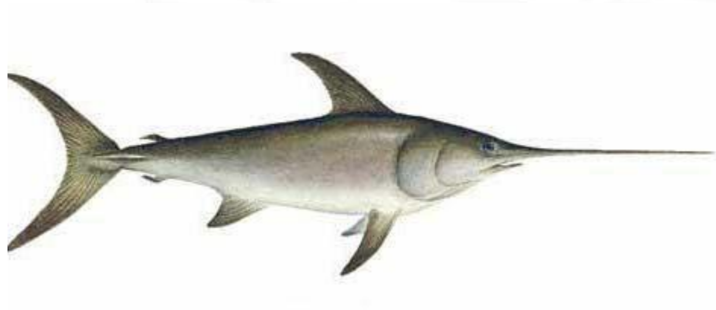
ISC/13/BILLWG-2/09

International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean

Application of Bayesian Production Model to Assess Pacific Blue Marlin (*Makaira nigricans*) in 2013[1]

Jon Brodziak, Joseph O'Malley, and Annie J. Yau
NOAA NMFS Pacific Island Fisheries Science Center
Honolulu, Hawaii, USA

Yi-Jay Chang
Joint Institute for Marine and Atmospheric Research
University of Hawaii
Honolulu, Hawaii, USA



[1]Working document submitted to the ISC Billfish Working Group Workshop, 20-28 May 2013, Shimizu, Shizuoka, Japan. Document not to be cited without author's written permission.

Application of a Bayesian Production Model to Assess Pacific Blue Marlin (*Makaira nigricans*) in 2013

Jon Brodziak[1], Joseph O'Malley[1], Annie J. Yau[1], Yi-Jay Chang[2]

[1] Pacific Islands Fisheries Science Center, NOAA Fisheries, 2570 Dole Street, Honolulu, HI 96822-2396, USA

[2] Joint Institute for Marine and Atmospheric Research, University of Hawaii, 2570 Dole Street, Honolulu, HI 96822-2396, USA

ABSTRACT

Bayesian surplus production models were developed for assessing the Pacific blue marlin population under alternative assumptions about catch-per-unit effort (CPUE) relative abundance indices. Alternative production models were developed and fit for two treatments of the annual intrinsic growth parameter (*r*): a single-*r* and a time-varying multiple-*r* with a different intrinsic growth rate parameter for each year. Biomass production was modeled with a 3-parameter production model that allowed production to vary from the symmetric Schaefer curve using an estimated shape parameter. Input data included nominal landings of Pacific blue marlin collected from all available sources during 1950–2011. Two alternative catch scenarios were investigated and fit with alternative production models: 1950-2011 and 1971-2011, the latter time period representing the period of the most consistent fishery data. Relative abundance indices for blue marlin consisted of standardized catch-per-unit effort (CPUE) for Japanese, Chinese-

Working paper submitted to the ISC Billfish Working Group Workshop, 20-28 May, 2013, Shimizu, Japan. Document not to be cited without author's written permission. Contact: Jon.Brodziak@noaa.gov.

Taipei, and USA longline fisheries. Annual coefficients of variation for CPUE were used to weight the annual observation error within each time series of relative abundance indices. Thus, the model fits to CPUE included heterogeneous annual observation errors.

A total of 36 model hypotheses were developed and fit to the alternative catch and CPUE data. Uninformative lognormal prior distributions for intrinsic growth rate and carrying capacity were assumed with coefficients of variation set at 100%. Goodness-of-fit diagnostics were developed for comparing the fits of alternative model configurations based on the root-mean squared error of CPUE fits, the standardized CPUE residuals, and the Deviance Information Criterion. Model selection results indicated that two models provided the most credible and best fits under the 1950-2011 catch scenario; these were the single-*r* and multiple-*r* models under the 1950-2011 catch scenario using the standardized Japanese CPUE estimates from 1975-1993 and 1994-2011 and the standardized Chinese Taipei CPUE estimates from 1979-1999 and 2000-2011. Model averaging was applied to summarize the results of the two credible models. Biomass estimates were not influenced by different P[1] prior values. The biomass status of blue marlin in 2011 suggest that the biomass was above the biomass (BMSY) to produce maximum sustainable yield (MSY) based on the model averaged values but the unconditional standard error suggests it may be slightly below BMSY. However, the model averaged estimates of harvest rate and the standard errors in 2011 did not exceed an overfishing threshold relative to MSY-based reference points. Overall, the production model results suggest that the blue marlin is likely not overfished relative to MSY-based reference points and did not experience overfishing during 1950-2011 with the possible exception of 2000-2006.

MATERIALS AND METHODS

Catch and CPUE Data Sources

Annual catch data for Pacific blue marlin were collected from all available sources including ISC countries, the WCPFC, and the IATTC on an annual basis (Figure 1) as reported in BILLWG (2013) and updated with corrections. The annual catch of Pacific blue marlin ranged from a low of 7,370 mt in 1954 to a high of 25,426 mt in 2003. Annual catches averaged 16,169 mt during 1952-2011, and in comparison, the catch in 2011 was reported to be 16,828 mt, or slightly above the long-term average. Blue marlin catch were predominated by Japanese catch until the mid-1990s, and at present Chinese Taipei produces the largest proportion of Pacific blue marlin catch by country.

CPUE abundance indices were collected from the standardizations reviewed and agreed upon in BILLWG (2013). There were a total of six CPUE abundance indices for Pacific blue marlin. These were the Japanese distant water longline 1975-1993 (CPUE 1) and 1994-2011 (CPUE 2) (Figure 2.1), the Hawaii longline (CPUE 3) (Figure 2.2), and the Chinese Taipei longline 1967-1978 (CPUE 4), 1979-1999 (CPUE 5), and 2000-2011 (CPUE 6) (Figure 2.3).

Production Model Assessment Method

Blue marlin production models were formulated as Bayesian-state space models with explicit observation and process error terms (see Meyer and Millar 1999). This Bayesian model has been used in some stock assessments where more complex assessment approaches were not applicable due to limited or conflicting data or other factors (see, for example, Brodziak et al. 2001, Brodziak et al. 2011). The biomass time series comprised the unobserved state variables, which were estimated from the observed relative abundance indices (i.e., CPUE) and catches using observation error likelihood function and prior distributions for model parameters (θ). In this case, the observation error likelihood measured the discrepancy between observed and predicted CPUE, and the prior distributions represented the relative degree of belief about the possible values of model parameters.

The process dynamics represented the fluctuations in exploitable blue marlin biomass based on density-dependent processes and fishery harvests. The production dynamics of biomass were based on a power function model with an annual time step. Two versions of the power function model were investigated: a single intrinsic growth rate model and a hierarchical multiple-*r* model in which there was a time-varying intrinsic growth parameter for each year T (r_T). The multiple- *r* model allows for time-dependence in the density-dependence experienced by the stock and is a generalization of the simpler single- *r* model. In what follows, we describe the structure of the multiple- *r* model and note that the single- *r* model has the year subscript removed except for the posterior distribution.

Under the process dynamics model, current biomass (B_T) depended on the previous biomass (B_{T-1}), catch (C_{T-1}), intrinsic growth rate r_{T-1}, carrying capacity (K), and a production shape parameter (M) for $T = 2,\ldots, N$.

(1) $$B_T = B_{T-1} + r_{T-1} \cdot B_{T-1}\left(1-\left(\frac{B_{T-1}}{K}\right)^M\right) - C_{T-1}$$

The production model shape parameter, *M*, determined where surplus production peaked as biomass varied as a fraction of carrying capacity. If the shape parameter was less than unity ($0 < M < 1$), then surplus production peaked when biomass was below ½ of K (i.e., a right-skewed production curve). If the shape parameter was greater than unity ($M > 1$), then biomass production peaked when biomass was above ½ of K (i.e., a left-skewed production curve). If the shape parameter was identically unity ($M = 1$), then the production model was identical to a discrete-time Schaefer production model where maximum surplus production occurred when biomass was equal to ½ of K. Thus, the shape of the biomass production curve could be symmetric, right- or left-skewed depending on the estimated value of M.

The power function model was reparameterized using the proportion of carrying capacity ($P = B/K$) to improve the efficiency of the Markov Chain Monte Carlo algorithm used to estimate parameters. Given this parameterization, the process dynamics for the power function model were

(2) $$P_T = P_{T-1} + r_{T-1} \cdot P_{T-1}\left(1 - P_{T-1}^{M}\right) - \frac{C_{T-1}}{K}$$

The values of biomass and harvest rate that maximize surplus production were relevant as biological reference points for maximum sustainable yield (MSY). For the discrete-time power function model, the biomass that produced MSY (B_{MSY}) was

(3) $$B_{MSY} = K \cdot \left(M+1\right)^{\frac{-1}{M}}$$

The corresponding harvest rate that produced MSY (H_{MSY}) was

(4) $$H_{MSY} = \bar{r}\left(1 - \frac{1}{M+1}\right)$$

where $\bar{r} = \frac{1}{N}\sum_{T} r_T$ was the average of the observed annual intrinsic growth rates and where the associated value of maximum sustainable yield (MSY) was

(5) $$MSY = \bar{r}\left(1-\frac{1}{M+1}\right)\cdot K\left(M+1\right)^{\frac{-1}{M}}$$

Thus, the production model produced direct estimates of biological reference points for blue marlin that are commonly used for determining stock status.

Observation Error Model

The observation error model related the observed fishery CPUE to the exploitable biomass of the blue marlin stock under each scenario. It was assumed that each CPUE index (I) is proportional to biomass with catchability coefficient Q

(6) $$I_T = QB_T = QKP_T$$

The observed CPUE dynamics for each fishing fleet were subject to natural sampling variation which was assumed to be lognormally distributed. The observation errors were distributed as $v_T = e^{V_T}$ where the V_T are iid normal random variables with zero mean and weighted variance $(w_T \cdot \tau)^2$ with standard deviation τ and weighting factor w_T. The weighting factors (w_T) of the annual CPUE variance terms reflected the relative uncertainty of the value of the CPUE index in year T and were scaled using the coefficient of variation (CV) of the difference between the observed and predicted log-transformed biomass indices. In particular, these annual weighting factors were calculated from the relative coefficients of variation of each annual CPUE index and the minimum observed CV of CPUE (min(CV[CPUE])) as w_T = CV[$CPUE_T$]/min(CV[CPUE]). Given the lognormal observation errors, the observation equations for each annual period indexed by $T = 1,\ldots,N$ were

(7) $$I_T = QKP_T \cdot \nu_T$$

This specified the general form of the observation error likelihood function $p(I_T \mid \theta)$ for each fishing fleet through time.

Process Error Model

The process error model related the dynamics of exploitable biomass to natural variability in demographic and environmental processes affecting the blue marlin stock. The deterministic process dynamics (Equation 2) were subject to natural variation as a result of fluctuations in life history parameters, trophic interactions, environmental conditions and other factors. In this case, the process error represented the joint effects of a large number of random multiplicative events which combined to form a multiplicative lognormal process under the Central Limit Theorem. As a result, the process error terms were independent and lognormally distributed random variables $\eta_T = e^{U_T}$ where the U_T were normal random variables with mean 0 and variance σ^2.

Given the process errors, the state equations defined the stochastic process dynamics by relating the unobserved biomass states to the observed catches and the estimated population dynamics parameters. Assuming multiplicative lognormal process errors, the state equations for the initial time period ($T = 1$) and subsequent periods ($T > 1$) were

(8) $$\begin{aligned} P_1 &= \eta_1 \\ P_T &= \left(P_{T-1} + r_{T-1} \cdot P_{T-1}\left(1 - P_{T-1}^{M}\right) - \frac{C_{T-1}}{K} \right) \cdot \eta_T \quad for\ T > 1 \end{aligned}$$

These coupled state equations set the conditional prior distribution for the proportion of carrying capacity, $p(P_T)$, in each time period *T*, conditioned on the proportion in the previous period.

Prior Distributions

Under the Bayesian paradigm, prior distributions are employed to quantify existing knowledge, or possible lack thereof, of the probable range of each model parameter. For the production model, the model parameters consisted of the carrying capacity, the intrinsic growth rate, the shape parameter, the catchability coefficients, the process and observation error variances, and the annual biomasses as a proportion of carrying capacity. Auxiliary information was incorporated into the formulation of the prior distributions when it was available.

Prior for Carrying Capacity

The prior distribution for the carrying capacity *p*(*K*) was a lognormal distribution with mean (μ_K) and variance (σ_K^2) parameters.

