot (1989) in his Stock Synthesis model. Forward-projecting age-structured models share many attributes with older tuned and untuned VPAs, but provide more flexibility and are widely used throughout the world (Patterson, 1994; Quinn and Deriso, 1999, Legault, et. al., 2001, SEDAR 2015). The North Sea of Okhotsk Pollock assessment model was custom programmed by the principal assessment scientist (Ilyin et al., 2013) and termed as the "Synthesis Model" (not to be confused with "Stock Synthesis" by Methot, 1989; 2009). The model and its implementation was reviewed internally by regional (TINRO) and central (VNIRO) Russian fisheries research institutes (Intertek Moody Marine, 2013). However, no external international review has been completed so far. This report does not constitute a full formal review of the assessment model as it's based on a limited description provided in two recent reports by KamchatnNIRO (Varkentin and Ilyin, 2014; 2015). Nonetheless, the information made available appears to be sufficient to describe major model structure and basic results.

Model structure

The "Synthesis" age structured model is based on the classic concept of exponential decline of cohort abundance through time and Baranov's catch equation (Baranov, 1918). Mathematical details are described in section 1.1.1.4 in Varkentin and Ilyin (2014, 2015). The principal concept of an age structured forward-projecting statistical model is to simulate a population that is projected forward in time by describing changes in abundance due to natural and fishing mortality, growth, recruitment to the stock, maturation, and spawning stock dynamics. Aspects of the fishing process, such as fishing gear selectivity, are also simulated. Model parameters to be estimated are systematically varied from starting values until the simulated population's characteristics match observed data on the real population as closely as possible. Such data include total catch by year, observed catch at age in numbers of fish by year, and observed indices of abundance or biomass from fishery dependent or fishery independent sources.

Data Inputs and Parameterization

The major characteristics of the model formulation were as follows:

Assessment dimensions

The assessment time period was 1963–2014. Alternative start year configurations were not explored or, at least, not reported. The population was assumed to represent a single stock, closed to immigration and emigration. The model included age classes 2-20, no plus group. Fishery removals are assumed to be completed by a single fleet.

Removals in numbers of fish by age

Removals included landings and dead discards. Annual landings, in weight, for three fishing fleets (midwater trawl, Danish seines and bycatch in other fisheries) were converted to numbers at age using annual age length keys and length frequency distributions and combined into a single catch at age matrix, by year.

Maturity

The percent of females mature was age and time varying. Multi-year mean values were used in place of missing data. Maturity was based on samples taken during the peak spawning period (January – February) at the start of the fishing year.

Natural mortality

Natural mortality (M) is age specific but time invariant. Age specific estimates of M are estimated through a two stage process. A long-term average annual M is calculated using the Gunderson and Dygert (1988) method based on a gonadosomatic index. This is assumed to represent the M on the most abundant ages (6 – 8) in the catch. The method of Blinov (1977) is then used to scale M for other age groups.

Sex ratio

The ratio of males to females was fixed at 1:1.

Fecundity

No fecundity data were used in the model. Pollock reproductive output was represented in terms of spawning stock biomass expressed in total weight of mature females.

Weights at age

The weights at age during spawning were year specific. Multi-year mean values were used in place of missing data.

Recruitment

Spawning was assumed to occur at the beginning of the year. Recruitment to age two was estimated in the assessment model for each year using either Ricker or Beverton - Holt stock-recruitment curves. No indication was given, though, of which stock recruitment model was used in the model run described in Varkentin and Ilyin (2014; 2015).

Fishing fleets

There are two principal fleets targeting Pollock in the Sea of Okhotsk – a fleet of large trawlers, utilizing midwater trawls, and a smaller, nearshore fleet using Danish seines. Bycatch in fisheries targeting other species (i.e. herring) is also accounted for. The assessment, however, does not model each fleet explicitly. Instead, the catch information from all data sources, including discards, is combined into a single catch at age matrix. Fishing mortality rates for fully selected age classes were estimated for each year by solving the catch equation.

Selectivity functions - fishery

The selectivity concept was used in the estimation of age and year specific fishing mortality rates where the fishing mortality rate F_{ij}, of age group i in year j is equal to the product of age specific selectivity rates S_{ij} and year specific full fishing mortality $F_{j.}$, $F_{ij} = S_{ij} * F_{j.}$, (Pope, Shepherd, 1982). Age specific selectivity was freely estimated in the model and no functional form was assumed. Two time period selectivity blocks were defined based on fishery regulations. It appears that for both time blocks, selectivity at age was dome shaped (Figure 1.1.4.5. in Varkentin and Ilyin, 2015).

Selectivity functions – indices

No selectivities for indices were considered because all indices were modeled as age aggregate.

