



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Fisheries Research 73 (2005) 79–97

FISHERIES
RESEARCH

www.elsevier.com/locate/fishres

Evaluation of a sex-specific age-structured assessment method for the swordfish, *Xiphias gladius*, in the North Pacific Ocean

Sheng-Ping Wang^a, Chi-Lu Sun^{a,*}, André E. Punt^b, Su-Zan Yeh^a

^a Institute of Oceanography, National Taiwan University, No. 1, Section 4, Roosevelt Road, Taipei City 10672, Taiwan

^b School of Aquatic and Fishery Sciences, University of Washington, 1122 N.E. Boat Street, Seattle, WA 98105, USA

Received 25 May 2004; received in revised form 3 January 2005; accepted 6 January 2005

Abstract

Swordfish are known to be sexually dimorphic. However, previous assessments of the status of swordfish in the North Pacific Ocean have not been based on population models that distinguish the sexes. A sex-specific age-structured assessment model was therefore constructed for the stock of swordfish in the North Pacific Ocean. However, prior to the application of this method to the data for the stock in question, its performance needs to be evaluated. Monte Carlo simulation is used to evaluate the impact of observation errors and errors in assumed parameters. The impact of observation error associated with catch-rate indices has a larger impact on estimation ability than errors in total catch and its length–frequency. Improved performance, both in accuracy and precision of the estimates of quantities of management interest, can be obtained by increasing the length–frequency sample size. The estimates of spawning stock biomass, MSY, and (to a lesser extent) fishing intensity are substantially more biased when the assessment does not consider sexual dimorphism. The ratios of current to unfished spawning stock biomass and to the spawning stock biomass corresponding to MSY were found to be the quantities estimated most robustly of those considered.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Sex-specific age-structured assessment method; Simulation; Evaluating model performance; Swordfish; North Pacific; *Xiphias gladius*

1. Introduction

Swordfish (*Xiphias gladius*, Linnaeus 1758) is a cosmopolitan species found in tropical, subtropical, temperate, and sometimes cold waters of all oceans and adjacent seas (Nakamura, 1985). In the North Pacific

Ocean (defined, for the purposes of this paper, to be the area of north of 10°N and west of 130°W; Fig. 1), the bulk of the swordfish catch has been taken by Japan, the United States and Taiwan, with very small quantities by Korea and China, whose swordfish catch is estimated to be less than 1% of total swordfish catch in the North Pacific (Anon, 2002).

Swordfish were targeted along with albacore in some areas of the North Pacific during 1952–1962 by the Japanese longline fleet when they were fishing at

* Corresponding author. Tel.: +886 2 23629842; fax: +886 2 23629842.

E-mail address: chilu@ntu.edu.tw (C.-L. Sun).

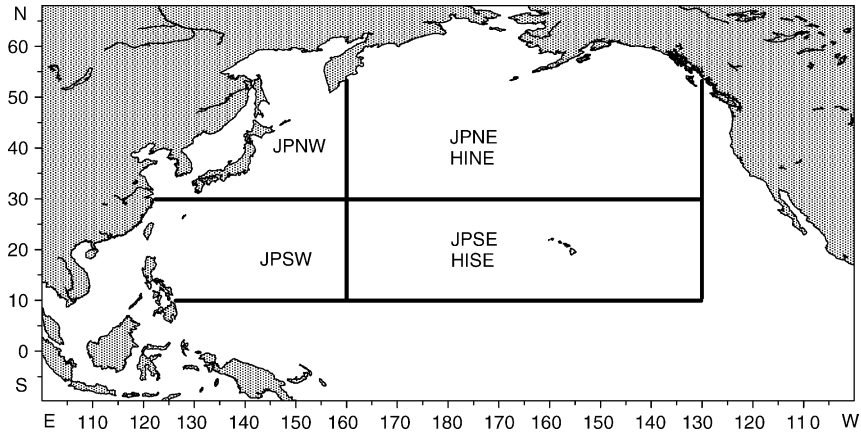


Fig. 1. The North Pacific Ocean showing the four regions considered in the analyses and the six “fleets” (four Japanese “fleets”—JPNW, JPSW, JPNE, JPSE and two US “fleets”—HINE, HISE).