(9) $$p(K) = \frac{1}{\sqrt{2\pi}K\sigma_K}\exp\left(-\frac{(\log K - \mu_K)^2}{2\sigma_K^2}\right)$$

The variance parameter was set to achieve a coefficient of variation (CV) for *K* of 100%, e.g., $CV[K] = \left(\exp(\sigma_K^2)-1\right)^{\frac{1}{2}}$ = 1 and the mean *K* parameter was set to be 150,000 mt for the entire Pacific. This mean value was chosen to reflect the magnitude of exploitable biomass likely needed to support the observed fishery catches.

Prior for Intrinsic Growth Rate

For the single-*r* production model, the prior distribution for intrinsic growth rate *p*(*r*) was a lognormal distribution with mean (μ_r) and variance (σ_r^2) parameters set to achieve a CV for *r* of 100%.

$$\textbf{(10.1)} \quad p(r) = \frac{1}{\sqrt{2\pi} r \sigma_r} \exp\left(-\frac{(\log r - \mu_r)^2}{2\sigma_r^2} \right)$$

The mean *r* parameter satisfied the equation $\ln(\mu_r) = -0.693$. While there was uncertainty about an appropriate prior mean for *r*, setting the prior mean to have a CV of 100% allowed sufficient flexibility to estimate the probable value of *r* given the observed catch and CPUE data.

For the multiple-*r* production model, the prior distribution for intrinsic growth rate $p(r_T)$ was also a lognormal distribution with mean (μ_{r_T}) and variance $(\sigma_{r_T}^2)$ parameters set to achieve a CV for r_T of 100%.

$$\textbf{(10.2)} \quad p(r_T) = \frac{1}{\sqrt{2\pi} r_T \sigma_{r_T}} \exp\left(-\frac{(\log r_T - \mu_{r_T})^2}{2\sigma_{r_T}^2} \right)$$

Under the multiple-*r* model the mean r_T parameter was assumed to be normally distributed with a mean satisfying $\ln(\mu_{r_T}) = -0.693$ and a CV of 100%. That is, the hyperprior $p(\mu_{r_T})$ for the mean r_T parameter was $p(\mu_{r_T}) \sim N(\ln(r), \ln(r)^2)$ where *r*=0.5. While there was uncertainty about an appropriate prior mean for r_T, setting the prior mean to satisfy $\ln\left(\mu_{r_T}\right) = 0.5$ with a CV of 100% allowed sufficient flexibility to estimate the probable value of r_T given the observed catch and CPUE data.

Prior for Production Shape Parameter

The prior distribution for the production function shape parameter $p(M)$ was a gamma distribution with scale parameter λ and shape parameter k:

$$\textbf{(11)} \qquad p(M) = \frac{\lambda^k M^{k-1} \exp\left(-\lambda M\right)}{\Gamma\left(k\right)}$$

The values of the scale and shape parameters were set to $\lambda = k = 2$. This choice of parameters set the mean of $p(M)$ to be $\mu_M = 1$, which corresponded to the value of M for the Schaefer production model. This choice also implied that the CV of the shape parameter prior was 71%. In effect, the shape parameter prior was centered on the symmetric Schaefer model as the default with enough flexibility to estimate a nonsymmetrical production function if needed.

Prior for Catchability

The prior for the catchability coefficient $p(Q)$ was chosen to be a diffuse inverse-gamma distribution with scale parameter λ and shape parameter k.

$$\textbf{(12)} \qquad p(Q) = \frac{\lambda^k Q^{-(k+1)}}{\Gamma\left(k\right)} \exp\left(\frac{-\lambda}{Q}\right)$$

The scale and shape parameters were set to be $\lambda = k = 0.01$. This choice of parameters implied that $1/Q$ has a mean of 1 and a variance of 1000 which is a relatively noninformative prior. Since $1/Q$ is unbounded at $Q = 0$, an additional numerical constraint that Q be no smaller than 0.0001 was imposed for the Markov Chain Monte Carlo (MCMC) sampling.

Priors for Observation and Process Error Variances

Priors for the process error variance $p\left(\sigma^2\right)$ and observation error variance $p\left(\tau^2\right)$ were chosen to be moderately informative inverse-gamma distributions with scale parameter $\lambda>0$ and shape parameter $k>0$:

(13) $$p\left(\sigma^2\right) = \frac{\lambda^k \left(\sigma^2\right)^{-k-1} \exp\left(\frac{-\lambda}{\sigma^2}\right)}{\Gamma(k)} \quad and \quad p\left(\tau^2\right) = \frac{\lambda^k \left(\tau^2\right)^{-k-1} \exp\left(\frac{-\lambda}{\tau^2}\right)}{\Gamma(k)}$$

The inverse-gamma distribution is a useful choice for priors that describe model error variances (see, for example, Congdon, 2001). The scale parameter was set to $\lambda = 0.1$ and the shape parameter was $k = 0.2$ for the process error variance prior. For this choice of parameters, the expected value of the inverse-gamma distribution is not bounded, and we used the mode for σ^2, denoted as MODE[σ^2] = 1/12 ≈ 0.083 to measure the central tendency of the distribution. For the observation error variance prior, the scale parameter was set to $\lambda = 0.1$ and the shape parameter was $k = 0.2$. As a result, the mode of τ^2 was MODE[τ^2] = 1/12 ≈ 0.083. The ratio of the modes of the observation error prior to the process error prior was MODE[τ^2]/MODE[σ^2] = 1 and the central tendency of the observation error variance prior was assumed to be the same as in the process error variance prior. The choice of the process error prior matched the expected scaling of process errors which were on the order of 0.1 to 1 for the state equations describing changes in the proportion of carrying capacity. In summary, the prior for the observation error variance was assumed to be similar in magnitude to the process error variance.

Priors for Ratios of Biomass to Carrying Capacity

The prior distributions for the time series of the ratio of biomass to carrying capacity, $p(P_T)$, were determined by the lognormal distributions for the process error dynamics. Alternative mean values for the initial ratio of biomass to carrying capacity were evaluated using a goodness-of-fit criterion to select a best-fitting model for each island group (see Alternative Production Models below).

Posterior Distribution

The posterior distribution was calculated to make inferences about the model parameters given the data, the likelihood, and the priors. In particular, the joint posterior distribution given catch and CPUE data D, $p(\theta | D)$, for the single-*r* model with set of CPUE time series indexed by *S* was proportional to the product of the priors and the joint observation error likelihood

(14.1)

$$p(\theta | D) \propto p(K)\,p(r)\,p(M)\,p(\sigma^2)\prod_{T=1}^{N} p(P_T)\prod_{s \in S}\left(p(Q_s)\,p(\tau_s^2)\prod_{T=1}^{N} p(I_{s,T} | \theta)\right)$$

Similarly, under the multiple-*r* model the joint posterior distribution $p(\theta | D)$ with set of CPUE time series indexed by *S* was proportional to the product of the priors, hyperprior, and the joint observation error likelihood

(14.2)

$$p(\theta \mid D) \propto p(K)\,p(r_T)\,p(\mu_{r_T})\,p(M)\,p(\sigma^2)\prod_{T=1}^{N} p(P_T)\prod_{s\in S}\left(p(Q_s)\,p(\tau_s^2)\prod_{T=1}^{N} p(I_{s,T} \mid \theta)\right)$$

There were no analytical expressions to calculated parameter estimates from the posterior distributions and we used standard MCMC methods to numerically draw samples from the posterior distribution.

Bayesian parameter estimation for multi-parameter nonlinear models, such as the blue marlin production model, is typically based on simulating a set of independent samples from the posterior distribution. For the production model, we used Markov Chain Monte Carlo (MCMC) simulation (Gilks et al. 1996) to numerically generate a sequence of samples from the posterior distribution. The WINBUGS software (version 1.4, Spiegelhalter et al. 2003) was applied to set the initial conditions, perform the MCMC calculations, and summarize the MCMC results.

MCMC simulations were conducted in an identical manner for each of the alternative models described below. Three chains of 260,000 samples were simulated in each model run. The first 10,000 samples of each chain were excluded from the inference process. This burn-in period removed any dependence of the MCMC samples on the initial conditions. Each chain was also thinned by 25 to remove autocorrelation. That is, every twenty-fifth sample was used for inference. As a result, there were 30,000 samples from the posterior for summarizing model results. Convergence of the MCMC simulations to the posterior distribution was checked using the Brooks-Gelman-Rubin (BGR) convergence diagnostic (Gelman and Rubin 1992, Brooks and Gelman 1998). This diagnostic was monitored using WinBUGS for key model parameters (intrinsic growth rate, carrying capacity, catchability, initial ratio of biomass to carrying capacity, process and observation error variances) with values near unity indicating convergence. Convergence of the MCMC samples to the posterior distribution was also checked using the Geweke (1992) and Heidleberger and Welch (1992) diagnostics as implemented in the R language (R Development Core Team 2008) and the CODA package (Plummer et al. 2006).

Alternative Production Models

For the single-*R* and multiple-*R* growth rate models, alternative production models based on different hypotheses of which CPUE data provided an unbiased index of relative abundance of blue marlin were investigated. The alternative models were developed to contrast the effect of differing assumptions about the CPUE indices and about the best available information on blue marlin catch biomass (Table 1).

The goodness of fit of the alternative production models to the observed data was evaluated using the root-mean square error of fits to the observer CPUE indices, inspection of the log-scale standardized CPUE residuals, and the Deviance information criterion (*DIC*, Spiegelhalter et al. 2002), a Bayesian analog of the Akaike information criterion. In particular, the production model with the minimum *DIC* value was judged to provide the best fit to the data with the caveat that *DIC* differences of roughly two units of deviance indicated that there was no substantial difference between model fits and that differences of more than seven units were substantial (e.g., Spiegelhalter et al. 2002).

Model Diagnostics and Selection

CPUE residuals were also used to rank the goodness of fit of the alternative production models. Residuals for the CPUE series are the log-scale observation errors $\varepsilon_{T:}$

$$\text{(15)} \qquad \varepsilon_T = \ln(I_T) - \ln(QKP_T)$$

Non-random patterns in the residuals were an indication that the observed CPUE may not conform to one or more model assumptions. The root mean-squared error (RMSE) of the CPUE fit provided another diagnostic of the model goodness-of-fit with lower RMSE indicating a better fit when models with the same number of parameters were compared.

The relative goodness of fit to the observed CPUE and complexity of the alternative models was evaluated using the Deviance information criterion Spiegelhalter et al. (2002) statistic based on the Markov Chain Monte Carlo simulations. The *DIC* values for the alternative models were calculated as

$$\text{(16)} \qquad DIC = 2 \cdot \overline{D} - D(\theta) = \overline{D} + p_D$$

where $\overline{D}$ was the posterior mean of the model deviance, $D(\theta)$ was the value of deviance evaluated at the posterior mean of the stochastic variables in the model, and p_D was the effective degrees of freedom in the model. The production model with the minimum *DIC* value provided the best fit to the CPUE data accounting for model complexity. The difference between the *DIC* values of the j^{th} ranked model and the best fitting model (ΔDIC_j) was

$$\text{(17)} \qquad \Delta DIC_j = DIC_j - DIC_{MIN}$$

where DIC_j was the *DIC* of the j^{th} alternative model and DIC_{MIN} was the minimum *DIC* values of the best fitting model. As a rough guide, values of ΔDIC_j less than 2 indicate that the two models provide relatively similar fits to the CPUE data while ΔDIC_j values greater than 2 indicate a poorer fit to the CPUE data (Spiegelhalter et al. 2002).

Model Averaging

Model averaging was applied to account for model selection uncertainty. We used the Deviance Information Criterion to measure the goodness of fit of the alternative production models. The DIC value was calculated from information from the MCMC simulations for each model.

We now briefly digress to provide some background on the logical basis for DIC as a measure of goodness of fit, or conversely, the amount of discrepancy between the model and the data. In this context, the deviance $D(y,\theta)$ is defined -2 times the log-likelihood for the data y and the parameters θ

(18) $$D(y,\theta) = -2\log\left(L(y\,|\,\theta)\right)$$

The expected deviance, computed by averaging $D(y,\theta)$ over the true but unknown sampling distribution $f(y)$, equals 2 times the Kullback-Leibler information up to a constant that does not depend on the parameters θ. This relation to the Kullback-Leibler information implies that the parameters that produce the lowest expected deviance will produce the maximum information and have the highest posterior probability. However, we do not know the true sampling distribution and must therefore develop an accurate way to estimate the expected deviance. To this end, note that the discrepancy between the model and the data depends on both data y and parameters being estimated θ. To get a deviance measure that depends only on the data, one can define $\overline{D}(y)$ conditioned on a point estimate of θ, denoted by θ, as

(19) $$\overline{D}(y) = \overline{D}\left(y,\theta(y)\right)$$

where the point estimate was the mean of the posterior MCMC simulations for θ. That is,

(20) $$\theta(y) = E\left[\theta \mid y\right] = \frac{1}{J}\sum_{j=1}^{J}\theta^{(j)}$$

where $\theta^{(j)}$ is the jth iterate of θ in a total of J posterior simulations. This is one estimate of the expected deviance for a fixed point estimate of θ calculated from the posterior simulations.