Discards

Discards were not modeled separately. Discard estimates, in weight, are converted into numbers at age based on observed size and age samples and added to the total catch at age matrix (Intertek Moody Marine, 2013).

Abundance indices

No abundance indices (numbers at age) were used in the model although data are available to derive such quantities from at least two fishery independent surveys. "Abundance indices" are used in Varkentin and Ilyin (2015) interchangeably with biomass s, but it appears that only biomass indices were used in the assessment.

Biomass indices

Several biomass indices were used in the model. Two ichthyoplankton surveys (KamchatNIRO's ichthyoplankton survey of the West Kamchatka shelf from 1972–2013 and TINRO-Center's ichthyoplankton survey of Pollock spawning stock in the northern part of the Sea of Okhotsk from 1998–2014) were used to develop indices of spawning stock biomass. TINRO-Center's acoustic survey in the northern part of the Sea of Okhotsk in 2001–2014 was used to produce an index of total Pollock biomass. In addition, two trawl surveys (TINRO-Center's trawl survey of Pollock biomass in the northern part of the Sea of Okhotsk from 1998–2014 and TINRO-Center's autumnal trawl survey of total Pollock biomass from 1995–2007) were used to generate total biomass indices. CPUE time series for large tonnage vessels were also considered to be an index of total population biomass.

Ageing uncertainty

Ageing uncertainty was not included in the assessment due to an absence of information on the true age of sampled fish. No age reading error measures were reported to be used in the assessment.

Fitting criterion

Estimates of population parameters are found on the basis of the best approximation of the model to observed data in accordance with error distribution assumptions. Information on catch, effort, indices of abundance or biomass, and the stock recruitment relationship is assumed to be observed with a random error that has a lognormal distribution.

The least squares objective function is a weighted sum of squared deviations:

$$Z = Z_C + \lambda_E Z_E + \lambda_R Z_R + \sum_I \lambda_I Z_I$$

where

$$Z_C = \sum_i (\ln(\hat{C}_{i,j}) - \ln(C_{i,j}))^2$$

is the sum of squared deviations of the catch of age *i* in year *j,*

$$Z_E = \sum_j (\ln(\hat{E}_j) - \ln(CE_j))^2$$

is the sum of squared deviations of the observed and predicted effort

$$Z_E = \sum_j (\ln(\hat{R}_j) - \ln(CR_j))^2$$

is the sum of squared deviations of the observed and predicted recruitment,

$$Z_I = \sum_j (\ln(\hat{I}_j) - \ln(I_j))^2$$

is the sum of squared deviations of the observed and predicted index of relative abundance or biomass, and λ*i* which are weight coefficients calculated as inversely proportional to the variance of the observation errors. Weights were standardized by assuming λ=*1* for deviations in catch data and calculating λi as a ratio of variances.

A notable difference of the “Synthesis” model from many other statistical catch at age applications is that it does not include age composition of the catch and indices of abundance in the objective function. This is probably the most significant drawback because such important and easily available information is not utilized in model fitting.

The least squares objective function is optimized by one of three numerical optimization methods: the Levenberg–Marquardt method, the steepest descent method or the Fletcher–Reeves conjugate graduate method (Bazaraa and Shetty, 1982). The simplified genetic algorithm (Holland, 1975) is applied to obtain an initial approximation for the vector of estimated parameters. The following model parameters are estimated:

initial numbers at age for the start year, recruitment for each year of simulation, fishing mortality rates of fully selected age groups for each year, selectivity parameters at age by time blocks, parameters of “stock – recruitment” relationship, and catchability coefficients for each survey.

Model testing

The “Synthesis” model has undergone some internal review within the KamchatNIRO and VNIRO research institutes (Intertek Moody Marine, 2013) but it is not known how rigorous these reviews were as no records were available to us. The general model structure is similar to the CAGEAN model (Deriso et al., 1985) that has been extensively peer reviewed. The model code and input data were not available to us for a review but would have been very useful as an additional measure of quality control. A combination of testing and verification procedures could serve as proof that the assessment model has been implemented correctly and provides an accurate assessment of Pollock in the northern part of the Sea of Okhotsk dynamics.

Model Diagnostics

Goodness of Fit

Goodness-of-fit is determined in the “Synthesis” assessment model by the residual sum of squares components in the objective function. The relative contribution among the objective function components is adjusted through the weighting terms and according to the error level for each data source. Goodness-of-fit is also judged for each data source through examination of the model residuals.

Patterns in the annual comparisons of the bubble plots of deviations between observed and predicted catch-at-age (Figure 1.1.4.10 in Varfkentin an Ilyin, 2015) indicate a good overall model fit to the observed data. There is no patterning observed in the bubble plot that caused concern. No year or year class effects were observed in the form of systematic positive or negative bias.

