INTER-AMERICAN TROPICAL TUNA COMMISSION

9TH STOCK ASSESSMENT REVIEW MEETING

LA JOLLA, CALIFORNIA (USA)
12-16 MAY 2008

DOCUMENT SARM-9-09

AN ESTIMATION METHOD FOR IN-SEASON MANAGEMENT

Mark N. Maunder, Alexandre Aires-Da-Silva, and Richard Deriso

ABSTRACT

An in-season abundance estimation method is developed based on at-sea reports of catch and effort (capacity at sea) data. The method uses autocorrelated random effects for catchability and start-of-year biomass to share information from completed years with the year for which the in-season estimate is being conducted. Start-of-year biomass is set equal to estimates from the full stock assessment for all years except the year for which the in-season estimate is being conducted, which is estimated as a model parameter. Harvest rates can be applied to this biomass estimate to determine the annual total allowable catch. The method is applied to data for yellowfin and bigeye tuna in the eastern Pacific Ocean. Cross-validation tests are used to evaluate the performance of the method. The method performs well with an average error in the range of about a 9%-15% for the estimates of biomass from one quarter of a year's data.

1. INTRODUCTION

Setting a total allowable catch (TAC) generally requires an estimate of abundance and a harvest rate policy. For short-lived species like yellowfin tuna, estimates of current biomass are problematic due to lags in data and fluctuations in recruitment to the population. Estimates of current biomass from full stock assessments (*e.g.* Maunder 2007) are frequently uncertain and biased. Therefore, annual setting of TACs may require alternative methods to estimate current abundance.

In-season estimates of biomass calculated using the most up-to-date information may provide an advantage over full stock assessments for setting annual TACs. The TAC could be updated during the season based on this in-season estimate of biomass. There are several approaches that have been developed for in-season estimation of abundance, and they are commonly used for estimating run sizes for salmon fisheries. Each fishery has its own characteristics and available data, therefore specific in-season estimators need to be developed for the fishery of interest.

We present an estimation method to obtain in-season updates of start-of-the-year biomass that uses biomass estimates from a full stock assessment and weekly reported at-sea catch and effort (capacity at sea) data. The method uses autocorrelated random effects for catchability and start-of-year biomass to improve estimation of biomass for the year of interest. We test the method using cross-validation and apply it to yellowfin and bigeye tuna in the eastern Pacific Ocean (EPO).

2. METHODS

We follow the approach of Maunder (unpublished manuscript) modified by 1) replacing the start-of-year biomass with that estimated in a full stock assessment model (except for the current year, which is estimated as a parameter); 2) treating catchability and start-of-year biomass as random effects; and 3) including autocorrelation in catchability and start-of-year biomass. The method is similar to that used for squid populations (Rosenberg *et al.* 1990; McAllister *et al.* 2004). The biomass through the year is modeled using a population dynamics model. The change in abundance is separated into the four main processes: growth (G), survival (S), recruitment (R), and catch. Growth and survival are assumed to be constant over time. Recruitment is incorporated in two ways: 1) implicitly in the estimated biomass at the start of the year, and 2) explicitly through a constant amount each week. Catch by week is assumed to be known. The population dynamics within a year is given by

$$B_{y,w+1} = B_{y,w}\left(1+G\right)S + R - C_{y,w} \qquad \text{Equation 1}$$

Where $B_{y,w}$ is the biomass in week w of year y, and $C_{y,w}$ is the catch in week w of year y.

Equation 1 can be used to implement a linear regression for abundance within a year by setting $G = 0$, $S = 1$, $C = 0$, and estimating R as the slope and the start-of-year biomass is the intercept. The linear regression can be used as an alternative to the model based on population dynamics .

When applying the method for in-season abundance estimation, for all years, except the year for which the in-season estimate is being made, the start-of-year biomass is set equal to the start-of-year biomass estimate from the full stock assessment model. The start-of-year biomass for the year for which the in-season estimate is being made is estimated as a model parameter.

The catch per capacity at sea (CPUE) is used as an index of relative abundance, and the constant of proportionality (q) is assumed constant within a year, but varies as an autocorrelated random effect among years. A normal-distribution-based likelihood function is used to fit the CPUE data

$$-\ln L = \sum_{y,w}\left\{\ln\left[\sigma_{CPUE}\right] + \frac{\left(CPUE_{y,w} - q_y B_{y,w}\right)^2}{2\sigma^2_{CPUE}}\right\} \qquad \text{Equation 2}$$

$$q_y = \mu_q + \varepsilon_{q,y}$$

Following the approach used by Punt (2002) to include autocorrelation in annual recruitment deviates, we model autocorrelation in catchability.

$$\varepsilon_{q,1} = \eta_{q,1}$$

$$\varepsilon_{q,y} = \rho_q \varepsilon_{q,y-1} + \sqrt{1-\rho_q^2}\,\eta_{q,y}$$

$$\eta_{q,y} \sim N\left(0;\sigma_q^2\right)$$

The penalty added to the objective function is

$$-\ln P\left(\eta_q\right) = \sum_{y}\left\{\ln\left[\sigma_q\right] + \frac{\eta_{q,y}^2}{2\sigma_q^2}\right\}$$

Using an analogous approach for start-of-year biomass, but modified because the biomass is known for all but one year, we model autocorrelation in start-of-year biomass.

$$\varepsilon_{B,y} = B_y - \mu_B = \rho_B \varepsilon_{B,y-1} + \sqrt{1-\rho_B^2}\,\eta_{B,y}$$

$$\eta_{B,y} = \frac{B_y - \mu_B - \rho_B \varepsilon_{B,y-1}}{\sqrt{1-\rho_B^2}}$$

$$\eta_{B,1} = B_y - \mu_B$$

$$\eta_{B,y} \sim N\left(0,\sigma_B\right)$$

The penalty added to the objective function is

$$-\ln P\left(\eta_B\right)=\ln\left[\sigma_B\right]+\frac{\left(B_1-\mu_B\right)^2}{2\sigma_B^2}+\sum_{y=2toY}\left\{\ln\left[\sigma_B\right]+\frac{\left(\frac{B_y-\mu_B-\rho_B\varepsilon_{B,y-1}}{\sqrt{1-\rho_B^2}}\right)^2}{2\sigma_B^2}\right\}$$