Another more natural approach to estimating the expected deviance would be to use the average D of the estimate of model discrepancy over the posterior distribution as our estimate

(21) $$D(y) = E\left[D(y,\theta) \mid y\right]$$

Of course, we do not have complete knowledge or information about the true posterior distribution and so an analytic calculation of the integral for the expected value is not generally possible except for simple problems. However we can use the natural plug-in estimate the expected deviance by taking the average of $D(y,\theta)$ over the posterior simulations to be D, where

(22) $$D(y) = \frac{1}{J}\sum_{j=1}^{J}D\left(y,\theta^{(j)}\right)$$

Given this background, we observe that the difference between the posterior mean of the deviance (D) minus the deviance evaluated at the posterior mean of the stochastic nodes

($\overline{D}$) provides a measure of the effect of model fitting and can be used as a measure of the effective degrees of freedom in the model (ρ_D)

(23) $$\rho_D = D - \overline{D}$$

The value of DIC was then calculated as twice the posterior mean of the deviance minus the deviance evaluated at the posterior mean of the stochastic nodes

(24) $$DIC = 2D - \overline{D} = D + \rho_D = \overline{D} + 2\rho_D$$

The alternative models, indexed by k (M_k), were ranked by their DIC values (DIC_k). The best fitting model (M^*) with the minimum value of DIC (DIC_{Min}) produced the best fit to the observed data. For each model, the difference in the model's value of DIC from the minimum was calculated (Δ_k) as

(25) $$\Delta_k = DIC_k - DIC_{Min}$$

Models with values of Δ_k less than 2 were selected as candidate models that provided a similar goodness of fit to the observed data as the best fitting production model. In this context, evidence indicated that the candidate models also provided adequate fits to the observed data and should be considered as viable alternative states of nature in comparison to the best fitting model. The results from the set of m total candidate models, denoted by $\underline{M}$, were model-averaged based on the likelihood $L_k\left(y \mid \theta\right)$ of each model $M_k \in \underline{M}$. In this context, the model likelihood was proportional to the exponential of the Δ_k value through the expression

(26) $$L_k\left(y \mid \theta\right) \propto \exp\left(-0.5 \cdot \Delta_k\right)$$

We applied Bayesian model averaging (BMA) to the set of candidate models. This produced a model-averaged set of results which were used for inference about stock status determination and population projections. To apply BMA, prior model probabilities (π_k) were developed to express the relative belief that the candidate models represented the true state of nature. We adopted an objective Bayesian approach for setting the prior model probabilities in the absence of any information for preferring one candidate model over another. Thus, we set the prior model probabilities to be equal based on the principle of indifference, where for all candidate models j and k

(27)

$$\pi_j = \pi_k = \frac{1}{m} \text{ and } \sum_j \pi_j = 1$$

Given the integrated likelihoods of the candidate models and the prior model probabilities, we calculated the posterior model probabilities (W_k) over the set of candidate models as

(28)

$$W_k = \frac{\pi_k \cdot \exp(-0.5\Delta_k)}{\sum_i \pi_i \cdot \exp(-0.5\Delta_i)}$$

The posterior model probability quantified the relative support contained in the sample data and the assumed prior model probabilities. Because we did not have any evidence to assume one model was a priori more likely than another, the posterior model probabilities were in effect based on the observed data. In this context, these BMA analyses represented an objective Bayesian approach.

The posterior model probabilities provide the essential model weighting information for BMA of any production model parameter θ. The model-averaged estimate of the parameter θ is denoted as $\overline{\theta}$ and its value depends on the estimates θ_k from each candidate model indexed by k. The expected value of the model-averaged estimate is

(29) $$E\left[\bar{\theta}\right] = \sum_k W_k \theta_k$$

and the variance of the model-averaged estimate is

(30) $$VAR\left[\bar{\theta}\right] = \left[\sum_k W_k \sqrt{VAR[\theta_k] + \left(\theta_k - E\left[\bar{\theta}\right]\right)^2}\,\right]^2$$

The variance of the model-averaged estimate includes two components, the first is the variance of the individual model estimates and the second is an expression for the variance contribution of model uncertainty in the point estimate of the parameters θ.

RESULTS

CPUE Index Correlation Analysis

The six CPUE indices (Japan 1975-1993, Japan 1993-2011, Hawaii 1995-2011, Taiwan 1967-1978, Taiwan 1979-1999, Taiwan 2000-2011) were examined for correlations. All CPUE indices were weakly or moderately positively correlated and had Pearson correlations ranging from $\rho = 0.15$ to $\rho = 0.46$, with the exception of the pairs of Hawaii 1995-2011 and Taiwan 1979-1999 ($\rho = -0.48$) and of Hawaii 1995-2011 and Taiwan 2000-2011 ($\rho = -0.27$) which were negatively correlated.

Model Convergence Diagnostics

Convergence diagnostics were calculated from the three chains used in the MCMC simulations for all models. The diagnostics were computed for the key model parameters: BMSY (exploitable biomass to produce MSY), HMSY (exploitation rate to produce MSY), MSY, *K*, *r* (single-*r* models), $\bar{r}$ average (for multiple-*r* models), P[1], and *Q*s. The Geweke Z-score diagnostic values were less than 2 in absolute value for the vast majority of the models parameters which indicated that there were no significant differences in means for the first and last sets of iterations of the chains. The Gelman and Rubin potential scale reduction factors were identically 1 for each of the models parameters which also indicated convergence in distribution of the MCMC samples to the joint posterior distributions. Last, each of the models parameters passed the Heidelberger and Welch stationary and half-width diagnostic tests with very few exceptions. Overall, the convergence results indicated that the MCMC chains produced representative samples from the joint posterior distribution of all production models parameters.

Model Fits to CPUE and Residual Tests

The predicted CPUE indices for each model were compared to the observed CPUE to determine model fit. Specifically, the standardized log residuals from the CPUE fit were visually examined for time trends and regression of the standardized log residuals on time were used to test that the residuals were normally distributed with constant variance. Several patterns were immediately apparent (Table 2):

- Models with just Japan and just Hawaii CPUE indices had at least one index that failed normality tests regardless of the catch time series (1950 or 1971).
- Taiwan CPUE 5 1979-1999 failed the normality tests in multiple-*r* models when just Taiwan CPUE indices were included.
- The Hawaii CPUE index had clear temporal patterning (positive residuals from 1995-2001 and negative residuals from 1996-2011) in all multiple- *r* models which contained Hawaii and another country's CPUE index. The same was true for single-*r* models except for the model containing Hawaii and Taiwan CPUE indices.

- The Taiwan 1967-1978 CPUE index had negative residuals in almost all years in all multiple-*r* models and clear non-random temporal patterning in all single-*r* models.

Model Selection

Several criteria were used to identify a final set of candidate models (Table 2). First, key model estimates (BMSY, HMSY, BSTATUS_2011, K, *r*) and residual diagnostic results were similar between models which differed in only the catch time series (1950-2011, 1971-2011) therefore all models with the truncated catch time series (1971-2011) were eliminated from contention. Second, models in which over half of the CPUE time series standardized log residuals contained obvious time trends or were non-normally distributed were also eliminated (Table 2). This left only 5 candidate models out of the 36 initial set (Table 1):

(1) mr_1950_all,

(2) mr_1950_Japan_Taiwan

(3) sr_1950_Japan_Taiwan

(4) sr_1950_Japan_Taiwan_5_6

(5) sr_1950_Japan_Taiwan_5_6

Of these five models, the mr_1950_Japan_Taiwan_5_6 and sr_1950_Japan_Taiwan_5_6 CPUE indices had the best fits to the observed CPUE indices as evidenced by a lack of residual time trends for any of the CPUE indices (Figure 3) and the most number of CPUE indices residuals being normally distributed (none for the multiple-*r* model, 1 for the single-*r* model) (Table 2). Examination of the individual CPUE indices root mean-square error also supported selection of these two models due to their relatively lower RMSE values (Table 2).

The final selection criteria utilized the *DIC* estimates (Table 2). This was possible because the two remaining candidate models contained the same catch data and CPUE indices. Results indicated that the minimum value of *DIC* equal to -96.25 was achieved for the single r model. However, the *DIC* of the multiple-*r* model was -94.18, which is a ΔDIC of 2.07 indicating that model averaging was appropriate.

Parameter estimates of *K*, *r* for single-*r* model and average *r* for multiple-*r* model, and *M* from the best-fitting model (sr_1950_Japan_Taiwan_5_6) and alternative model (mr_1950_Japan_Taiwan_5_6) were examined for correlations. Negative correlations were found amongst all combinations of parameters for both models (Tables 3, 4).

P[1] Prior Sensitivity Analysis

A sensitivity analysis of the P[1] prior was conducted to understand its influence on the best-fitting (sr_1950_Japan_Taiwan_5_6) and alternative model (mr_1950_Japan_Taiwan_5_6) exploitable biomass estimates. In addition to the initial P[1] prior of 0.5 (CV=1.0), P[1] prior values of 0.25 and 0.75, both with CV = 1.0, were examined.

The sensitivity analysis of the P[1] prior revealed a minimal influence on the exploitable biomass estimates. The biomass estimates in 1950 were lowest when the P[1] prior = 0.25, intermediate the P[1] prior = 0.5, and greatest with a P[1] prior = 0.75 for both the single r and multiple r models (Fig. 3). However, the biomass estimates of the models with the different P[1] priors quickly converged to a single point by 1961.

Posterior Estimates of Model Parameters, Exploitable Biomass, Exploitation Rate and Reference Points

The production model shape parameter from the two credible models suggested M ranged from 0.94 to 1.25 with a model average value of 1.17. Carrying capacity estimates indicated that K ranged from 73.8 to 118.1 thousand metric tons with a model averaged value of 85.43. The posterior means for intrinsic growth rate suggested that the estimate of *r* was 1.33 for the single-*r* model and the average *r* for the multiple-*r* model was $\bar{r}=1.13$ with a model averaged value of 1.28. Estimates of the initial ratio of biomass to carrying capacity in 1950 were between 0.44 and 0.48 with a model average value of 0.45. Model averaged mean estimates of biological reference points were: BMSY = 43.03 (SE =13.62), HMSY = 0.59 (SE =0.19), and MSY = 23.04 (SE =3.57).

Estimates of blue marlin exploitable biomass have fluctuated around 60,000 mt since 1950 (Table 5, Figures 5 and 7). Biomass fluctuated around 65,000 mt from 1953 to 1995 and then abruptly declined and fluctuated around 50,000 mt through 2011. Estimates of exploitation rate rose steadily from 0.20 in 1952 to a high of 0.57 in 2003 (Table 5, Figures 6 and 8). Since then the exploitation rate has steadily declined to 0.38 in 2011 (Table 5, Figure 6).

Estimates of relative biomass indicate that the initial biomass of blue marlin was likely above BMSY, however after the abrupt decline in biomass in 1995 the lower bounds of the model averaged unconditional standard error suggest that biomass may have dropped below BMSY and remain there today (Table 5, Figures 7, 8, and 9). Relative biomass B/BMSY was closest to 1 in 2002 (1.05) but has since increased to 1.18 in 2011.

Similarly, estimates of relative exploitation rate indicate that the annual harvest rate was increasing but below HMSY until 2002 when it was slightly above 1 (1.05) (Figure 8). The upper bounds of the model averaged unconditional standard error suggest that the relative harvest rate was above HMSY 2001-2007 but have since lowered to below HMSY.

The biomass status of blue marlin in 2011 suggests that the biomass was above BMSY based on the model averaged values but the unconditional standard error shows there is a small chance that it may be slightly below BMSY (Table 5, Figures 5, 7, and 9). However, the best point estimate, or mean of the model averaged estimates of harvest rate and its standard error in 2011 show that it is very likely that the harvest rate did not exceed the overfishing threshold (Table 5, Figures 6, 8, and 9). Overall, the production model results suggest that the blue marlin is likely not overfished and did not experience overfishing during 1950-2011 with the possible exception of 2000-2006.

ACKNOWLEDGMENTS

We thank the members of the ISC Billfish Working Group for their help and diligent efforts to prepare the information needed for this assessment. We also thank Lennon Thomas for her assistance with the graphics.