The observed and predicted fishing effort values fit extremely well, possibly due to a high weighting (λ) value in the objective function (lambda values were not shown in the report).

Observed and predicted biomass indices were compared for the base model run (Figures. 1.1.4.11– 19). Visual examination of the fit suggests that the overall pattern fit reasonably well for most of the data points. The general patterns are captured. For all indices, logarithmic deviations of observed values from predicted values were reported to be distributed satisfactorily which is confirmed by the Kolmogorov–Smirnov test ($p>0.05$).

Estimating Precision

The “Synthesis” model allows calculation of the inverse Hessian matrix which provides approximate precision of estimated model parameters (Table 1.1.4.1 in Varkentin and Ilyin, 2015). However, in this case where some key values were fixed (e.g., natural mortality), it is believed that the precision measures from the inverse Hessian matrix are likely to be underestimates of the true precision.

Sensitivity Analyses

Sensitivity runs are a very useful tool to explore the effects of uncertainty in input data, changes to the model configuration, and historic stability of model results as part of the retrospective analyses. Sensitivity of the model was investigated only in terms of retrospective analysis.

Retrospective Analyses

Retrospective analyses were completed by running the “Synthesis” model in a series of runs sequentially omitting all data, year by year, from 2014 to 2011. Retrospective analyses are meant to demonstrate the behavior of the model to additional years of data. If additional years of data are outside the range of data observed in the past, sometimes patterns may exist in the retrospective runs, resulting in bias – underestimation or overestimation in fishing mortality, abundance or biomass. Retrospective analysis results for the North Sea of Okhotsk Pollock included only one plot for the spawning biomass (Figure 1.1.4.20 in Varkentin and Ilyin, 2015). It appears that there is either very little or no bias in SSB estimates but this conclusion is made based on the visual inspection of the graph as no numerical estimates were reported. It would be helpful to have separate plots of retrospective runs for total abundance, SSB, fishing mortality and recruitment on absolute and relative scale. Despite its limited scope in presentation, retrospective analysis suggests stability of model estimates, at least with respect to the spawning biomass.

Biological Benchmarks (Reference Points)

Current interim benchmarks adopted for Pollock in the northern part of the Sea of Okhotsk include F_{msy} as F target and $F_{35\%}$ as a threshold. B_{msy} is currently used as a target reference point for spawning biomass and B_{loss}, adjusted for uncertainty, is adopted as a threshold. No uncertainty in reference point estimates is currently considered

Projections

Projections were run for either a two year period to calculate the TAC or for a ten year period to evaluate long-term stock sustainability under the selected harvest strategy. The projection methodology is based on the same mathematical concept as the assessment model where the numbers, by age group, are projected forward. The starting conditions of the projection analysis include numbers at age, N_a, estimated in the terminal year. Numbers at age after the initial year were calculated as:

$$N_{a+1,y+1} = N_{a,y} e^{-M_a - F_t S_a}$$

where $\boldsymbol{M_a}$ was age and year specific natural mortality, $\boldsymbol{F_t}$ is the target fishing mortality rate, and $\boldsymbol{S_a}$ is the selectivity at age $\boldsymbol{a}$.

To account for the uncertainty in the input data, an output from the nonparametric bootstrap runs of the "Synthesis" model is used. A Monte Carlo (MC) random resampling method is applied in each model run and each year of the projection, with a random error being added to the mean value of age specific weight, percentage of mature individuals, and selectivity. Recruitment was projected with an underlying stock-recruitment function with additional error estimated through the bootstrap procedure during the base run. A large number of runs are completed to generate possible outcomes. Although no specifics were reported in the documentation as to how many MC runs were completed, it is expected to be at least in the hundreds. Based on the MC runs, means and standard errors are calculated for variables of interest, such as total biomass, spawning biomass, and fishing mortality and their statistical distributions are generated for each year of the projection.

Uncertainty characterization in the "Synthesis" assessment model

There are four major sources of uncertainty recognized in stock assessments: (1) measurement errors, (2) model parameters, (3) model structure and (4) physical and biological processes affecting the stocks (Hilborn, 1997). Uncertainty caused by measurement errors is associated with unrepresentative samples of biological data (length, weight, maturity, etc.) or inaccurate catch. Thus, the reliability of the assessment is greatly dependent on the reliability of the data. Model parameters are estimated with some error due to imperfect data model structure, while the model structure is only a simplified mathematical concept of population dynamics that may not fully or precisely describe underlying processes. Furthermore, physical and biological processes, such as multidecadal dynamics in the climate, marine environment or ecosystem affect population productivity, which also translates into errors in estimated population parameters, particularly when they are assumed to be constant through time.