3. APPLICATIONS TO YELLOWFIN AND BIGEYE TUNAS IN THE EPO

The in-season estimator is applied to data for yellowfin and bigeye tunas in the EPO during 1991-2006 and 1995-2006, respectively. The shorter period for bigeye is due to the limited amount of purse-seine catch of the species before 1993 and the rapid expansion of the floating-object fishery that occurred between 1993 and 1995. The parameters for growth (G = 0.02, 0.015) and survival (S = 0.985, 0.992) are fixed (for yellowfin and bigeye, respectively) based on values used in the stock assessment (Maunder 2007; Aires-Da-Silva and Maunder 2007). Biomass at the start of the year is set equal to the biomass of individuals one and a half years and 9 months of age and older estimated in the full stock assessments of yellowfin and bigeye, respectively (Maunder 2007; Aires-Da-Silva and Maunder 2007).

Data are taken from weekly catch reports provided by observers aboard purse-seine vessels at sea. Effort is represented by fishing capacity (in cubic meters of well capacity) at sea. These reports do not include all the catch from the EPO (*e.g.* unreported purse-seine, longline and pole-and-line catches) and the catch has not been adjusted for species composition errors. In addition, the biomass measure used does not represent the vulnerable biomass, and the effect of the catch on the measure of biomass may be more or less than expected. Therefore, an additional estimated parameter that scales the catch, δ, is included to adjust for these factors.

First, the best model is selected by applying the method to all years of data and fixing the start-of-year biomass to that estimated in the full stock assessment for all years of data. The model parameters are estimated in a stepwise fashion and the likelihood ratio test is used to determine which parameters to include in the model. This is repeated until no new parameter significantly improves the fit to the model. The model with the selected parameters is then used in the tests described below.

The model is coded in *AD Model Builder*, and the random effect is implemented using the Laplace approximation method (Skaug and Fournier 2006)

To test the performance of the method for in-season assessment, the method is repeated using each year as the in-season estimation year. The estimated start-of-year biomass for the in-season estimation year is compared to that estimated by the full stock assessment model. This cross-validation test probably overestimates the performance of the in-season estimator because it is not based on one-step-ahead predictions, which is how the in-season estimator would be applied. Therefore, the biomass and catchability for the year of interest are autocorrelated with the equivalent values in years both before and after the year of the in-season estimate. In an actual application of the in-season estimator, only the prior years would be available. However, cross-validation tests using only prior years are limited, because this reduces the number of years available to estimate the autocorrelation and other model parameters. To maximize the number of years of data used, we also apply the cross-validation test by excluding either the years before the year of interest when the year of interest is in the first half of the time series, or the years after the year of interest when the year of interest is in the second half of the time series.

The CPUE data for bigeye increase within a year. This does not follow the pattern of an exploited population with annual spawning described by the population dynamics model. In addition, fitting CPUE well may not mean that start-of-year abundance is estimated well. Therefore, as an alternative to selecting a model using the likelihood ratio tests, we chose a model that appears reasonable for the situation, a linear regression with autocorrelation in biomass and catchability.

4. RESULTS

4.1. Yellowfin tuna

The model selected based on the likelihood ratio test includes a random effect for both catchability and biomass, autocorrelation parameters for both catchability and biomass, and a catch-scaling parameter (Table 1). Estimating the growth, mortality, and recruitment parameters did not improve the fit. The best model was substantially better than assuming constant biomass within a year, but only slightly better than a linear regression with autocorrelated random effects (Table 1).

The model estimates that the biomass declines during the year and increases substantially from the end of one year to the start of the next (Figure 1). The model fits the general pattern in the CPUE data, but there is a large amount of variation in the CPUE data around the predictions (Figure 1). The model estimated that there is substantial autocorrelation in both biomass and catchability (Table 2). The catch-scaling factor was less than 1 (Table 2), indicating that the model predicted that the catch has less effect on the modeled biomass than would be normally assumed.

The model estimates the abundance well, using either the full,quarter or half year of data (Figure 2). The performance is degraded based on the test that drops years to account for using the data to predict the biomass in the current year (Figure 3). The average percent absolute error is 9%, 9%, and 10%, for a full, quarter, and half year of data, respectively (Figure 4). The average percent absolute error based on the test that excludes years to account for using the data to predict the biomass in the current year (using data for one quarter of a year) is higher (15%) (Figure 5).

4.2. Bigeye tuna

The model selected based on the likelihood ratio test includes a random effect for both catchability and biomass, a catch-scaling parameter, and recruitment (Table 2). Autocorrelation was not significant for either biomass or recruitment. Estimating the growth or mortality parameters did not improve the fit. The best model was substantially better than assuming constant biomass within a year, but only slightly better than a linear regression with autocorrelated random effects (Table 3).

The model estimates that the biomass increases during the year and decreases substantially from the end of one year to the start of the next (Figure 6). The model fits the general pattern in the CPUE data, but there is a large amount of variation in the CPUE data (Figure 6). The catch-scaling factor was substantially greater than 1 (Table 4), indicating that the model predicted that the catch has a much greater effect on the modeled biomass than would be normally assumed.

The model estimates the abundance only moderately well, using either the full, quarter, or half year of data (Figure 7). The average percent absolute error is 15%, 16%, and 14%, for a full, quarter, and half year of data, respectively (Figure 8). Because no autocorrelated parameters were included in the selected model, the cross-alidation test that excludes years of data was not appropriate.

The linear regression model with autocorreled random effects produces estimates of biomass closer to the true estimates (Figure 9), with an average percent absolute error of 9% (Figure 10). The average percent absolute error based on the test that excludes years to account for using the data to predict the biomass in the current year (using data for one quarter of a year) is higher (13%) (Figure 10).