REFERENCES

Brodziak, J., Holmes, E., Sosebee, K., and Mayo, R. 2001. Assessment of the silver hake resource in the northwest Atlantic in 2000. *NEFSC Ref. Doc. 01-03*, 134 p. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0103/

Brodziak, J., Courtney, D., Wagatsuma, L., O'Malley, J., Lee, H-.H., Walsh, W., Andrews, A., Humphreys, R., and DiNardo, G. 2011. Stock assessment of the main Hawaiian Islands Deep7 bottomfish complex through 2010. U.S. Dep. Commer., *NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-29*, 176 p. + Appendix. Available at: http://www.pifsc.noaa.gov/library/pubs/tech/NOAA_Tech_Memo_PIFSC_29.pdf

Brooks, S. P., and Gelman, A. 1998. Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7:434-455.

Gelman, A., and Rubin, D. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7: 457-511.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics*. Edited by J. Bernardo, J. Berger, A. Dawid, and A. Smith. Vol. 4, Claredon Press, Oxford, U.K. pp. 169-194.

Gilks, W. R., Richardson, S., and Spiegelhalter, D.J. [Eds.] 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London. 486 pp.

Heidelberger, P. and Welch, P. 1992. Simulation run length control in the presence of an initial transient. *Op. Res.* 31: 1109-1144.

Meyer, R., and Millar, R. BUGS in Bayesian stock assessments. 1999. *Can. J. Fish. Aquat. Sci.* 56:1078–1086.

Plummer, M., Best, N., Cowles, K., and Vines, K. 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*. 6: 7-11.

R Development Core Team. 2008. R: a language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. 2002. Bayesian measures of model complexity and fit. *J. R. Statist. Soc. B*, 64:583-639.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. 2003. *WinBUGS User Manual*. Available at: http://www.mrc.bsu.carn.ac.uk/bugs/winbugs/manual14.pdf

Table 1. Alternative Pacific blue marlin Bayesian production models. The intrinsic growth rate of the population, 'r', varied every year ('mr') or was a constant throughout time ('r'). The starting year varied between 1950 and 1971, and included CPUE indices also varied.

Model name	**r**	**Catch**	**CPUE****
mr_1950_all	1 per year	1950-2011	1,2,3,4,5,6
mr_1950_HI	1 per year	1950-2011	3
mr_1950_HI_Taiwan	1 per year	1950-2011	3,4,5,6
mr_1950_Japan	1 per year	1950-2011	1,2
mr_1950_Japan_HI	1 per year	1950-2011	1,2,3
mr_1950_Japan_Taiwan	1 per year	1950-2011	1,2,4,5,6
mr_1950_Japan_Taiwan_5_6	1 per year	1950-2011	1,2,5,6
mr_1950_Taiwan	1 per year	1950-2011	4,5,6
mr_1950_Taiwan_5_6	1 per year	1950-2011	5,6
mr_1950_weightedCPUE*	1 per year	1950-2011	1,2,3,4,5,6
sr_1950_all	1 total	1950-2011	1,2,3,4,5,6
sr_1950_HI	1 total	1950-2011	3
sr_1950_HI_Taiwan	1 total	1950-2011	3,4,5,6
sr_1950_Japan	1 total	1950-2011	1,2
sr_1950_Japan_HI	1 total	1950-2011	1,2,3
sr_1950_Japan_Taiwan	1 total	1950-2011	1,2,4,5,6
sr_1950_Japan_Taiwan_5_6	1 total	1950-2011	1,2,5,6
sr_1950_Taiwan	1 total	1950-2011	4,5,6
sr_1950_Taiwan_5_6	1 total	1950-2011	5,6
sr_1950_weightedCPUE*	1 total	1950-2011	1,2,3,4,5,6
mr_1971_all	1 per year	1971-2011	1,2,3,4,5,6
mr_1971_HI	1 per year	1971-2011	3
mr_1971_HI_Taiwan	1 per year	1971-2011	3,4,5,6
mr_1971_Japan	1 per year	1971-2011	1,2
mr_1971_Japan_HI	1 per year	1971-2011	1,2,3
mr_1971_Japan_Taiwan	1 per year	1971-2011	1,2,4,5,6
mr_1971_Taiwan	1 per year	1971-2011	4,5,6
mr_1971_Taiwan_5_6	1 per year	1971-2011	5,6
mr_1971_weightedCPUE*	1 per year	1971-2011	1,2,3,4,5,6
sr_1971_all	1 total	1971-2011	1,2,3,4,5,6
sr_1971_HI	1 total	1971-2011	3
sr_1971_HI_Taiwan	1 total	1971-2011	3,4,5,6

* CPUE weighted by each country's proportion of total catch

** Key to CPUE longline time series:

1	Japan	1975-1993	4	Taiwan	1967-1978
2	Japan	1994-2011	5	Taiwan	1979-1999
3	USA (HI)	1995-2011	6	Taiwan	2000-2011

Table 1. Continued.

sr_1971_Japan	1 total	1971-2011	1,2
sr_1971_Japan_HI	1 total	1971-2011	1,2,3
sr_1971_Japan_Taiwan	1 total	1971-2011	1,2,4,5,6
sr_1971_Japan_Taiwan_5_6	1 total	1971-2011	1,2,5,6
sr_1971_Taiwan	1 total	1971-2011	4,5,6
sr_1971_Taiwan_5_6	1 total	1971-2011	5,6
sr_1971_weightedCPUE*	1 total	1971-2011	1,2,3,4,5,6

****** CPUE weighted by each country's proportion of total catch

******* Key to CPUE longline time series:

1	Japan	1975-1993	4	Taiwan	1967-1978
2	Japan	1994-2011	5	Taiwan	1979-1999
3	USA (HI)	1995-2011	6	Taiwan	2000-2011

Table 2. Diagnostics used in final model selection. mr represents multiple-*r* models, sr represents models with a single-*r*. Indices with time trends are the CPUE indices with obvious time trends apparent via gross visual examination. Indices with non-normality are the CPUE indices with Shapiro–Wilk test P-value <0.05. *DIC* is the deviance information criteria.

Model	Indices with time trends	Indices with non-normality	Root Mean Square Error						*DIC*
			S1	S2	S3	S4	S5	S6	
mr_1950_all	S3,S4		0.08	0.65	0.15	0.01	0.02	0.05	-167.83
mr_1950_HI		S3			0.11				-27.72
mr_1950_HI_Taiwan	S3,S4,S6	S5			0.12	0.01	0.03	0.08	-181.21
mr_1950_Japan		S1	0.09	0.63					24.27
mr_1950_Japan_HI	S2,S3	S1	0.09	0.78	0.14				12.92
mr_1950_Japan_Taiwan	S4		0.08	0.62		0.01	0.02	0.05	-157.57
mr_1950_Japan_Taiwan_5_6	S5		0.09	0.60			0.02	0.05	-94.18
mr_1950_Taiwan	S4	S5				0.01	0.03	0.06	-160.99
mr_1950_Taiwan_5_6		S5					0.03	0.06	-102.67
mr_1950_weightedCPUE	S1, S3,S4	S6	0.08	0.40	0.01	0.01	0.01	0.04	-363.30
sr_1950_all	S1,S4	S5	0.11	0.61	0.16	0.01	0.02	0.05	-169.67
sr_1950_HI		S3			0.11				-27.85
sr_1950_HI_Taiwan	S3,S4,S6	S5			0.13	0.01	0.03	0.08	-191.23
sr_1950_Japan		S1	0.09	0.61					23.25
sr_1950_Japan_HI	S2,S3	S1	0.09	0.74	0.14				11.19
sr_1950_Japan_Taiwan	S1,S4	S5	0.11	0.59		0.01	0.02	0.05	-159.61
sr_1950_Japan_Taiwan_5_6		S5	0.10	0.57			0.02	0.05	-96.25
sr_1950_Taiwan	S4					0.01	0.02	0.06	-175.97
sr_1950_Taiwan_5_6							0.02	0.05	-110.50
sr_1950_weightedCPUE	S3,S4	S5	0.08	0.49	0.01	0.01	0.01	0.04	-370.55

Table 2. Continued.

Model	Indices with time trends	Indices with non-normality	Root Mean Square Error S1	S2	S3	S4	S5	S6	*DIC*
mr_1971_all	S3, S4		0.08	0.65	0.15	0.02	0.02	0.05	182.34
mr_1971_HI		S3			0.11				302.52
mr_1971_HI_Taiwan	S3,S4,S6	S5			0.11	0.02	0.03	0.08	194.37
mr_1971_Japan		S1	0.08	0.63					353.82
mr_1971_Japan_HI	S2, S3	S1	0.08	0.78	0.14				341.67
mr_1971_Japan_Taiwan	S4		0.08	0.62		0.02	0.02	0.05	194.25
mr_1971_Japan_Taiwan_5_6			0.09	0.60			0.02	0.05	233.43
mr_1971_Taiwan	S4					0.02	0.03	0.06	193.61
mr_1971_Taiwan_5_6		S5					0.03	0.06	228.02
mr_1971_weightedCPUE	S1,S3,S4	S6	0.08	0.41	0.01	0.07	0.01	0.04	0.00
sr_1971_all	S3,S4		0.08	0.61	0.15	0.02	0.02	0.05	185.11
sr_1971_HI		S3			0.11				302.34
sr_1971_HI_Taiwan	S3,S4,S6	S5			0.12	0.02	0.03	0.08	170.27
sr_1971_Japan		S1	0.08	0.60					351.21
sr_1971_Japan_HI	S2,S3	S1	0.08	0.73	0.14				338.56
sr_1971_Japan_Taiwan	S4		0.08	0.58		0.02	0.02	0.05	194.59
sr_1971_Japan_Taiwan_5_6		S5	0.09	0.57			0.02	0.05	231.06
sr_1971_Taiwan	S4					0.02	0.02	0.06	185.95
sr_1971_Taiwan_5_6							0.02	0.05	220.16
sr_1971_weightedCPUE	S3,S4	S5	0.08	0.50	0.01	0.04	0.01	0.04	1.10

Table 3. Correlation matrix of parameters *K*, *r* and *M* for best fitting model (sr_1950_Japan_Taiwan_5_6).

	K	*r*	*M*
K		-0.405	-0.128
r	-0.405		-0.612
M	-0.128	-0.612	

Table 4. Correlation matrix of parameters *K*, *r* and *M* for alternative model (mr_1950_Japan_Taiwan_5_6).

	K	$\bar{r}$	*M*
K		-0.273	-0.112
$\bar{r}$	-0.273		-0.588
M	-0.112	-0.588	

Table 5. Model averaged estimates of biomass, exploitation rate, relative biomass and relative exploitation rate.

Year	Exploitable Biomass (B)		Exploitation Rate (H)		Relative Biomass (B/BMSY)		Relative Exploitation (H/HMSY)	
	B Mean	SE	H Mean	SE	Bstatus Mean	SE	Hstatus Mean	SE
1950	39.22	35.89	0.01	0.01	0.90	0.73	0.01	0.01
1951	51.94	31.64	0.01	0.00	1.23	0.65	0.01	0.01
1952	65.13	30.30	0.16	0.10	1.56	0.61	0.30	0.20
1953	65.45	29.11	0.18	0.11	1.54	0.52	0.32	0.19
1954	67.64	28.30	0.13	0.08	1.59	0.48	0.24	0.13
1955	71.17	28.27	0.14	0.07	1.68	0.46	0.25	0.12
1956	71.51	28.66	0.13	0.07	1.68	0.45	0.22	0.10
1957	73.05	28.85	0.19	0.09	1.71	0.44	0.33	0.13
1958	69.66	28.80	0.21	0.10	1.62	0.42	0.38	0.14
1959	68.82	28.00	0.21	0.09	1.60	0.40	0.37	0.13
1960	68.86	27.72	0.19	0.08	1.60	0.40	0.33	0.11
1961	70.14	27.56	0.28	0.11	1.64	0.40	0.49	0.16
1962	64.82	27.19	0.34	0.14	1.50	0.37	0.60	0.19
1963	61.57	26.27	0.38	0.16	1.43	0.37	0.68	0.23
1964	58.56	25.51	0.35	0.15	1.36	0.37	0.61	0.22
1965	59.68	25.30	0.28	0.12	1.39	0.39	0.49	0.20
1966	62.60	25.38	0.26	0.11	1.47	0.41	0.46	0.20
1967	63.66	25.42	0.23	0.10	1.50	0.41	0.41	0.18
1968	65.30	25.45	0.23	0.10	1.54	0.42	0.42	0.18
1969	64.99	25.46	0.24	0.10	1.53	0.41	0.43	0.19
1970	64.77	25.31	0.27	0.11	1.53	0.41	0.48	0.21
1971	62.78	24.85	0.17	0.08	1.48	0.41	0.31	0.15
1972	67.45	24.85	0.18	0.08	1.61	0.44	0.34	0.16
1973	65.58	25.04	0.22	0.11	1.56	0.45	0.40	0.19
1974	62.48	23.89	0.22	0.10	1.49	0.46	0.39	0.18
1975	54.97	19.26	0.22	0.09	1.30	0.34	0.40	0.15
1976	55.67	18.38	0.23	0.08	1.33	0.35	0.42	0.16
1977	54.43	18.22	0.25	0.09	1.30	0.32	0.46	0.16
1978	61.94	20.79	0.26	0.09	1.46	0.30	0.47	0.14
1979	71.89	23.69	0.24	0.08	1.68	0.27	0.42	0.11
1980	66.56	22.54	0.27	0.09	1.55	0.25	0.47	0.12
1981	69.25	23.13	0.27	0.09	1.62	0.25	0.47	0.12
1982	67.12	22.90	0.29	0.10	1.56	0.25	0.52	0.13
1983	64.27	22.55	0.30	0.10	1.49	0.25	0.52	0.13
1984	71.99	25.36	0.31	0.11	1.67	0.29	0.54	0.14

Table 3. Continued.