The most common methods for computing uncertainty in the assessment include bootstrapping (Restrepo et al., 1992), maximum likelihood (Pollacheck at al, 1993), and Bayesian statistical analysis (Punt and Hilborn, 1997; McAllister and Kirkwood, 1998). Another additional popular tool used in exploration of uncertainty in model results is a sensitivity analysis, which constitutes an investigation of the effect of the parameter of interest on model outcome. This is done by re-running the model multiple times with the range of selected parameter values ("what if" scenarios).

The uncertainty in the North Sea of Okhotsk Pollock stock assessment is characterized in several steps.

First, at the model structure level, the assumption is made that for some variables or model parameters, observed data deviate from predicted ones due to random error with a lognormal distribution. In the "Synthesis" model, the lognormal error distribution is assumed for total catch in weight, the indices of abundance, fishing effort and two stock recruitment parameters (see section on Model Structure above and section 1.1.4. in Varkentin and Ilyin, 2015). This is a standard assumption made when statistical catch at age models are used. The "Synthesis" model differs from similar applications only by the fact that some models assume a lognormal distribution for a larger number of parameters. Therefore, observation errors are reflected in standard deviations of the catch, effort, recruitment and indices of abundance. The weighted sum of squared deviations of these quantities constitutes the objective function. When the objective function is minimized, model parameters are selected that minimize uncertainty of the first (observation error) and the second (process error) type. Assigning weight coefficients to the objective function components being inversely proportional to the variance leads to the components with higher variance being assigned a lower weight and vice versa. Thus, the level of uncertainty in the data sources is accounted for and quantitatively adjusted when optimizing the objective function. Based on the calculation of the inverse Hessian matrix, the model provides standard error estimates for model parameters, this being the first step of quantification of model uncertainty.

To further quantify the uncertainty in model results, the “Synthesis” model uses a Monte Carlo and nonparametric bootstrap methods. Monte Carlo and bootstrap methods (Efron and Tibshirani 1993; Manly, 1997) are often used to characterize uncertainty in ecological studies, and the mixed approach, Monte Carlo - Bootstrap (MCB) has been applied successfully in stock assessment throughout the world (Restrepo et al.,1992; Legault et al., 2001 Magnusson et al., 2013;. Gjøsæter et al.,2002). The approach is among those recommended for use as standard methodology to characterize the uncertainty in assessment models in some regions of the USA (SEDAR Procedural Guidance 2010).

The nonparametric bootstrap approach translates uncertainty in model input into uncertainty in model output by re-sampling the deviations (residuals) of observed from modeled data and adding them to the original data, thus creating a new “observed” data set. The model is fit many times with different values of “observed” data and key input parameters. For each model run, one set of parameter estimates is obtained and once a large number of model runs are completed, a distribution of the values is generated for each estimated parameter. A chief advantage of this approach is that the results describe a range of possible outcomes so that uncertainty is characterized more thoroughly than it could be by any single fit or handful of sensitivity runs. As mentioned earlier, this is one of the most frequent, “standard” approaches widely used by assessment scientists throughout the world (Magnusson et al., 2013). The resulting estimates from the MCB runs usually are graphically summarized, showing the selected % confidence region (i.e.95%). The MCB analysis should be interpreted as providing an approximation to the uncertainty associated with each output. The results are approximate for two related reasons. First, not all combinations of Monte Carlo parameter inputs are equally likely as biological parameters might be correlated. Second, all runs are given equal weight in the results yet some might provide better fits to the data than others.

A standard outcome of such an application is a number of time series plots of fishing mortality, exploitable and spawning stock biomass or recruitment with a selected level of confidence intervals (Figures 1.1.4.6 – 1.1.4.9 in Varkentin and Ilyin, 2015) or probability density plots for the terminal year of the assessment (not shown in the report). By comparing these probability distributions to corresponding reference points, a probability of exceeding a target or threshold fishing mortality or spawning stock biomass can be calculated. This is the essence of the third step of uncertainty evaluation – accounting for uncertainty in the harvest control rule (management strategy evaluation).

To evaluate the effectiveness of the current harvest strategy, uncertainty in the North Sea of Okhotsk Pollock stock assessment is presented in the form of probability of potential changes in stock condition relative to the biological reference points. More specifically, a probability that the spawning stock will fall below the biomass threshold is calculated assuming that total removals will be equivalent to the TAC recommended using the harvest control rule (HCR). Probabilistic interpretation of uncertainty in the scientific recommendations on the TAC is presented in the form of risk. Risk analysis is based on the implementation of the Monte-Carlo method in the forecasting model for a period of two years for the evaluation of TAC or long-term projections (10 – 50 years) to evaluate long term sustainability given the elected management strategy.