5. DISCUSSION

The in-season estimation procedure performs well for yellowfin, but less well for bigeye. Yellowfin shows declining CPUE within a year, which might be expected for a population that is exploited and has annual recruitment. Yellowfin spawn throughout the year in areas where the sea surface temperature is above 24º (Schaefer 1998); therefore, it is not expected that recruitment occurs as a single event each year. Despite this, a depletion model used to estimate the abundance of yellowfin in the EPO (Maunder unpublished manuscript) and the in-season estimator presented here perform well at estimating yellowfin abundance.

The in-season estimator performed only moderately well for bigeye. The CPUE increases during the year, which is counter to an exploited population with annual recruitment that is represented by the population dynamics model. Recruitment during the year could explain increasing CPUE, but not the decrease in CPUE between consecutive years. The estimate of the scaling factor for catch is unrealistic, and therefore the model based on population dynamics is considered only as a method to estimate a trend, and not a representation of the true population dynamics. The linear regression model performs better; however, there is only moderate autocorrelation in the biomass and catchability. The contrast in bigeye abundance is much less than in yellowfin, and it is not known if the method would work well for bigeye if the abundance decreased or increased greatly. The overestimation of abundance in recent years is of particular concern if this method is to be applied to set catch limits for bigeye. The CPUE at the start and end of the year is distorted by the apparent reduction in effort by vessels targeting yellowfin associated with dolphins. Separating the effort (capacity at sea) into vessels that fish for dolphin-associated tuna and those that fish on floating objects (*e.g.* by area or by the vessel's targeting practice) may improve the performance of the in-season estimator.

The estimates do not appear to improve greatly as more weeks of data are used. One quarter of a year's data provides a good estimate of abundance. This would allow an annual catch limit to be set just after the first quarter of the year. The model could be used to put confidence intervals on the estimate of abundance to provide a measure of uncertainty that could be taken into consideration when setting limits.

The performance of the model is degraded when tested by excluding years so that the year of interest has only one neighboring year. However, by excluding years, this test uses less data to estimate the model parameters than would be used to estimate the most recent abundance, which likely degrades the performance of the model. Therefore, the real performance of the method probably lies in between the two tests.

The method performs well for yellowfin and, in a modified format, for bigeye in the EPO, based on data from weekly catch reports and capacity at sea. It is therefore a practical method for in-season estimation of biomass and and for setting TACs. However, care should be taken when applying the method to bigeye, due to the limited contrast in abundance under which the method has been tested. Improvements to the methodology may improve the performance of the method.

Acknowledgements

Joanne Boster and Joydelee Marrow provided the weekly report catch and capacity at sea data.

References

Aires-da-Silva, A. and Maunder, M.N. (2007). Status of bigeye tuna in the eastern Pacific Ocean in 2006 and outlook. Inter-Amer. Trop. Tuna Comm., Stock Assessment Report, 8: 105-228.

Maunder, M.N. (unpublished manuscript) A depletion estimator for in-season management of yellowfin tuna

Maunder, M.N. (2007). Status of yellowfin tuna in the eastern Pacific Ocean in 2006 and outlook. Inter-Amer. Trop. Tuna Comm., Stock Assessment Report, 8: 3-104.

McAllister, M.K., Agnew, D.J., Kirkwood, G.P., Beddington, J.R., Hill, S.L. 2004. A Bayesian hierarchical formulation of the De Lury stock assessment model for abundance estimation of Falkland Islands' squid (*Loligo gahi*). Can. J. Fish. Aquat. Sci. 61: 1048–1059.

Rosenberg, A.A., Kirkwood, G.P., Crombie, J.A., and Beddington, J.R. 1990. The assessment of stocks of annual squid species. Fish. Res. 8: 335–350.

Punt, A. E. 2002. SSC default rebuilding analysis: technical specifications and User Manual Version 2.4. School of Fisheries, University of Washington.

Schaefer, K.M. 1998. Reproductive biology of yellowfin tuna (Thunnus albacares) in the eastern Pacific Ocean. Inter-Amer. Trop. Tuna Comm., Bull. 21: 205-272.

Skaug HJ, Fournier DA (2006) Automatic approximation of h marginal likelihood in non-Gaussian hierarchical models. Computational Statistics and Data Analysis 51:699-709