Year	Exploitable Biomass (B)		Exploitation Rate (H)		Relative Biomass (B/BMSY)		Relative Exploitation (H/HMSY)	
	B Mean	SE	H Mean	SE	Bstatus Mean	SE	Hstatus Mean	SE
1985	66.29	23.33	0.26	0.09	1.54	0.26	0.46	0.12
1986	65.23	22.20	0.31	0.10	1.52	0.26	0.54	0.14
1987	62.74	21.45	0.44	0.15	1.46	0.24	0.78	0.20
1988	58.93	21.35	0.41	0.15	1.37	0.24	0.71	0.18
1989	60.52	21.48	0.34	0.12	1.41	0.25	0.60	0.16
1990	60.72	20.98	0.30	0.10	1.42	0.25	0.53	0.14
1991	62.82	21.28	0.31	0.11	1.47	0.26	0.55	0.15
1992	61.38	21.13	0.37	0.13	1.43	0.25	0.65	0.17
1993	66.75	23.24	0.37	0.13	1.55	0.26	0.65	0.17
1994	63.77	21.75	0.39	0.13	1.49	0.24	0.69	0.18
1995	63.66	21.92	0.41	0.14	1.49	0.29	0.73	0.21
1996	48.10	16.93	0.36	0.13	1.12	0.21	0.63	0.18
1997	55.11	18.52	0.35	0.12	1.30	0.27	0.63	0.19
1998	53.33	17.74	0.37	0.12	1.26	0.26	0.67	0.21
1999	49.12	16.19	0.38	0.12	1.16	0.23	0.67	0.20
2000	49.25	16.04	0.45	0.15	1.17	0.26	0.82	0.26
2001	46.23	14.91	0.54	0.17	1.09	0.22	0.98	0.29
2002	44.31	14.42	0.57	0.17	1.05	0.21	1.02	0.30
2003	49.53	16.62	0.57	0.18	1.16	0.23	1.02	0.30
2004	51.16	17.89	0.49	0.17	1.20	0.24	0.87	0.25
2005	55.67	19.43	0.48	0.17	1.30	0.27	0.86	0.26
2006	51.11	18.12	0.47	0.16	1.20	0.25	0.83	0.25
2007	48.55	17.12	0.43	0.15	1.14	0.25	0.77	0.24
2008	48.46	16.77	0.41	0.14	1.14	0.26	0.74	0.24
2009	51.16	17.87	0.40	0.14	1.21	0.29	0.72	0.24
2010	53.19	18.15	0.41	0.14	1.26	0.29	0.74	0.25
2011	49.84	17.12	0.38	0.13	1.18	0.27	0.68	0.24
average	60.35	22.80	0.30	0.11	1.42	0.34	0.54	0.18
recent average (2007-2011)	50.24	17.40	0.41	0.14	1.19	0.27	0.73	0.24

Figure 1.1. Pacific blue marlin catch biomass (mt), 1950-2011.

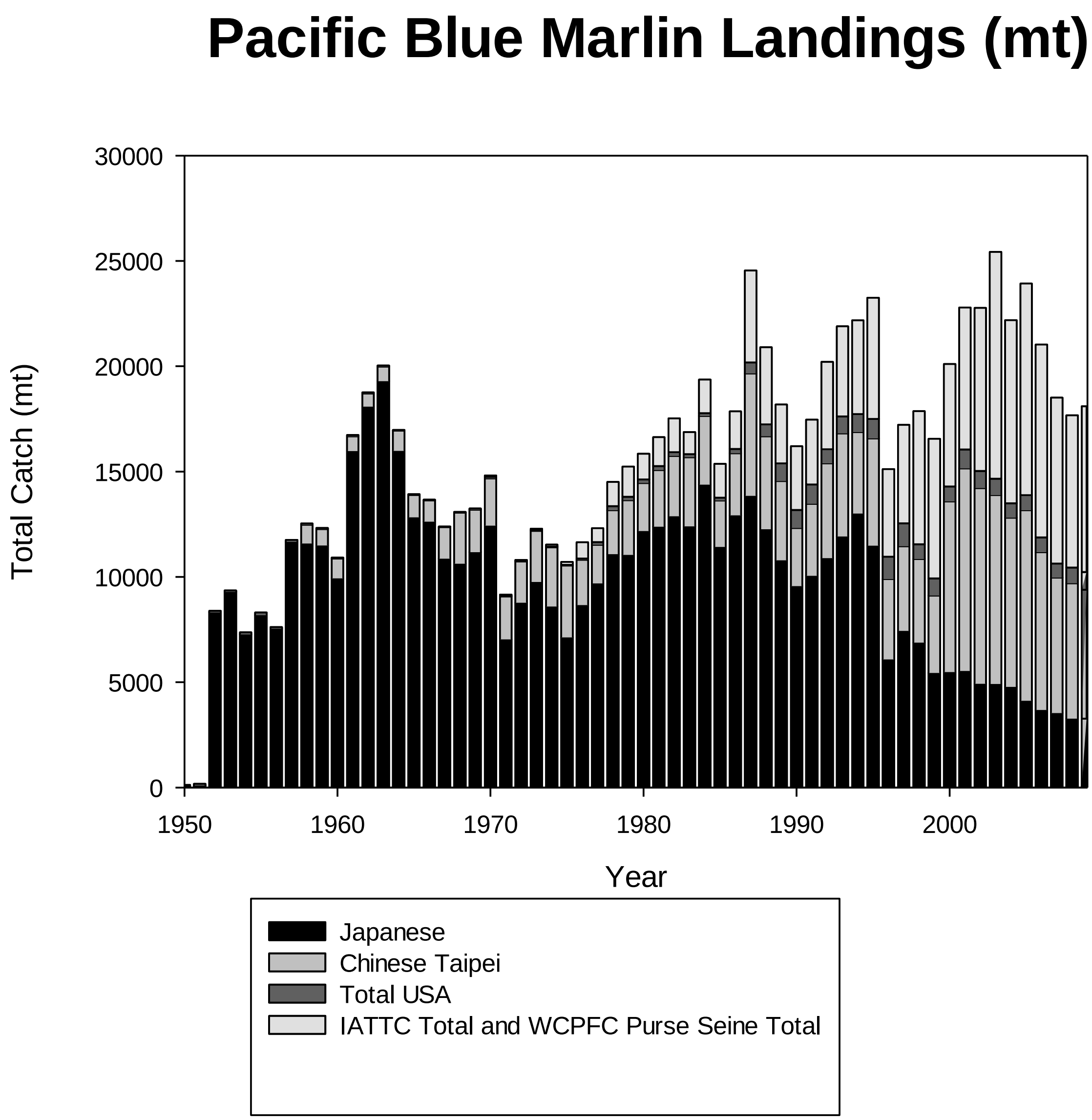


Figure 1.2. Annual proportion of Pacific blue marlin catch biomass by country, 1950-2011.

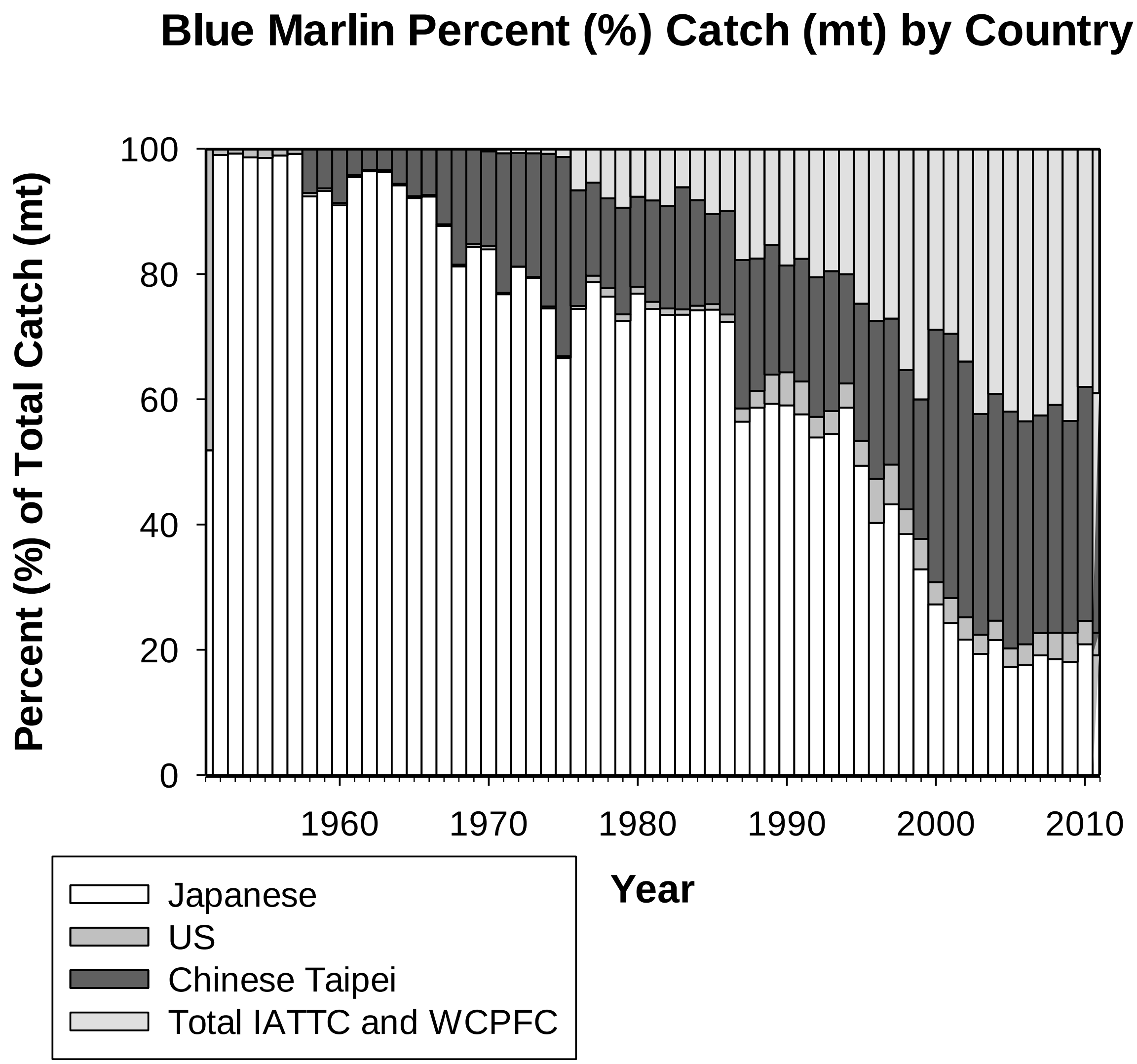


Figure 2.1. Standardized Japanese longline fishery CPUE, 1975-1993 and 1994-2011.