In the case of the TAC evaluation, population abundance and biomass is projected forward multiple times (>>100), assuming that total removals will be equivalent to the TAC. Each time the projection is run, model parameters are drawn randomly from distributions derived by the bootstrap method during previous "Synthesis" model runs or some preset distributions based on external evaluation. Random errors are generated for starting values of yearly class numbers, weight, percentage of mature individuals, selectivity factors and recruitment values in forecast years. Based on projection results, a cumulative probability distribution for spawning stock biomass and fishing mortality is generated (Figure 1.1.6.3 in Varkentin and Ilyin, 2015). The conclusion on the appropriateness of the recommended TAC is made based on maximum allowable risk, which is pre-determined for the event under consideration. Specifically, the current management strategy specifies that the probability of fishing mortality to exceed the threshold or spawning biomass to fall below the biomass limit should not exceed the recommended level of 10% (Babayan, 2000). This appears to be a strong, risk averse requirement, ensuring that under the recommended level of removals there is a very low risk of recruitment overfishing due to low spawning stock biomass. It should be noted, however, that this risk assessment is built around the threshold reference point B_{lim}, while more frequently risk is evaluated in relation to B_{target} (which in this case is equivalent to B_{msy}), thus focusing more on maintaining the population near the target spawning biomass, rather than avoiding spawning stock collapse. Similarly, given the selected level of removals, an evaluation of the probability of the spawning stock falling below the biomass limit is evaluated based on the long-term projection using the same methodology

(Figure 1.1.6.2.). Thus, all three major types of uncertainties are taken into account when evaluating the risk of the current management strategy (HCR) application.

It has to be noted however, that sensitivity analysis, a rather useful method of uncertainty exploration, has not been used much in the assessment, with the exception of retrospective analysis. Sensitivity analysis is an analysis that refits the assessment model with different values for the assumed parameters to examine how much, if any, the outputs change. It also allows for testing of alternative hypotheses. A thorough sensitivity analysis may include changing the following parameters or assumptions: natural mortality, the stock – recruitment relationship, selectivity, environmental factors, or using alternative indices. Useful outputs to compare for such sensitivity analyses include biomass (total and spawning), fishing mortality, recruitment, selectivity curves, reference points, and stock status. Sensitivity analysis can be very useful, particularly if it shows that the model is not sensitive to the uncertainty in a particular parameter, improving confidence in the model results. Alternatively, if the model is sensitive to a certain parameter, additional work can be done to define its possible range and distribution type, thus narrowing the range of possible outcomes.

Consistency of methods and levels of uncertainty taken into consideration in stock assessment

During the original certification process, it was judged that the "Synthesis" model recognized the major sources of uncertainty as being present in the catch at age, the stock – recruitment relationship, and the fishing mortality via the fishing mortality – effort and catch rate – fishable biomass relationships (Intertek Moody Marine, 2013). However, relative weight of these uncertainties incorporated into the model as terms in the objective function was determined arbitrarily by expert judgment. While other sources of uncertainty were recognized (e.g. M and ageing) they were not explicitly modeled. This lead the assessment team to the conclusion that "*while the assessment identifies major sources of uncertainty, it is only taking these into account to a limited degree*" (Intertek, 2013)). At that time, the Monte Carlo method was used in the characterization of uncertainty in time series estimates of abundance and biomass, but it appeared to understate the uncertainty due to the use of a predetermined error level prescribed to catch information.

Since 2012, following the assessment experts' advice, the KamchatNIRO scientists adopted a nonparametric bootstrap approach, with sampling of residuals associated with the catch at age and stock abundance indices (e.g. CPUE and survey indices), to more fully characterize uncertainty in the model. Assigning weight coefficients to the objective function components being inversely proportional to the variance was another important modification at the same time. Both changes seemed to be an improvement to the assessment model. The bootstrap based estimates for model parameter variance are further used in the Monte Carlo analysis to characterize uncertainty in projections of the future stock size and determine the risk following the application of the existing harvest control rule. The assessment methodology has not been modified since that time, therefore a consistency in methodology has been present in the course of at least four years (since 2012). A consistency in methods accounting for the uncertainty translates into consistency in the level of uncertainty taken into account, meaning that consistent application of the same method allows to capture the same sources of uncertainty within the limits of the applied methodology.