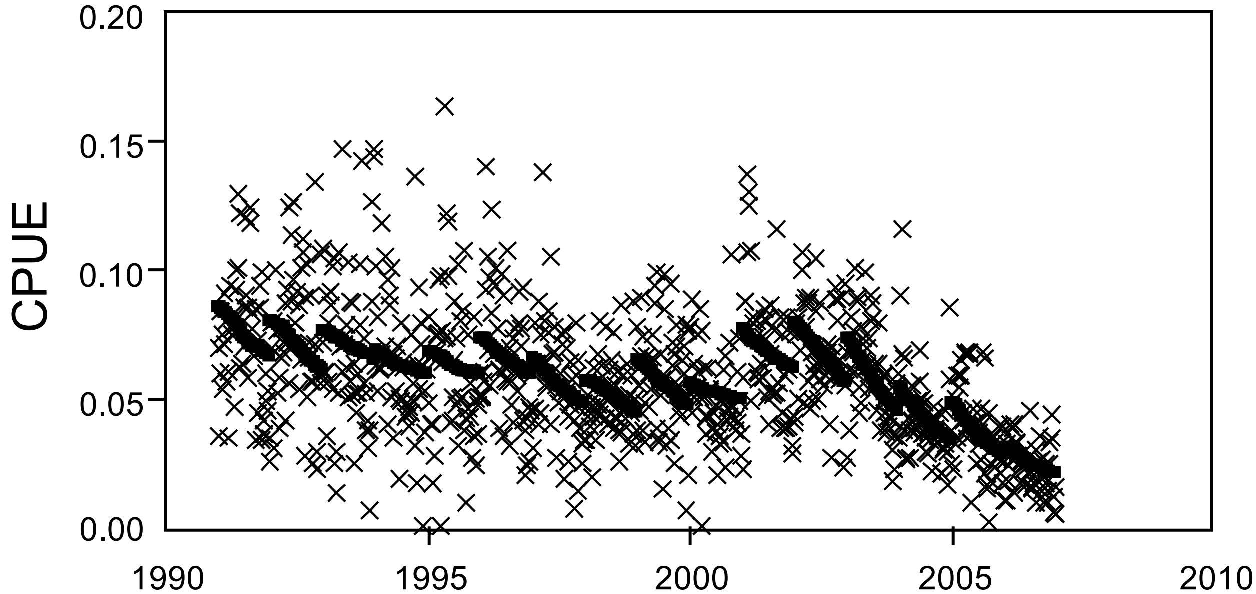

FIGURE 1. Fit of the selected model for yellowfin tuna with known start-of-year biomass (solid points) to CPUE data (crosses).

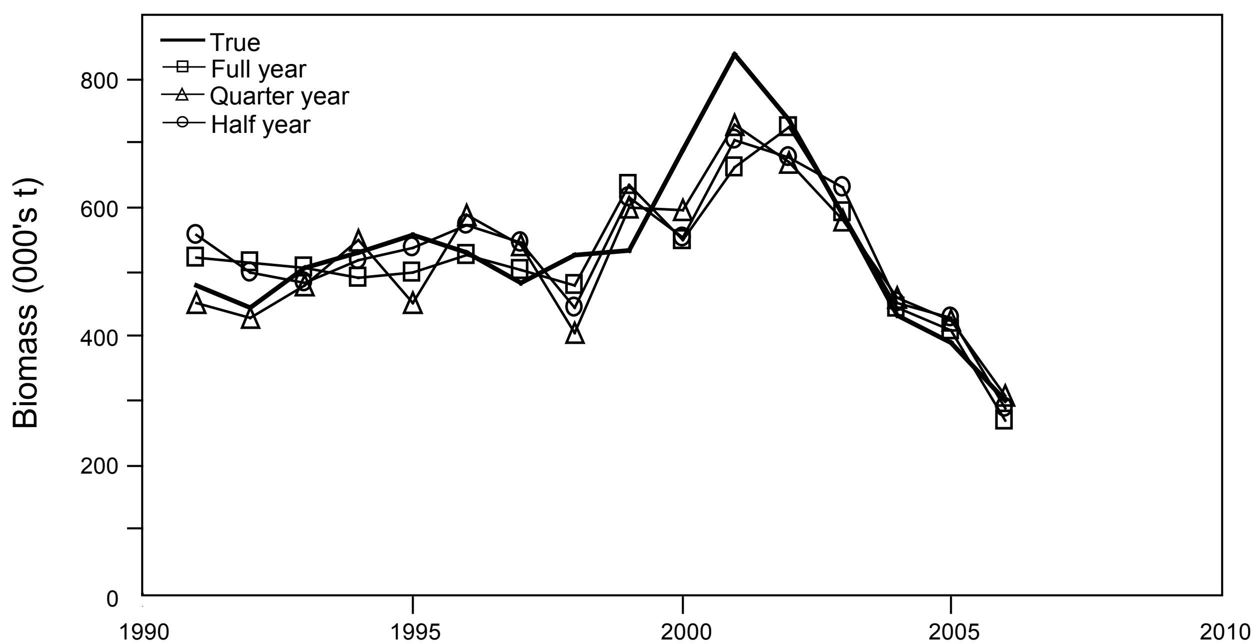

FIGURE 2. Estimates of yellowfin biomass from the selected model, using a full year, quarter year, and half year of data, compared to the "true" biomass (as estimated by the full stock assessment).

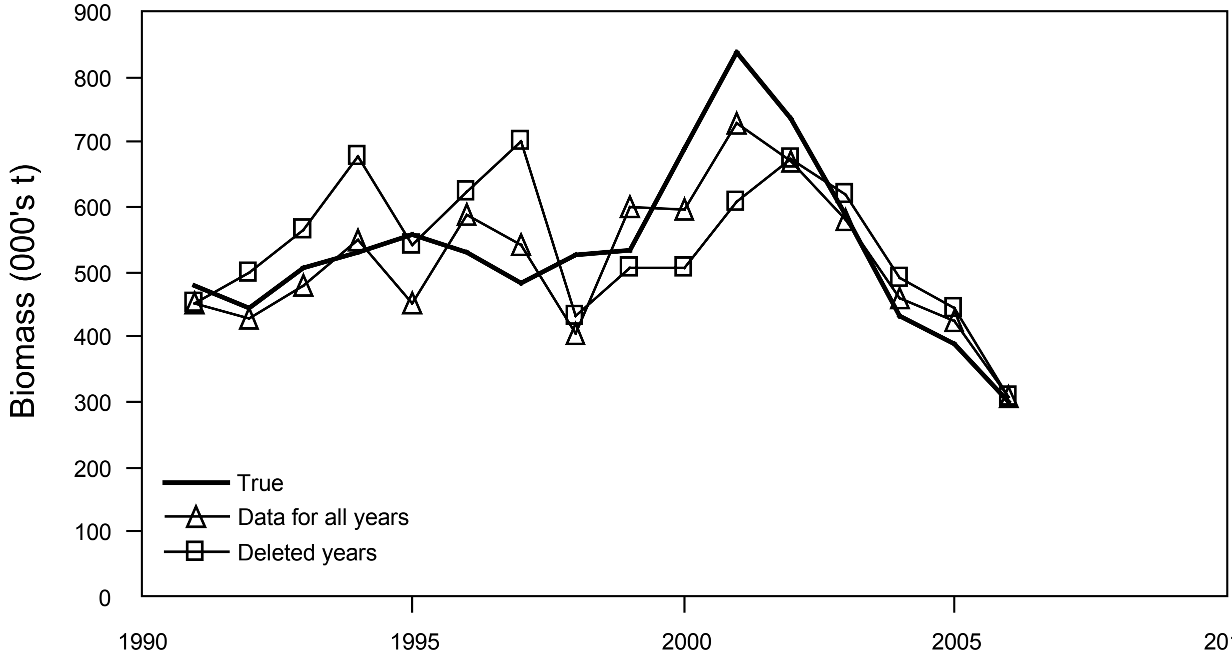

FIGURE 3. Estimates of yellowfin biomass from the selected model, using a quarter year of data, calculated with either data for all years or deleting some years (as explained in the text) compared to the "true" biomass (as estimated by the full stock assessment).