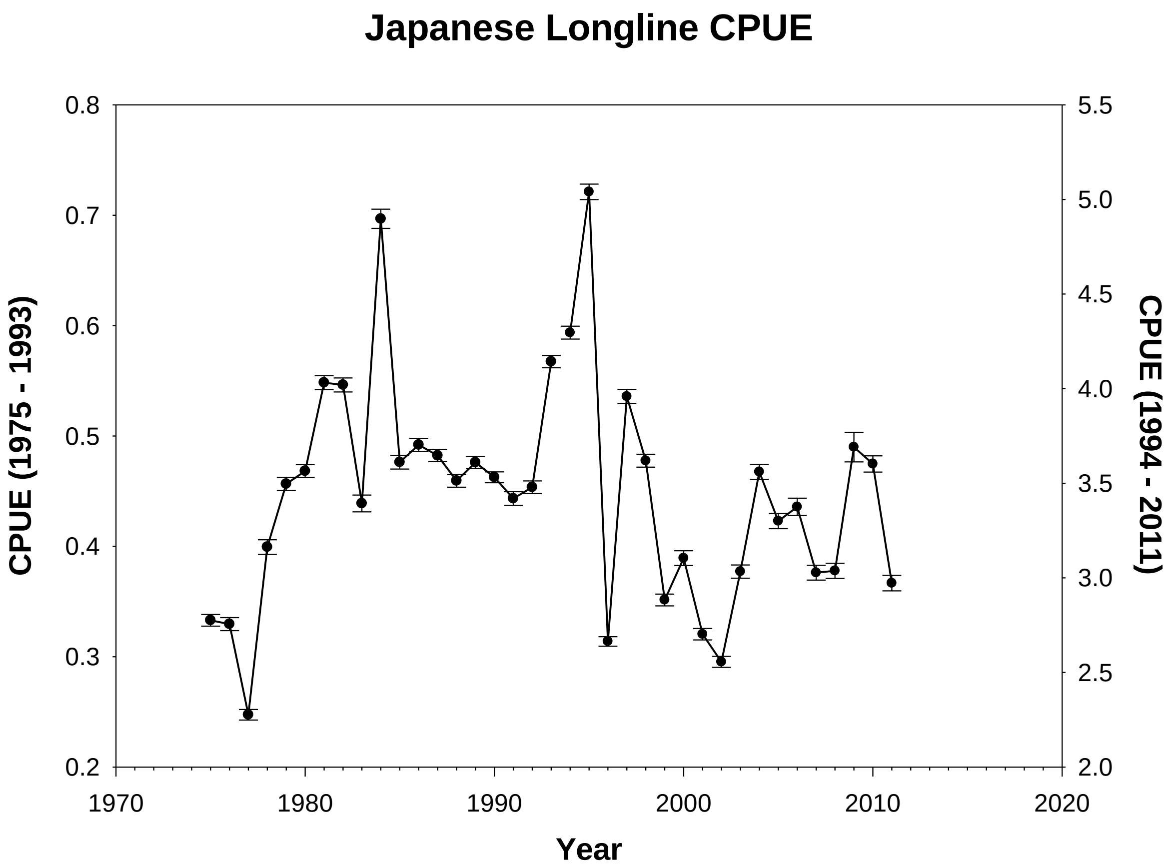


Figure 2.2. Hawaii-based longline fishery CPUE, 1994-2011.

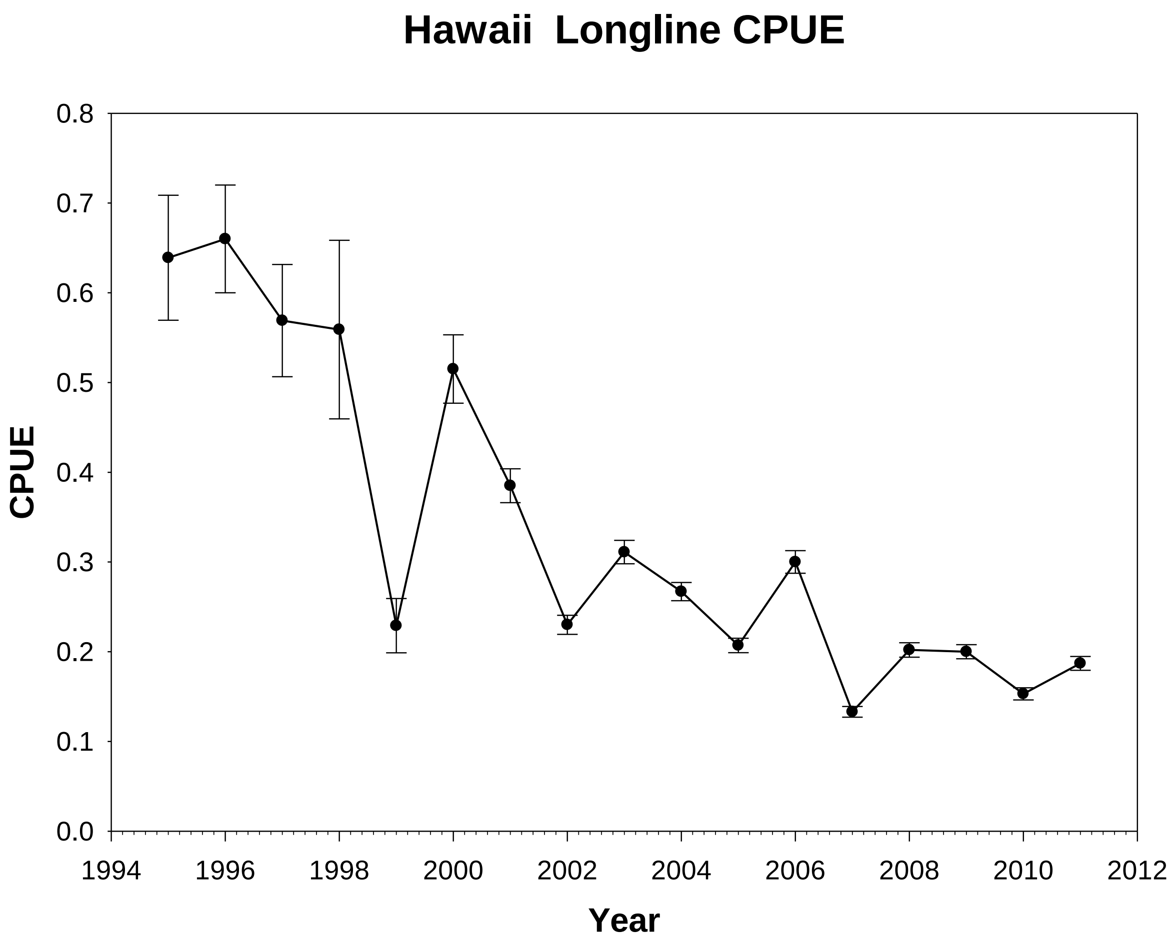


Figure 2.3. Chinese Taipei longline fishery CPUE, 1994-2011.

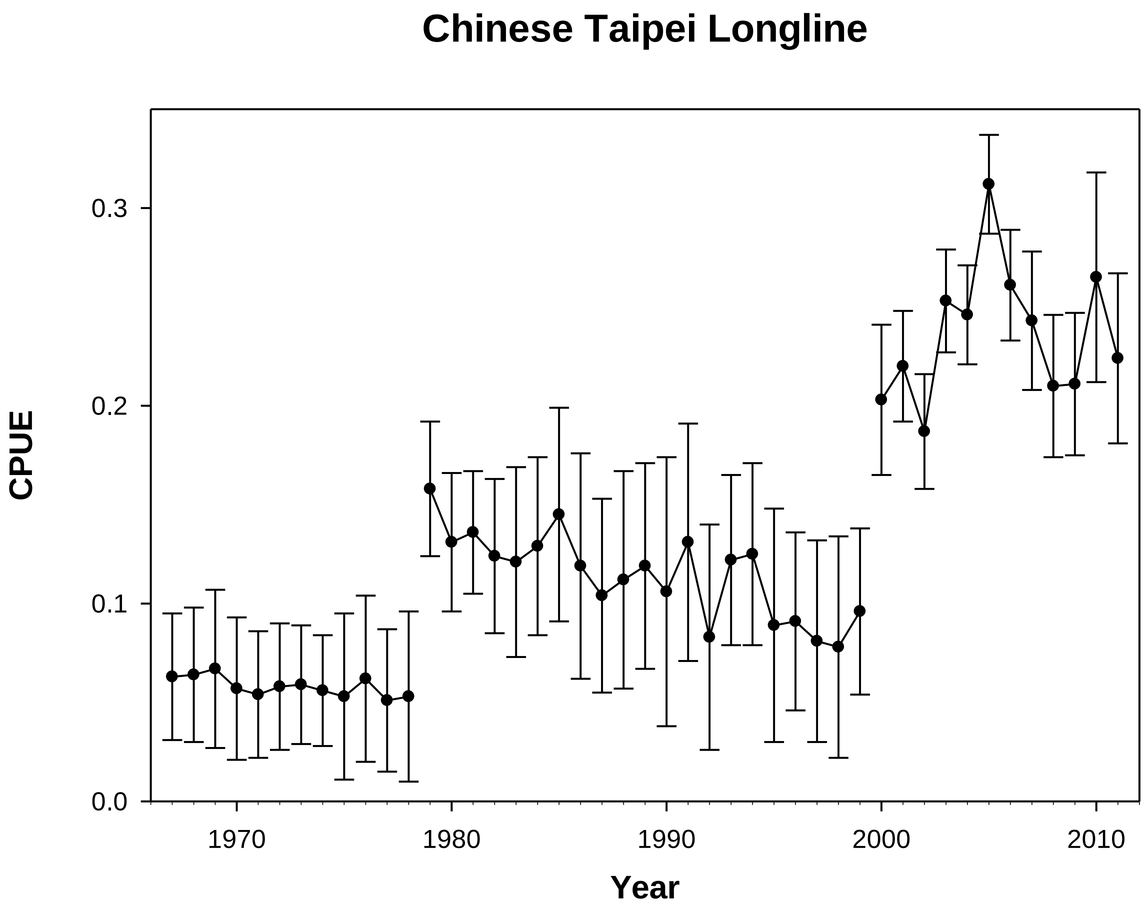


Figure 3. Exploitable biomass estimates for the best-fitting and alternative production models with different P[1] priors.