Evaluation of effectiveness of the current harvest strategy of the Sea of Okhotsk Pollock

For the purpose of this review, we interpret effectiveness as the success of the current management strategy in achieving its management goals, which are generally defined as stock preservation and rational use (Intertek Moody Marine, 2013). Specifically, current management strategy relies on the harvest control rule (HCR) adopted in 2010 which is designed to maintain the fishing mortality and spawning stock biomass at their corresponding target levels. This is achieved by setting the TAC according to the HCR which includes accounting for the uncertainty in the stock assessment and risk of exceeding the limit reference points. A plot of the HCR diagram with the annual point estimates of spawning stock biomass and full fishing mortality, based on the most recent assessment, indicates that the spawning stock biomass has been above SSB_{target} from 2008 through 2014 (Figure 1.1.6.4. in Varkentin and Ilyin, 2015), while fishing mortality was at or below the target in six out eight years and slightly above the F_{target} but well below the F_{lim} in two years (2009 and 2010). This suggests that recommended harvest level resulted in at least 50% or higher probability that SSB was above the SSB_{msy} and fishing mortality was below F_{msy}. Thus, setting a TAC based on the status of the stock and exploitation level according to the HCR and applying a low risk principle (less then 10 % chance of exceeding limit reference point) appears to be a successful harvest strategy. Since the inception of the current HCR, total removals were appropriately controlled and the population continued to fluctuate near SSB_{target}. It has to be noted that the SSB has never dropped below the SSB target since the inception of the current HCR and, therefore, further testing of the current HCR efficiency may occur when and if the SSB drops below the target, which would require a reduction in fishing mortality. Such an event may happen, for example, if a series of low recruitment events occur in the future which will lead to a decline in SSB. However, based on the five year period of the HCR use, the current strategy appears to be working within the observed range of recruitment fluctuation. The harvest strategy is responsive to the state of the stock and there is clear evidence that shows that it is achieving its objective to maintain the stock at target level.

Conclusions

The “Synthesis” model used for the assessment of the North Sea of Okhotsk Pollock is a version of the statistical catch at age class of models which are used extensively throughout the world. Thus, the foundation of the assessment methodology is solid. The “Synthesis” model accounts for measurement errors in catch data, fishing effort and indices of abundance, as well as uncertainty in the stock recruitment relationship when solving the objective function. Minimization of the objective function also reduces the process error, provided that the selected model appropriately describes the population dynamics. Weight coefficients of the objective function components are adjusted inversely proportional to the precision of the data so the contribution of various sources of measurements error is accounted for in the model according to their reliability. Both sources of uncertainty are further accounted for when the bootstrap method is used to generate a probability distribution for every important population parameter in order to quantify the uncertainty and evaluate the risk of exceeding the target or threshold reference points.

Uncertainties in model parameters and uncertainty in natural variability are accounted for in evaluating the efficiency of the harvest control rule. Model estimated errors in model parameters are used in the MCMC re-sampling method where a large number of randomly generated values of numbers and weight at age, proportion of mature females, selectivity at age and recruitment values are generated for the forecasted years.

Overall, the Northern Sea of Okhotsk Pollock stock assessment does characterize major sources of uncertainty, such as uncertainty caused by measurement errors, uncertainty in the model approximation of population dynamics, and uncertainty in the natural variability of the Northern Sea of Okhotsk ecosystem. The methods used for uncertainty characterization are similar to those used by stock assessment teams at national and international levels throughout the world. However, not all available information is currently used in the uncertainty characterization, although it could be used and fully integrated into the model. In particular, information on age composition is currently not included in the objective function which is probably the only important missing component. Accounting for age composition could be easily achieved by disaggregating the indices of abundance into age specific indices and including the deviations of observed from predicted age proportions into the objective function. This will lead to a more complete capture of age composition information in the assessment. “Synthesis” is an age structured model that tracks individual cohorts. Utilizing “true signals” on year class strength from the survey indices may significantly improve the assessment. Because the age composition is currently not included in the objective function, the uncertainty level is likely to be underestimated. However, the model is flexible and can easily accommodate the addition of age composition data. Other sources of uncertainty that could improve uncertainty characterization in the model are listed in the recommendation section.

Recommendations for further model modifications to improve the handling of model uncertainty

Disaggregate fishery independent survey indices of abundance into age specific. Model age structure assuming a multinomial distribution for catch at age, discard at age, index proportion at age, and include age composition in the objective function. This is the most immediate change that can cover the most significant gap in the estimation of uncertainty.

Consider the introduction of ageing error in the model. Ageing error matrices (AEM) are often used to account for ageing uncertainty during fisheries stock assessment and define age misclassification rates for the observed age range from sub-sampled harvest (Richards et al. 1992; Punt et al. 2008). Alternatively, a fixed level of ageing error (i.e. a constant CV estimated from an age verification study) can be set.

Consider different weighting of the objective function components. This provides important information about the robustness of the assessment. Not all stock assessments are sensitive to changes in data weighting but we can’t know about any such sensitivity unless we investigate alternative weightings (Francis, R. C., 2011).