night with squid as bait. Since then, this fishery changed to a primarily day operation using a mixture of bait types targeting bluefin tuna, yellowfin tuna and bigeye tuna for the high grade sashimi market. The Japanese fleets began “deep longlining” in 1974 to increase the catch-rate of bigeye tuna (Bartoo and Coan, 1989; Sakagawa, 1989). The proportion of the total catch taken by Japan decreased from more than 95% in 1969 to about 75% during 1970–1990 because of these changes to fishing practices, the introduction in 1963 of Taiwanese vessels into the fishery, and increased catches by the US in the late 1960s and early 1970s (Fig. 2). The proportion of the total catch taken by Japan declined

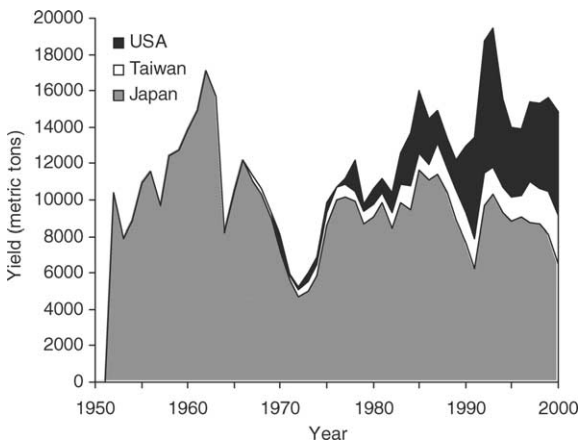


Fig. 2. Annual catch of swordfish from the North Pacific Ocean by Japan, the USA and Taiwan (1952–2001).

even further after 1990 (to about 45–65%) because of the rapid expansion of the Hawaii-based swordfish fisheries (Di Nardo and Kwok, 1998; Holts and Sosa-Nishizaki, 1998; Ito et al., 1998; Skillman, 1998; Anon, 2002).

Although not presently subject to international management, effective management of this commercially important resource will require information on stock status relative to target and limit reference points (FAO, 1995). Common fisheries reference points include maximum sustainable yield, MSY, the biomass at which MSY is achieved, S_{MSY} , and the fishing mortality corresponding to MSY, F_{MSY} . The latter quantity has been identified as a limit reference point in the U.N. Fish Stocks Agreement (United Nations, 1995).

Most previous assessments of swordfish in the North Pacific Ocean have been based on trends in catch-rates (i.e. catch-per-unit-effort, CPUE) (Yamanaka, 1958; Palko et al., 1981; Bartoo and Coan, 1989; Hanan et al., 1993; Sosa-Nishizaki and Shimizu, 1991; Di Nardo and Kwok, 1998; Kleiber and Bartoo, 1998; Hinton and Deriso, 1998; Nakano, 1998; Uosaki, 1998) and the results of production model assessments (Anon, 1999), although Kleiber and Yokawa (2002, 2004) attempted to apply MULTIFAN-CL (Fournier et al., 1998) to catch, effort and length–frequency data collected from Japanese and Hawaiian vessels. All of these analyses indicate that the swordfish stock in the North Pacific Ocean is not over-exploited and that it has been relatively stable at current levels of exploitation.

Swordfish are known to be sexually dimorphic (females grow faster and to a larger size than males (Sun et al., 2002), females mature later than males and the sex-ratio varies with length (DeMartini et al., 2000; Wang et al., 2003)). Although the studies referred to above provide a general impression of stock status for swordfish in the North Pacific, they remain preliminary in particular because they have not been based on population dynamics models that distinguish the sexes.

The absence of information about the age-structure of the catches precludes the use of standard age-structured methods of fisheries stock assessment such as tuned virtual population analysis (e.g. Pope and Shepherd, 1985) and ADAPT (Gavaris, 1988). In contrast, it is possible to assess swordfish using ‘length-based synthesis’ (e.g. Methot, 2000; Cope et al., 2004). This approach to fisheries stock assessment involves developing an appropriate (and biological realistic) population dynamics model and then estimating the values for its parameters by maximizing an objective (or likelihood) function. In this case, the data include the length–frequency of the catches.