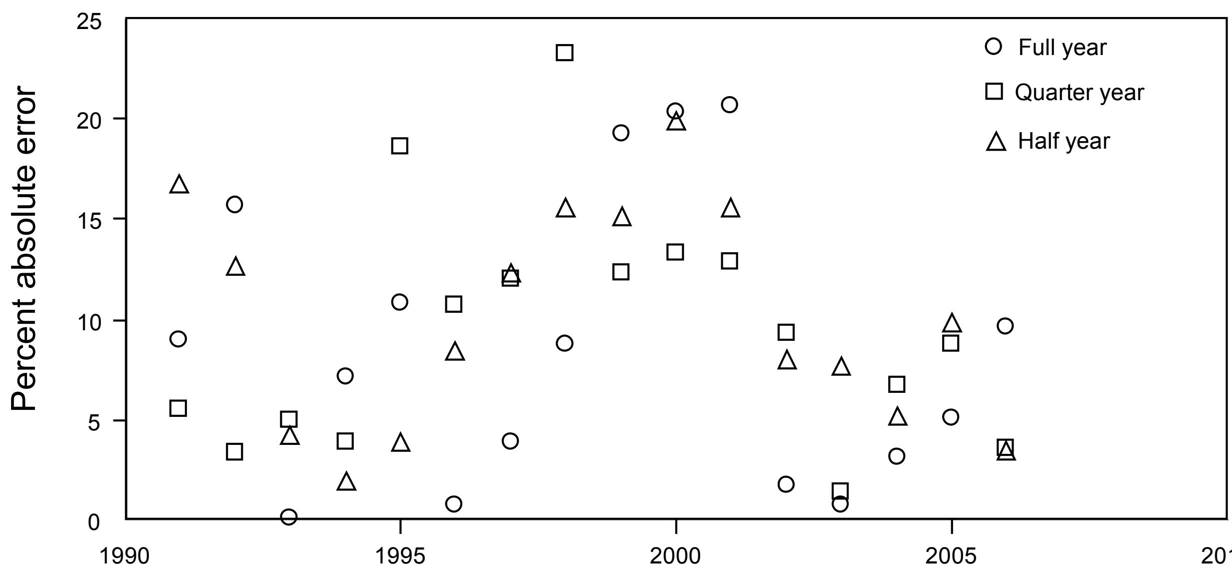

FIGURE 4. Percent absolute error in estimating start-of-year biomass in the corresponding year on the x-axis using a full year, quarter year, or half year of data for that year. Results are for the selected model for yellowfin.

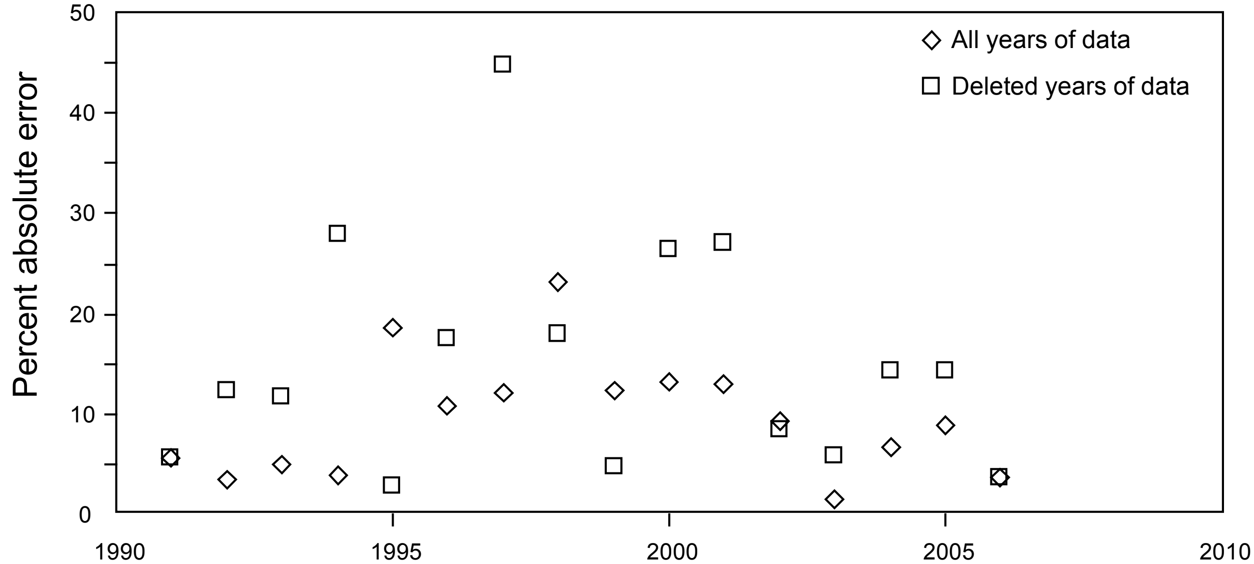

FIGURE 5. Percent absolute error for estimating start-of-year biomass in the corresponding year on the x-axis using a quarter year of data for that year. Results are from the selected model for yellowfin calculated with either data for all years or with years deleted either before or after the year of interest (as described in the text).