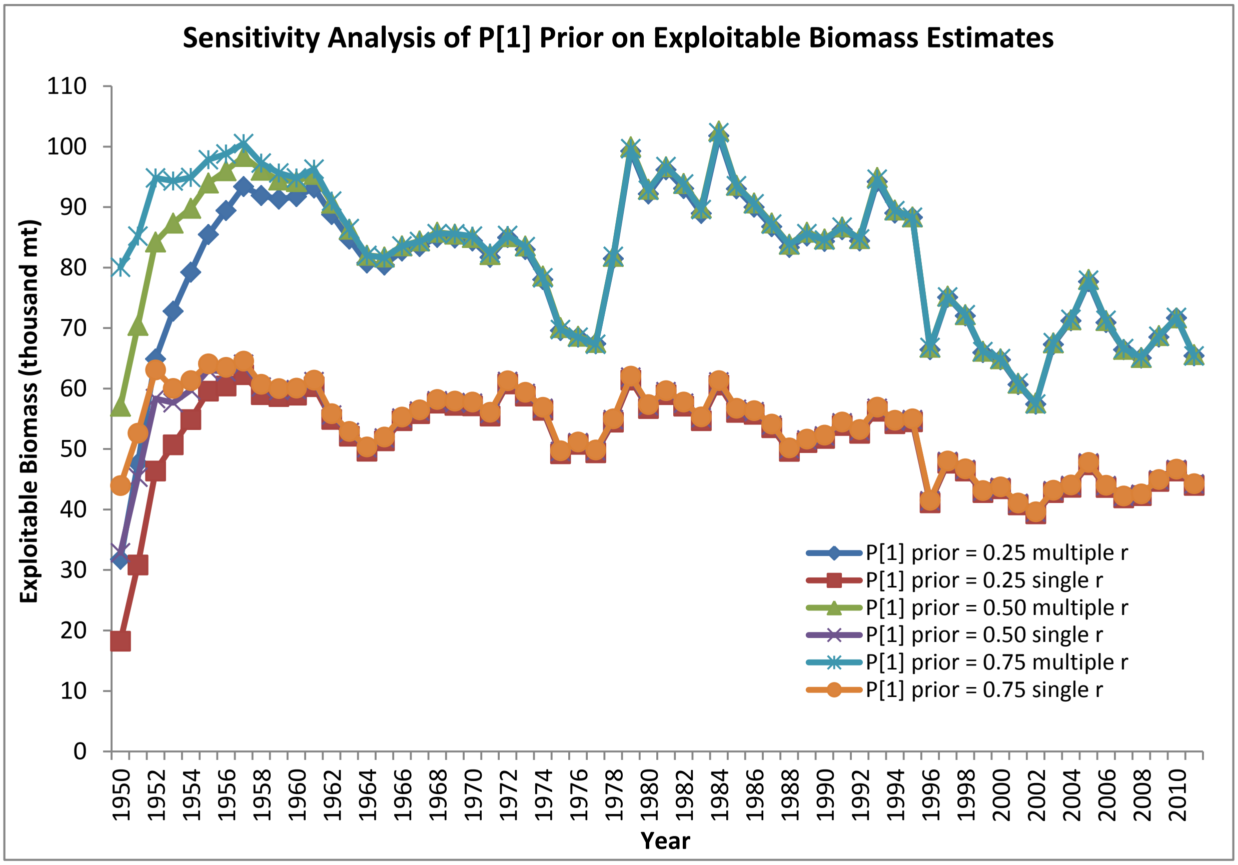
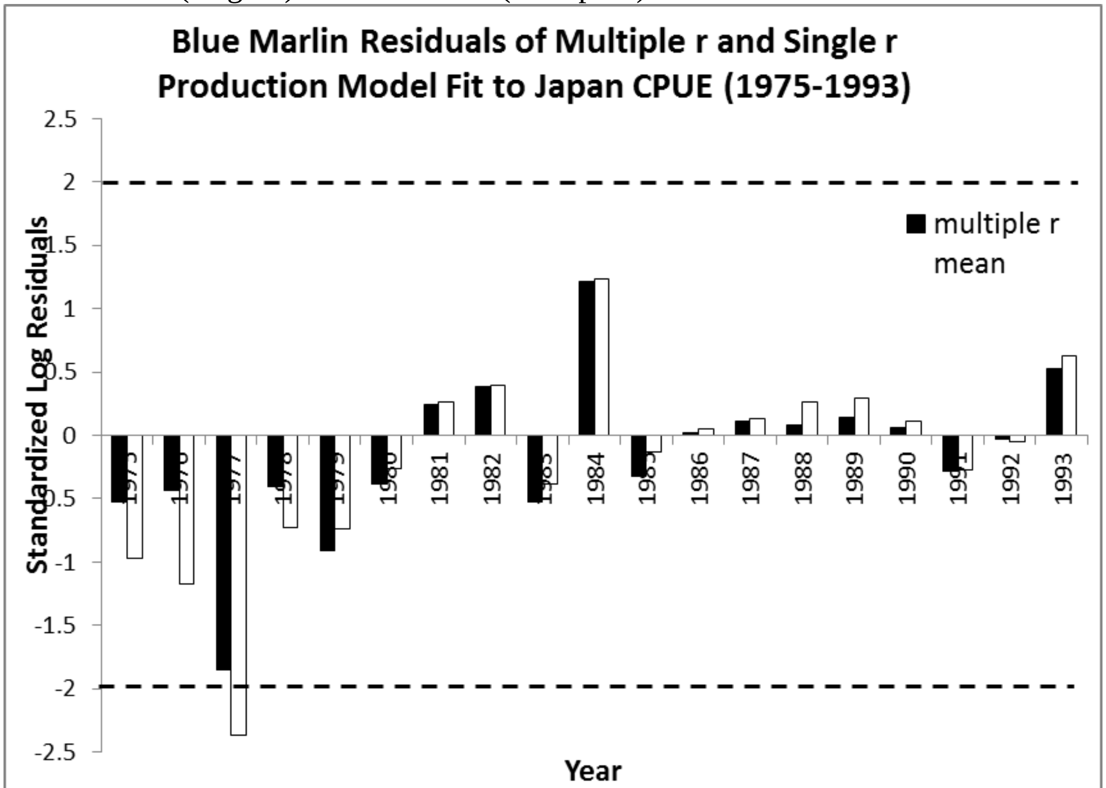
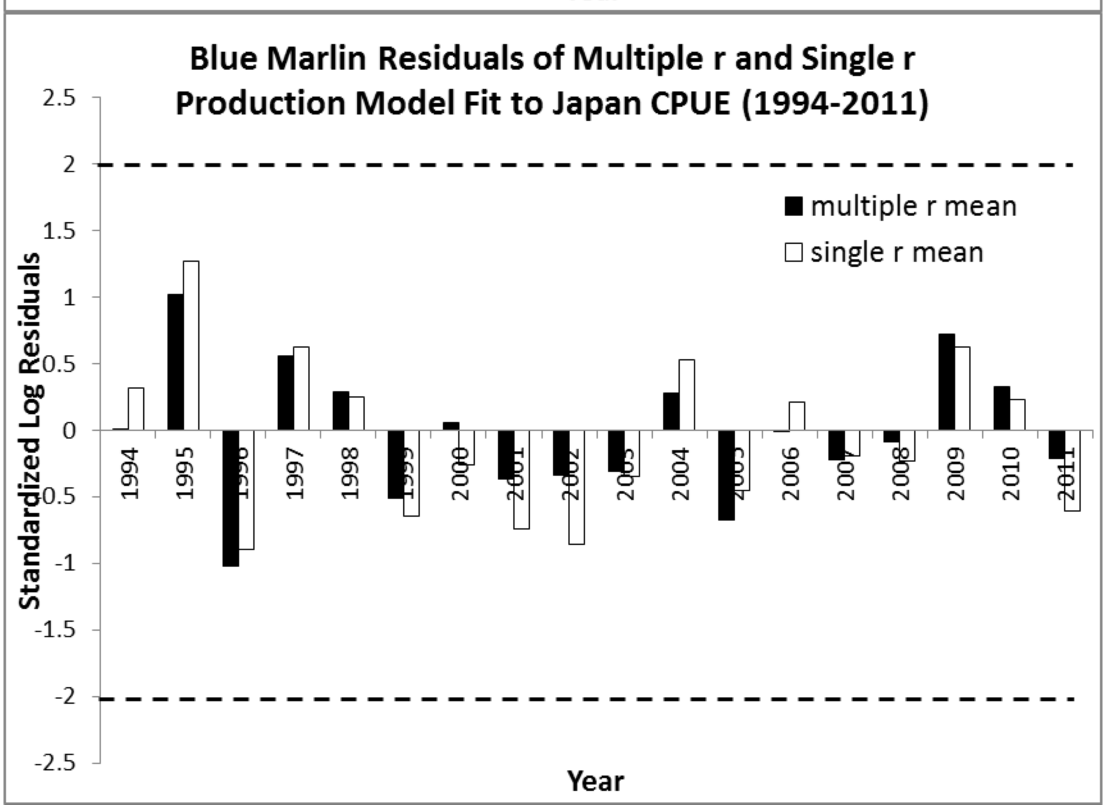
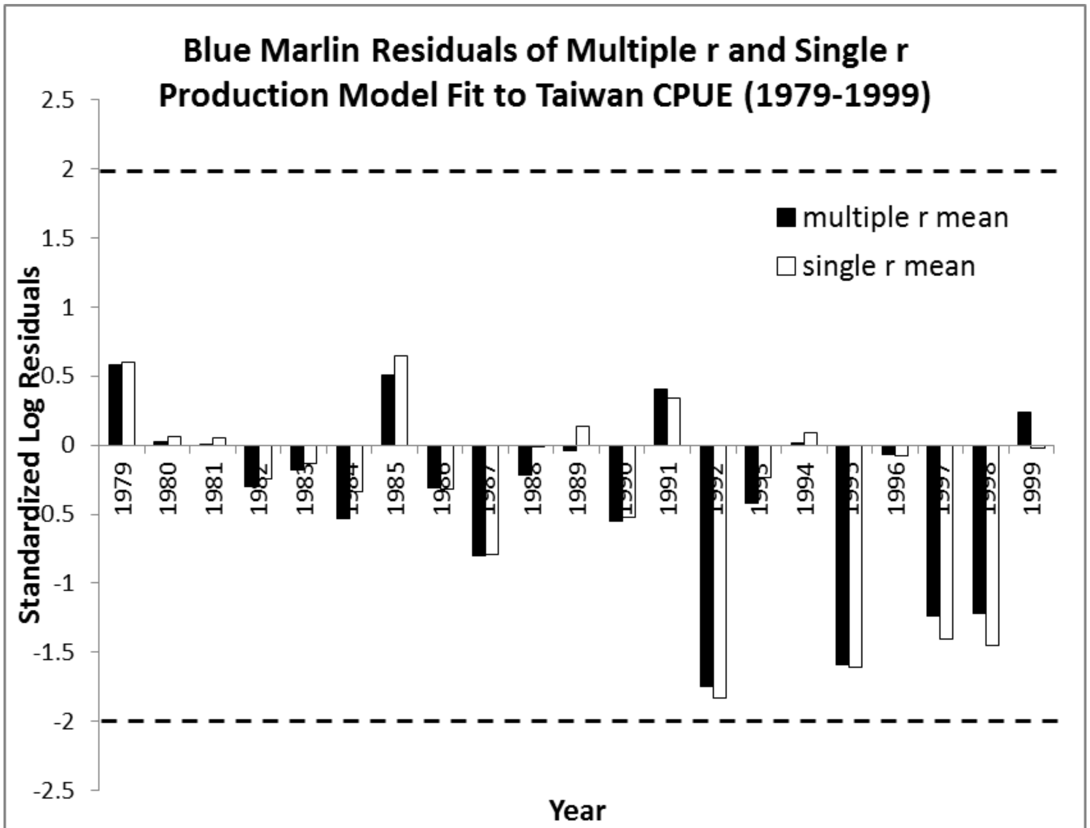


Figure 4. Standardized residuals of the production model fits to the CPUE time series of the base case (single r) and alternative (multiple r) models.






Blue Marlin Residuals of Multiple r and Single r Production Model Fit to Taiwan CPUE (1979-1999)
Standardized Log Residuals
Year
multiple r mean
single r mean
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999

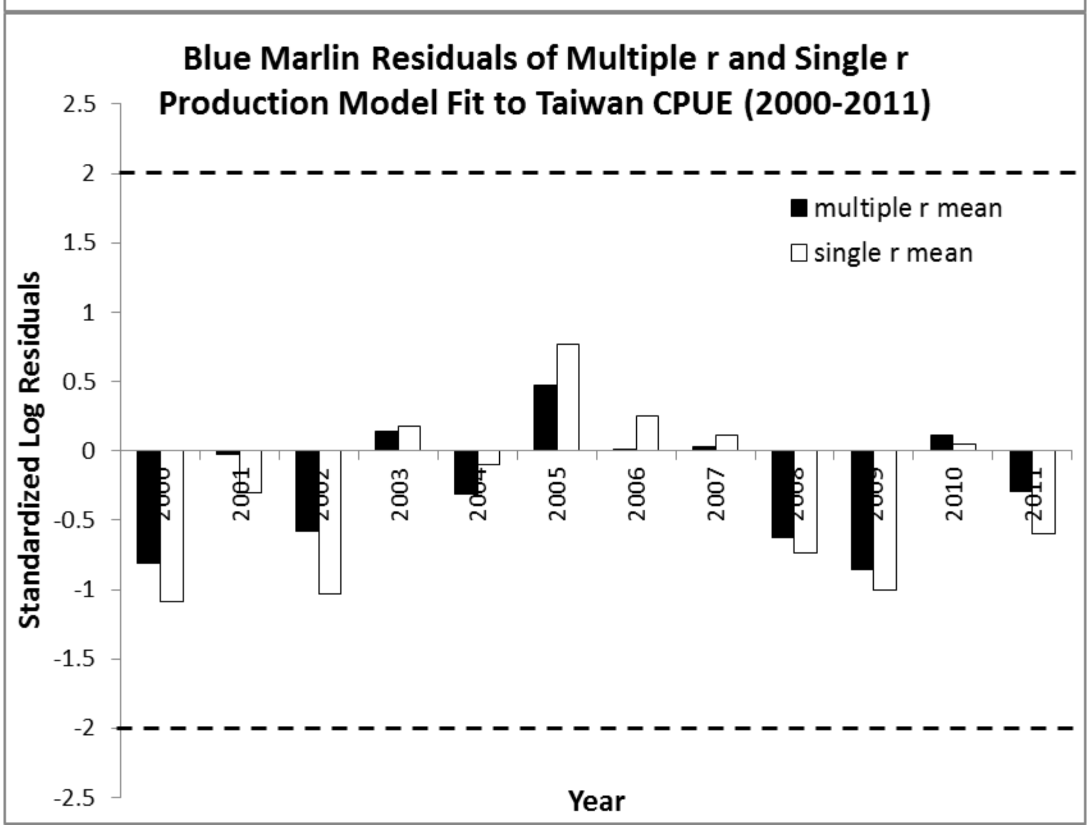

Blue Marlin Residuals of Multiple r and Single r Production Model Fit to Taiwan CPUE (2000-2011)
Standardized Log Residuals
Year
multiple r mean
single r mean
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011

Figure 5. Blue marlin model averaged estimates of exploitable biomass relative to BMSY. Error bars represent unconditional standard errors.