Conduct sensitivity analyses to investigate model sensitivity to a variety of assumptions, including but not limited to: natural mortality, growth, maturity, the stock – recruitment relationship, selectivity, environmental factors, and alternative indices.

Consider modeling fishing fleets separately. Each fleet may have unique selectivity vectors so modeling them separately may lead to a reduction in uncertainty. Discards could be modeled separately as well.

Currently risk evaluation is considering only the probability of falling below biomass threshold B_{lim}. Including of the probability of achieving SSB_{target} will strengthen further current precautionary approach and correspond to fuller extent to the MSC requirement on appropriate handling of uncertainty in harvest control rule.

Consider accounting for the uncertainty in reference points. Examples of such an approach include Prager et. al. (2003), Grabowsky and Chen (2004), and Shertzer et. al. (2008).

A more substantial recommendation that will require a significant effort, but may also provide a very substantial benefit, is to recode the model in the AD Model Builder environment (Fournier et al., 2012) or adopt the existing, tried and tested versions of Statistical Catch at Age Model such as ASAP (Legault 2008), BAM (Williams and Shertzer, 2015) or namesake "Stock Synthesis" model by Methot (2013). ADMB is the most widely used software package for the development of state-of-the-art fisheries stock assessment methods. It can be downloaded without charge from a public web site, http://admb-project.org. Published benchmark assessments have shown that it provides faster and more reliable parameter estimation than other generic function minimizers. The advantage of such a move is twofold. First, this software provides a very powerful modeling environment and second, it has become a standard environment for assessment scientists throughout the world, making any future peer review of model structure and performance easy and transparent.

Moving the model into the Stock Synthesis framework (Methot, 2013) will even further advance the stock assessment of Pollock due to extreme flexibility of the model and its ability to accommodate all sources of data (e.g. size frequency information from the catch and the surveys that current model does not utilize). The Stock Synthesis model also uses maximum likelihood estimation and incorporates options for Bayesian analysis which would allow more complete analysis of model uncertainty.

Literature cited

Babayan V.K. 2000. Precautionary approach to estimation of total allowable catch (TAC). Analysis and recommendations on application. M.: VNIRO Publishing House. 190 p.

Baranov F.I. 1918. On biological fundamentals of fisheries. Proceedings of fish farming, scientific and industrial research department. Vol. 1, Issue 1. P. 84–128.

Bazaraa M., Shetty C. 1982. Non-linear programming. Theory and algorithms. M.: Mir Publishers. 583 p.

Blinov V.V. 1977. Modeling of natural mortality of younger age fish groups. Voprosy Ichthyologii. Vol. 17. Issue 3. P. 572–578.

Deriso, R. B., Quinn Ii, T. J., & Neal, P. R. (1985). Catch-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences*,*42*(4), 815-824.

Efron B, T. R. (1993). An introduction to the bootstrap. *New York, NY: Champan & Hall*.

Fournier D.A., C.P. Archibald. 1982. A general theory for analyzing catch-at-age data. Can. J. Fish. Aquat. Sci. 39: 1195–1207.

Fournier, D.A., H.J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M.N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27:233-249.

Francis, R. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(6), 1124-1138.

Grabowski, R., & Chen, Y. (2004). Incorporating uncertainty into the estimation of the biological reference points F 0.1 and F max for the Maine green sea urchin (Strongylocentrotus droebachiensis) fishery. *Fisheries research*, *68*(1), 367-371.

Gunderson D.R., P. H. Dygert. 1988. Reproductive effort as a predictor of natural mortality rate. J. Cons. Int. Explor. Mer. 44: 200-209.

Gjøsæter, H., Bogstad, B., & Tjelmeland, S. (2002). Assessment methodology for Barents Sea capelin, Mallotus villosus (Müller). *ICES Journal of Marine Science: Journal du Conseil*, *59*(5), 1086-1095.

Hilborn, R. (1997). Uncertainty, risk and the precautionary principle. In*American Fisheries Society Symposium* (Vol. 20, pp. 100-106).

Holland J.H. 1975. Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor.

Ilyin O.I., Sergeyeva N.P., Varkentin A.I. 2013. Stock assessment and TAC planning for East Kamchatka pollock (*Theragra chalcogramma*) based on precautionary approach // Collection of VNIRO scholarly writings. Vol. 151.

Intertek Moody Marine. 2013. Russian Sea of Okhotsk Mid-water Trawl Walleye Pollock (Theragra chalcogramma) Fishery. Final certification report. Robert O'Boyle, David Japp, Andrew Payne, Steve Devitt. Intertek Moody Marine Merlin House Stanier Way The Wyvern Business Park Derby. DE216BF

Intertek. 2013. Fishery Managmenet Certificate of Registration, MML – F -152. Russian Sea of Okhotsk Mid-water Trawl Walleye Pollock (Theragra chalcogramma) Fishery. Access through https://www.msc.org.