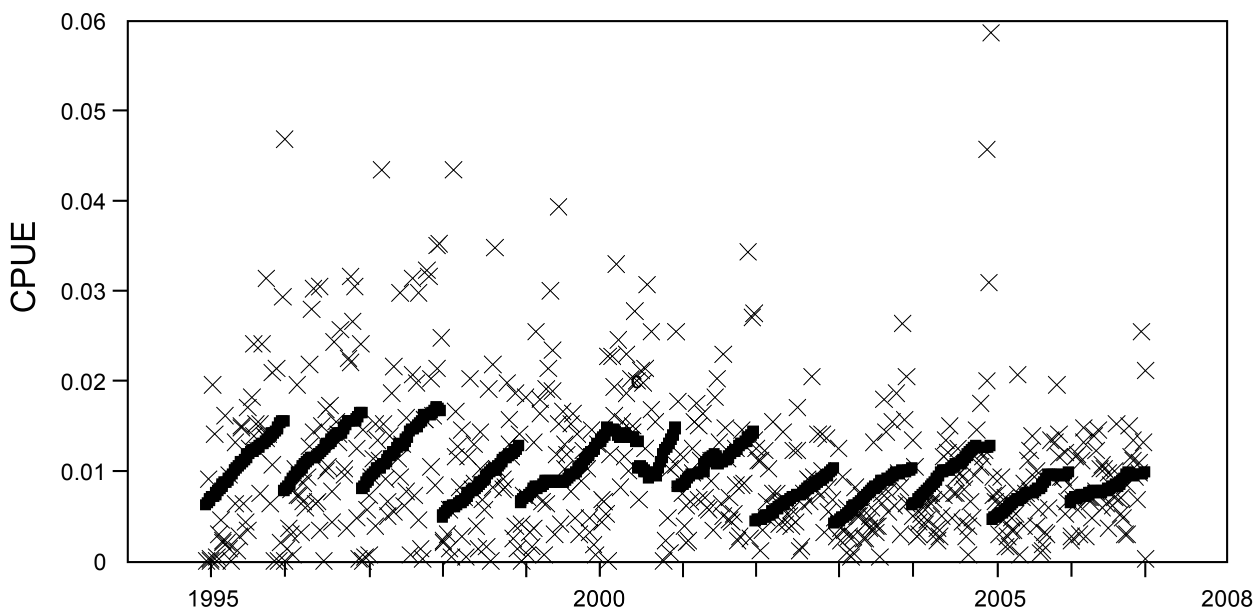

FIGURE 6. Fit of the selected model for bigeye with known start-of-year biomass (solid points) to CPUE data (crosses).

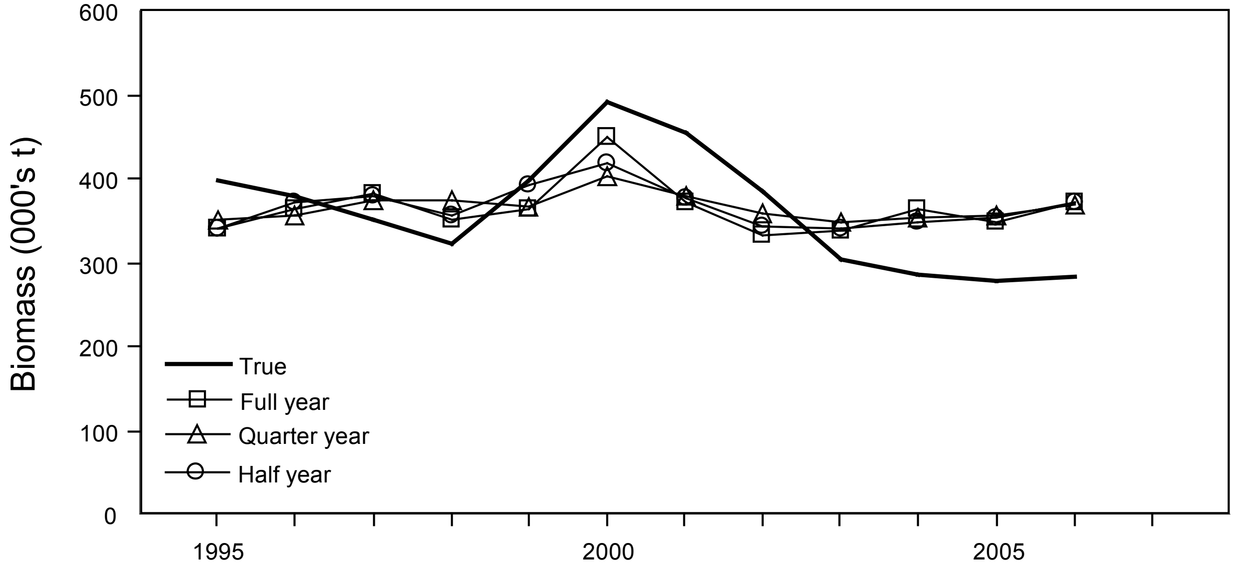

FIGURE 7. Estimates of bigeye biomass from the selected model, using a full year, quarter year, and half year of data, compared to the "true" biomass (as estimated by the full stock assessment).