Although it is possible to apply ‘length-based synthesis’ to the data for swordfish in the North Pacific, it is not clear whether the results will be meaningful given the relatively small sample sizes for length–frequency. Although considerable effort has been directed towards evaluating the ability of production models and age-based methods of fisheries stock assessment to estimate quantities of interest to managers such as MSY , S_{MSY} and F_{MSY} (e.g. Patterson and Kirkwood, 1995; Punt, 1997; Sampson and Yin, 1998), relatively little attention has been focused on methods which rely primarily on length–frequency data (Quinn et al., 1998; Zheng et al., 1998; Sigler, 1999).

Therefore, prior to applying length-based synthesis to the data for North Pacific Ocean swordfish, this study evaluates the likely performance of this method of assessment using Monte Carlo simulation. This is because the performance of this method cannot be inferred from the results of previous simulation evaluations of estimator performance because the previous studies have not focused on assessment methods that fit to length–frequency data and because most previous studies have focused on highly depleted resources. In contrast, the bulk of the simulations of this paper relate to a situation in which the resource is not highly depleted.

2. Materials and methods

2.1. Data used

Following Kleiber and Yokawa (2002), the North Pacific is divided into four fishing regions at 30°N and 160°E (Fig. 1). This is primarily because the US fisheries (HINE and HISE) are then restricted to the north-east and southeast regions.

Catch and effort data are available for the Japanese longline fisheries (1952–2000) and for the Hawaii-based longline fisheries (1990–2000). The data for each catch-effort record includes the catch in number of fish, the effort in hooks, and the operating type. Operating type is defined as the number of hooks between each float (i.e. hooks per basket) and is available for the Japanese longline fisheries from 1975 and for the Hawaii-based longline fisheries from 1995. The data for the Japanese longline fisheries are aggregated by 10° × 20° block, year, month, and are divided into three categories [surface operations (3–4 hooks per basket), regular operations (5–9 hooks per basket), and deep operations (more than 10 hooks per basket)]. The data for the Hawaii-based longline fisheries are aggregated by 5° × 5° block, year, month, and are also divided into three categories (less than 3 hooks per basket, 3–5 hooks per basket, and more than 5 hooks per basket).

The nominal catch-rates (defined as the number of fish per 1000 hooks) are standardized using a general linear model (GLM) approach (Gavaris, 1980; Kimura, 1981; Hilborn and Walters, 1992), with the assumption that the errors are log-normally distributed (see Wang (2004) for details).

Length–frequency data are available for the Japanese longline fisheries since 1975 and for the Hawaii-based longline fisheries since 1994. Information on the sex-structure of the catch by Japan is, however, only available after 1984. The sample sizes for the JPNW, JPSW and JPSE fleets (see Fig. 1) are too small to construct reliable sex-specific length–frequency distributions so the length–frequency data for these three fleets are aggregated over sex when treated as input to population models.

2.2. Evaluating performance using simulation

The steps considered when evaluating the performance of a stock assessment method using simulation

are:

1. Specify the population dynamics model for the true population (the operating model) and set the values for some of its parameters based on auxiliary information.
2. Fit the operating model to the data for the situation in question (i.e. North Pacific swordfish) to obtain values for the remaining parameters of this model. The results of this model fit provide the ‘true’ values for the quantities of interest (listed below).
3. Generate artificial data sets by adding noise to the predictions obtained by fitting the operating model to the data (step 2).
 - (1) Catch data: The simulated catch data (C_t^{f}) are generated by adding log-normally distributed observation error (ν_C ; $\nu_C \sim N(0, \sigma_C^2)$) to the ‘true’ catch data:

$$C_t^{f} = C_t^f e^{\nu_C - \sigma_C^2/2} \quad (1)$$
 - (2) Catch-rate data: The simulated catch-rate data (I_t^{f}) are generated by adding log-normally distributed observation error (ν_I ; $\nu_I \sim N(0, \sigma_I^2)$) to the ‘true’ catch-rate data (see Eq. (B.11)):

$$I_t^{f} = I_t^f e^{\nu_I - \sigma_I^2/2} \quad (2)$$
 - (3) Length–frequency data: Following Chen (1996), the simulated catch length–frequency data for each year, sex and fleet ($P_{t,l}^{s,f}$) are generated from the ‘true’ catch length–frequencies ($P_{t,l}^{s,f}$) by sampling multinomially from the catches-at-length with a sample size of n .
4. The estimation model (taken here to be the same as the operating model) is fit to the artificial data sets generated at step 3 to provide estimates of the quantities of interest.
5. Steps 3 and 4 are repeated 100 times—100 repetitions was sufficient to enable comparisons to be made among different length–frequency sample sizes and coefficients of variation for the catch and catch-rate data.