Legault C.M. 2008. Technical Documentation for ASAP Version 2.0 NOAA Fisheries Toolbox (http://nft.nefsc.noaa.gov/).

Legault, C. M., J. E. Powers, and V. R. Restrepo. 2001. Mixed Monte Carlo/bootstrap approach to assessing king and Spanish mackerel in the Atlantic and Gulf of Mexico: Its evolution and impact. Amercian Fisheries Society Symposium 24:1–8.

Magnusson, A., Punt, A. E., & Hilborn, R. (2013). Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and MCMC. *Fish and Fisheries*, *14*(3), 325-342.

Manly B. F. J. (1997) Randomization, Bootstrap and Monte Carlo Methods in Biology, 2nd edn. Chapman & Hall, London.

McAllister, M. K., & Kirkwood, G. P. (1998). Bayesian stock assessment: a review and example application using the logistic model. *ICES Journal of Marine Science: Journal du Conseil*, *55*(6), 1031-1060.

Methot, R. D. (1989). Synthetic estimates of historical abundance and mortality for northern anchovy. In *Am. Fish. Soc. Symp* (Vol. 6, pp. 66-82).

Methot, R. D., & Wetzel, C. R. (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management.*Fisheries Research*, *142*, 86-99.

Patterson K.R. 1994. Technical reference for the Integrated Catch-at-Age Programmes, Version 1.2 //SOAFD Marine Laboratory. Aberdeen. 13 p.

Quinn, T.J., Deriso R.B. 1999. Quantitative Fish Dynamics. Oxford University Press. New York. 542 pp.

Polacheck, T., Hilborn, R., & Punt, A. E. (1993). Fitting surplus production models: comparing methods and measuring uncertainty. *Canadian Journal of Fisheries and Aquatic Sciences*, *50*(12), 2597-2607.

Pope J.G., Shepherd J.G. 1982. A simple method for consistent interpretation of catch-at-age data. J. Cons. Cons. Int. Explor. Mer. 40: p. 176-184.

Prager, M. H., Porch, C. E., Shertzer, K. W., & Caddy, J. F. (2003). Targets and limits for management of fisheries: a simple probability-based approach.*North American Journal of Fisheries Management*, *23*(2), 349-361.

Punt, A. E., & Hilborn, R. A. Y. (1997). Fisheries stock assessment and decision analysis: the Bayesian approach. *Reviews in Fish Biology and Fisheries*, *7*(1), 35-63.

Punt, A.E., D. C. Smith, K. KrusicGolub, and S. Robertson. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. Canadian Journal of Fisheries and Aquatic Sciences 65: 1991-2005.

Restrepo, V. R., Hoenig, J. M., Powers, J. E., Baird, J. W., & Turner, S. C. (1992). A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. *Fishery Bulletin*, *90*(4), 736-748.

Richards, L. J., J. T. Schnute, A. R. Kronlund, and R. J. Beamish. 1992. Statistical methods for the analysis of ageing error. Canadian Journal of Fisheries and Aquatic Sciences 49: 1801-1815.

Shertzer, K. W., Prager, M. H., & Williams, E. H. (2008). A probability-based approach to setting annual catch levels. *Fishery Bulletin*, *106*(3), 225-232.

SEDAR Procedural Guidance, 2010. SEDAR Procedural Workshop IV: Characterizing and Presenting Assessment Uncertainty.

SEDAR. 2015. SEDAR 40 – Atlantic Menhaden Stock Assessment Report. SEDAR, North Charleston SC. 643 pp. available online at: http://www.sefsc.noaa.gov/sedar/Sedar_Workshops.jsp?WorkshopNum=40

Varkentin A.I. and Ilyin O.I. 2014 “Analysis of efficiency of the pollock fishery strategy in North Sea of Okhotsk. Uncertainty considerations for pollock stock assessment and TAC planning the North Sea of Okhotsk”. Research report. June 2014, KamchatNIRO.

Varkentin A.I. and Ilyin O.I. 2015 “Analysis of efficiency of the pollock fishery strategy in North Sea of Okhotsk. Uncertainty considerations for pollock stock assessment and TAC planning the North Sea of Okhotsk”, June 2014, KamchatNIRO.

Williams E.H. and K.W. Shertzer. 2015. Technical documentation of the Beaufort Assessment Model (BAM). US Departmetn of Commerce. NOAA Technical Memorandum. NMFS- SEFSC - 671. 433 p.