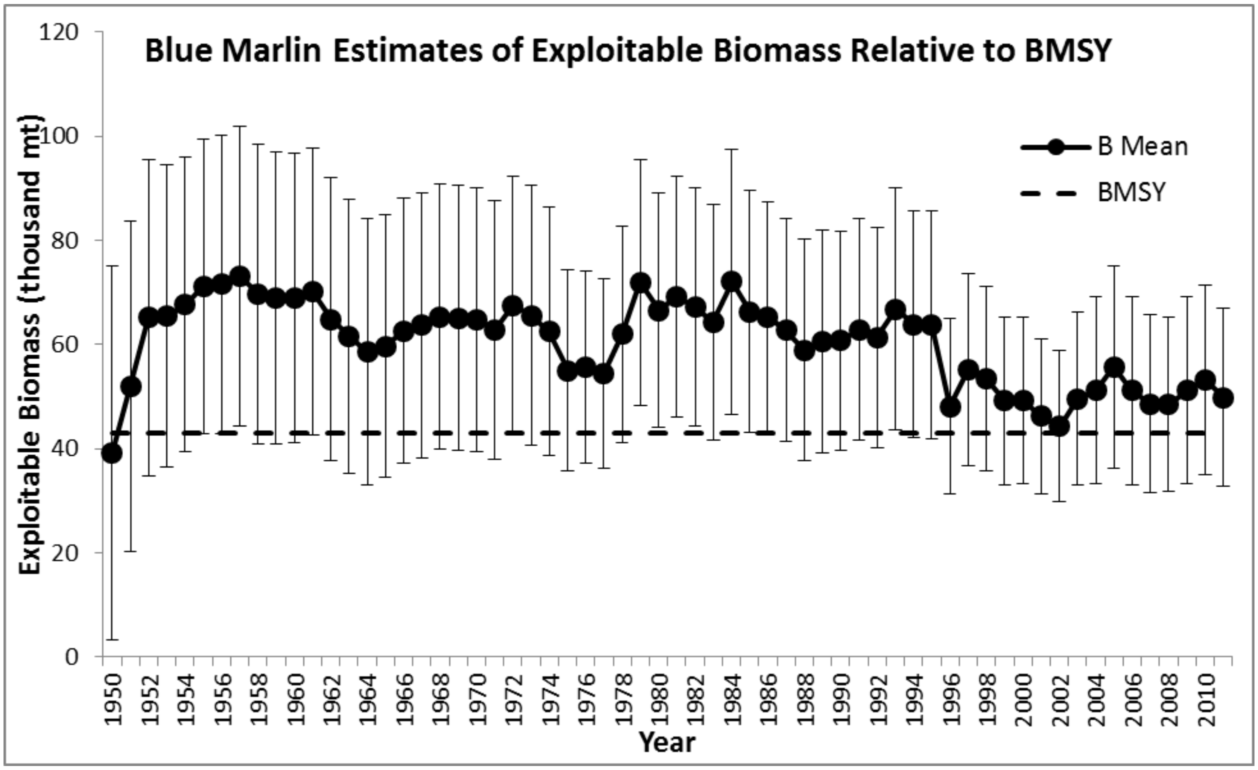


Figure 6. Blue marlin model averaged estimates of harvest rate relative to HMSY. Error bars represent unconditional standard errors.

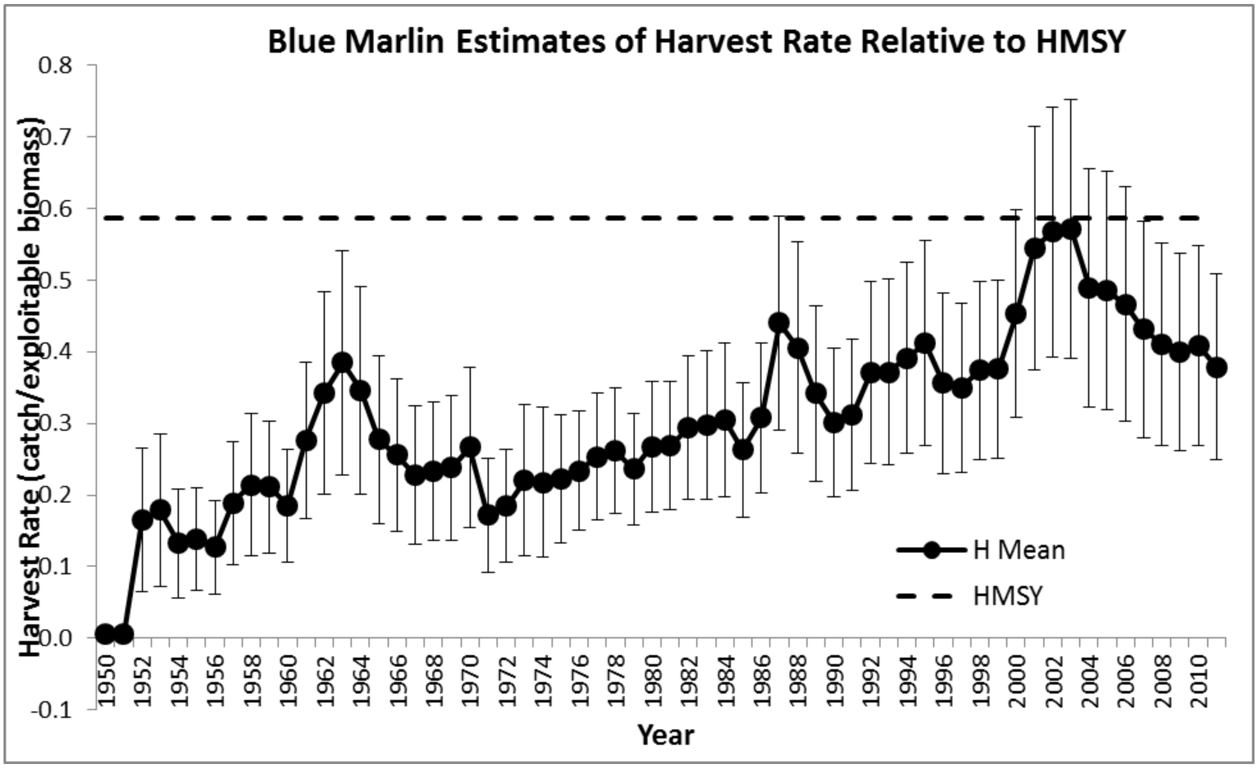


Figure 7. Blue marlin model averaged trends in biomass status. Error bars represent unconditional standard errors.

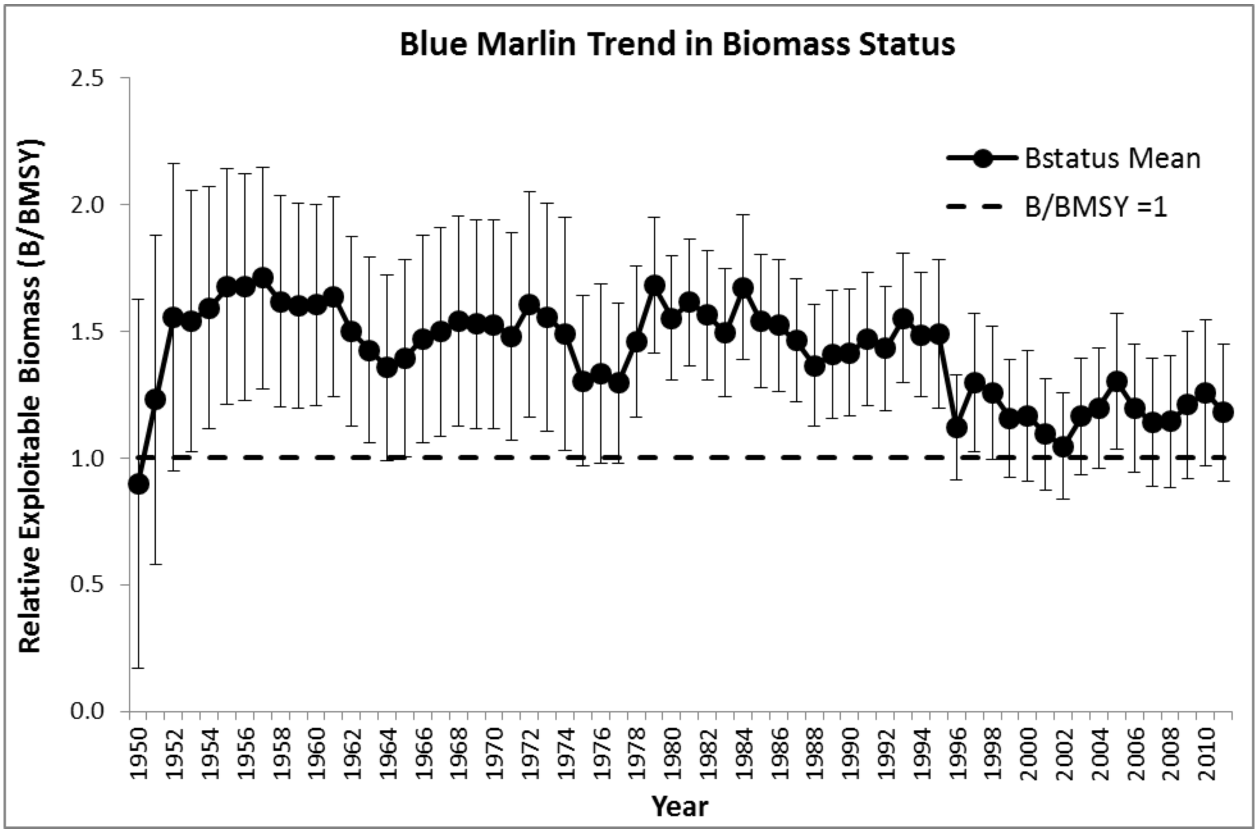


Figure 8. Blue marlin model averaged trends in harvest status. Error bars represent unconditional standard errors.

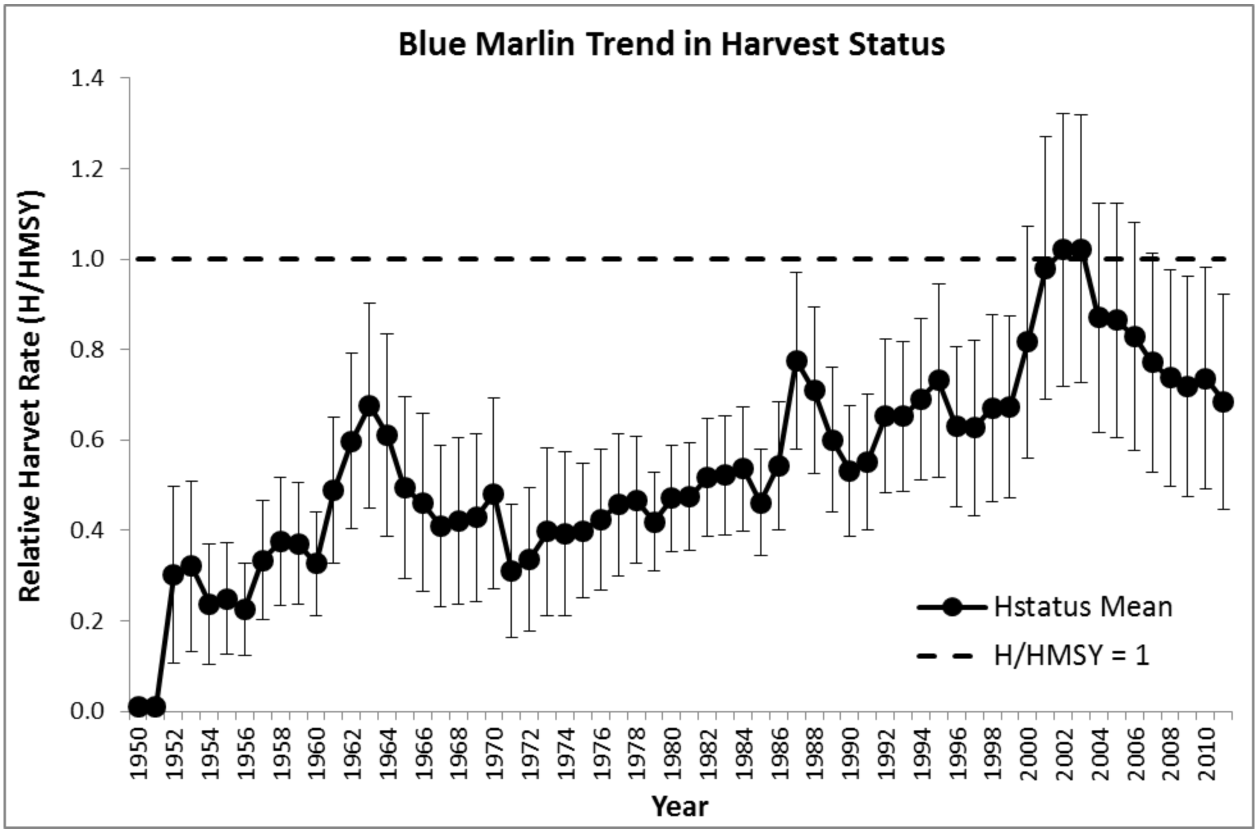


Figure 9. Blue marlin model averaged Kobe plot of relative biomass and harvest rate.