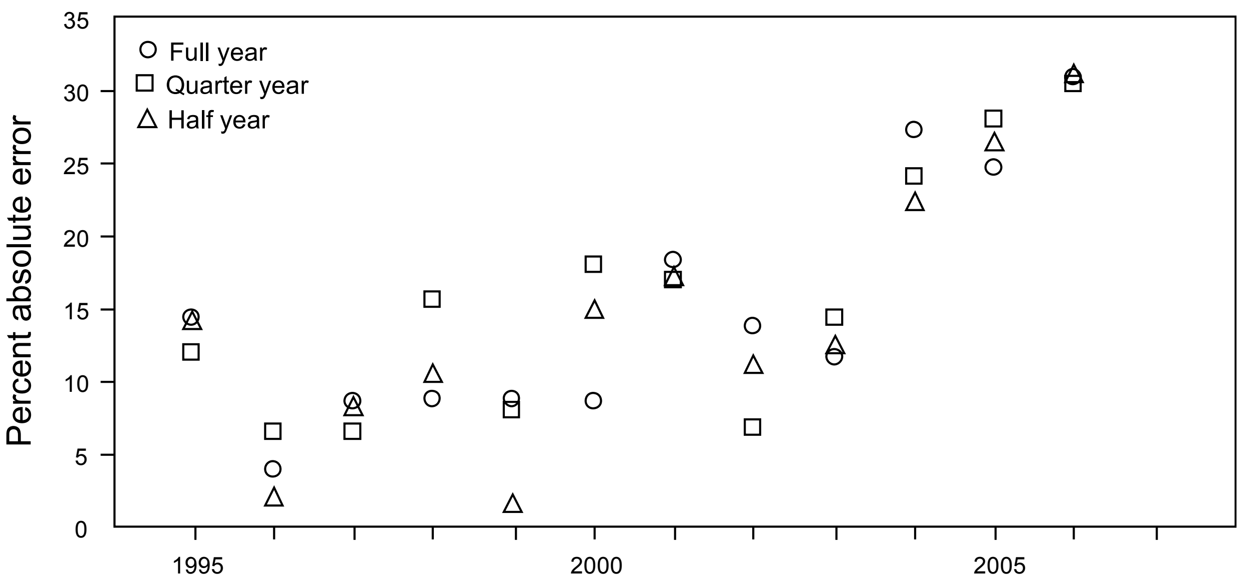

FIGURE 8. Percent absolute error for estimating start-of-year biomass in the corresponding year on the x-axis using a full year, quarter year, or half year of data for that year. Results are for the selected model for bigeye.

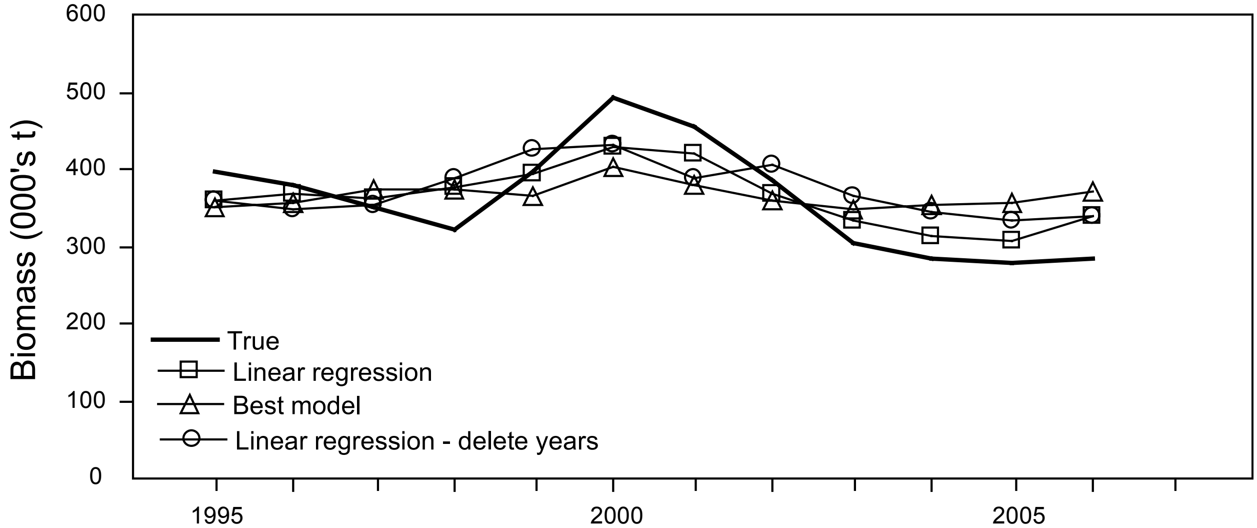

FIGURE 9. Estimates of bigeye biomass calculated from the selected model and the linear regression with data for all years and with years deleted either before or after the year of interest (as described in the text).

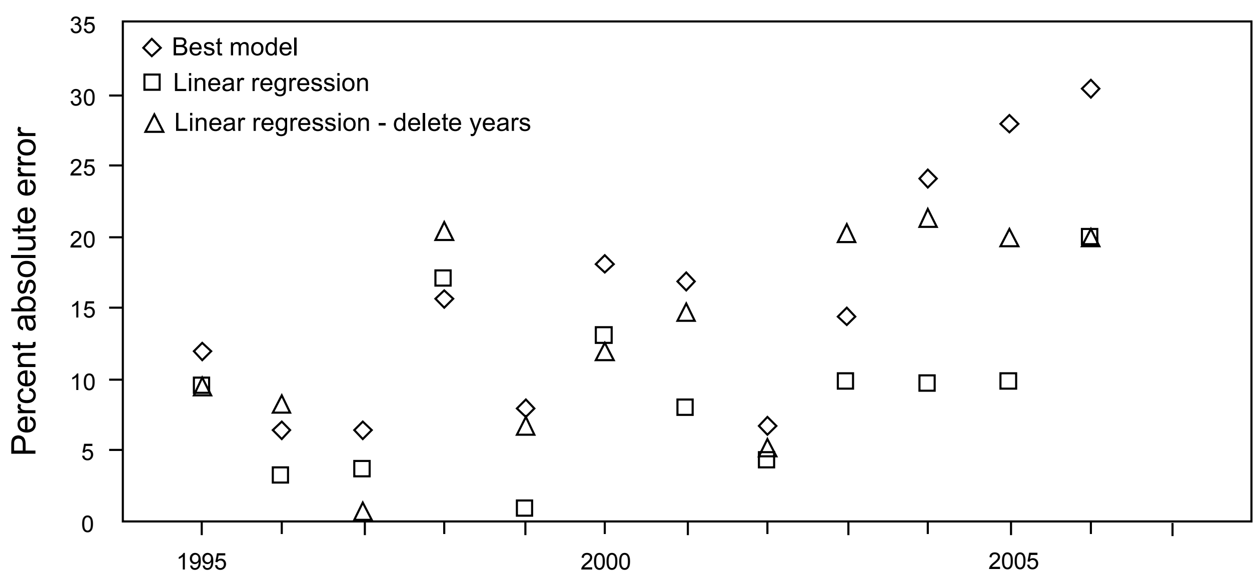

FIGURE 10. Percent absolute error for estimating start-of-year biomass in the corresponding year on the x-axis using a quarter year of data for that year. Results for bigeye using the selected model, the linear regression model with autocorrelated random effects. The results for the linear regression model are calculated with data for all years and with years deleted either before or after the year of interest (as described in the text).

TABLE 1. Likelihood values for models run to select the best model for yellowfin tuna. "Yes" indicates that the parameter is estimated and "No" that it is not. A zero indicates that the parameter was fixed at zero. "Fixed" indicates that the parameter was fixed at the appropriate value. The shaded cells indicate the parameter that was chosen in the forward stepwise approach of including parameters.

	σ_q	ε_q	ρ_q	ρ_B	δ	G	S	R	-lnL	-lnL+ lnL(best model)
	no	0	0	0	fixed	fixed	fixed	0	-2344.11	114.00
	no	0	0	Yes	fixed	fixed	fixed	0	-2346.79	111.32
	yes	yes	0	0	fixed	fixed	fixed	0	-2449.38	8.73
	no	0	0	0	fixed	fixed	fixed	yes	-2352.18	105.93
	no	0	0	0	yes	fixed	fixed	0	-2346.22	111.89
	yes	yes	yes	0	fixed	fixed	fixed	0	-2452.81	5.30
	yes	yes	0	yes	fixed	fixed	fixed	0	-2452.06	6.05
	yes	yes	0	0	yes	fixed	fixed	0	-2451.63	6.48
	yes	yes	0	0	fixed	fixed	fixed	yes	-2451.96	6.15
	yes	yes	yes	yes	fixed	fixed	fixed	0	-2455.49	2.62
	yes	yes	yes	0	yes	fixed	fixed	0	-2455.43	2.68
	yes	yes	yes	0	fixed	fixed	fixed	yes	-2455.36	2.75
	yes	yes	yes	yes	yes	fixed	fixed	0	-2458.11	0.00
	yes	yes	yes	yes	fixed	fixed	fixed	yes	-2458.04	0.07
	yes	yes	yes	yes	yes	fixed	fixed	yes	-2458.12	-0.01
Biological	yes	yes	yes	yes	yes	yes	fixed	0	-2458.64	-0.53
parameters	yes	yes	yes	yes	yes	fixed	yes	0	-2458.64	-0.53
Linear	yes	yes	yes	yes	small	0	1	0	-2439.97	18.14
regression	yes	yes	yes	yes	small	0	1	yes	-2456.13	1.98

TABLE 2. Parameter estimates for the selected model for yellowfin with known start-of-year biomass.

μ_q	σ_q	σ_{CPUE}	δ	μ_B	σ_B	ρ_B	ρ_q
1.31E-07	4.15E-08	0.023	0.871	534812	108233	0.437	0.736

TABLE 3. Likelihood values for models run to select the best model for bigeye tuna. "Yes" indicates that the parameter is estimated and "No" that it is not. A zero indicates that the parameter was fixed at zero. "Fixed" indicates that the parameter was fixed at the appropriate value. The shaded cells indicate the parameter that was chosen in the forward stepwise approach of including parameters.

	σ_q	ε_q	ρ_q	ρ_B	δ	G	S	R	-lnL	-lnL+ lnL(best model)
	no	0	0	0	fixed	fixed	fixed	0	-2366.33	18.60
	no	0	0	yes	fixed	fixed	fixed	0	-2368.18	16.75
	yes	yes	0	0	fixed	fixed	fixed	0	-2375.94	8.99
	no	0	0	0	fixed	fixed	fixed	yes	-2370.76	14.17
	no	0	0	0	yes	fixed	fixed	0	-2367.49	17.44
	yes	yes	yes	0	fixed	fixed	fixed	0	-2376.12	8.81
	yes	yes	0	yes	fixed	fixed	fixed	0	-2377.79	7.14
	yes	yes	0	0	yes	fixed	fixed	0	-2377.12	7.81
	yes	yes	0	0	fixed	fixed	fixed	yes	-2380.87	4.06
	yes	yes	yes	0	fixed	fixed	fixed	yes	-2380.96	3.97
	yes	yes	0	yes	fixed	fixed	fixed	yes	-2382.72	2.21
	yes	yes	0	0	yes	fixed	fixed	yes	-2384.93	0.00
	yes	yes	yes	0	yes	fixed	fixed	yes	-2384.98	-0.05
	yes	yes	0	yes	yes	fixed	fixed	yes	-2386.79	-1.86
	yes	yes	yes	yes	yes	fixed	fixed	yes	-2386.83	-1.90
Biological	yes	yes	0	0	yes	yes	fixed	yes	-2385.41	-0.43
parameters	yes	yes	0	0	yes	fixed	yes	yes	-2385.41	-0.43
Linear	yes	yes	0	0	small	0	1	0	-2370.21	14.77
regression	yes	yes	0	0	small	0	1	yes	-2381.18	3.80
	yes	yes	yes	yes	small	0	1	yes	-2383.14	1.84

TABLE 4. Parameter estimates for the selected model for bigeye with known start-of-year biomass.

μ_q	σ_q	σ_{CPUE}	δ	μ_B	σ_B	ρ_B	ρ_q	R
1.82E-08	6.81E-09	0.0096	10.506	370242	67189	0	0	13150

TABLE 5. Parameter estimates for the linear regression bigeye tuna model with known start of year biomass.

μ_q	σ_q	σ_{CPUE}	μ_B	σ_B	ρ_B	ρ_q	R
1.87E-08	4.19E-09	0.0098	370435	57579	0.41	0.19	6590