The sensitivity of the results to all combinations of three levels of error when quantifying the catch (σ_C : 0.01, 0.05 and 0.10), four levels of error about the relationship between catch-rate and abundance (σ_I : 0.01, 0.1, 0.25 and 0.5) and five levels for the extent of sampling for length–frequency (sample sizes, n , of 10, 50,

100, 250 and 500 for each year, fleet and sex) is examined. The values for σ_C , σ_I and n considered in the simulations are likely to bound the actual values of these parameters for North Pacific swordfish.

2.3. The population dynamics model

The population dynamics model that forms the basis for the operating model and the estimation model is described in Appendix A. This model is sex- and age-structured and considers both sexes from age 0 to age 15 (age 15 being treated as a ‘plus group’). The model assumes that recruitment is related to spawning stock biomass according to a Beverton–Holt stock–recruitment relationship and that the deviations about this relationship are log-normally distributed. Selectivity is assumed to be a combination of logistic and dome-shaped components to adequately reflect the length-composition of the JPSE fleet. Different selectivity patterns are assumed for the various fleets (combination of logistic and dome-shaped for the JPSE fleet and logistic for the other fleets; Fig. 3). Owing to lack of data, the recruitment deviations prior to 1971 and those thereafter are treated differently. For the operating model, the former are generated from $N(0; \sigma_v^2)$ while the latter are estimated. For the estimation model, however, the recruitment deviations for the years prior to 1971 are all set to zero because there are no data which could inform year-class strength for these years.

2.4. Parameter estimation

The parameters of the model can be divided into those for which auxiliary information is available (Tables 1 and 2) and those which need to be estimated from the monitoring data (Table 3). The data available for assessment purposes are: (1) the catches (assumed known without error), (2) the annual length–frequencies by sex and fleet (pooled across sex in some instances), (3) the sex-ratio data by fleet (the ratio of the catch (in numbers) of animals sexed to be females to the total catch of males and females), and (4) the catch-rate-based indices of abundance. The objective function minimized¹ to find the estimates of

¹ Using AD Model Builder (version 5.0.2) (Otter Research Ltd., 2000).

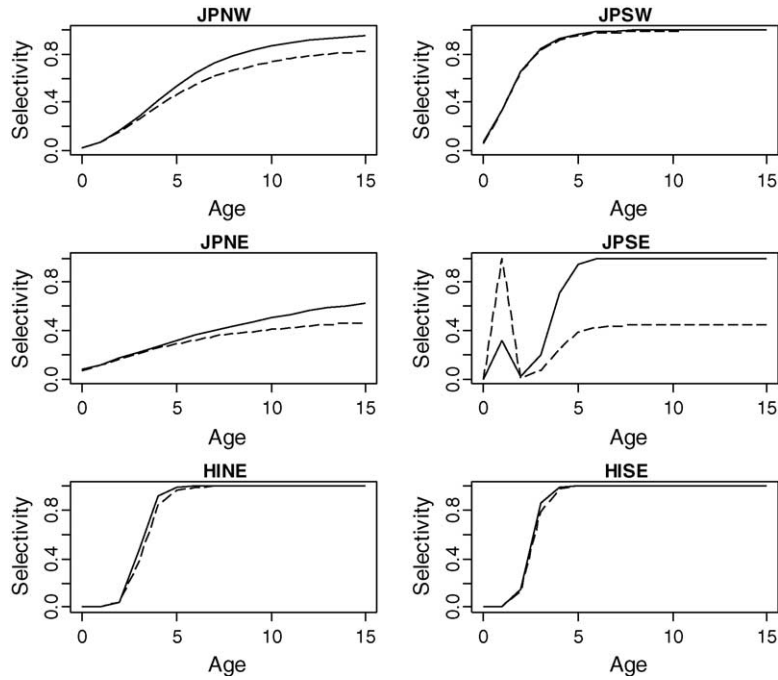


Fig. 3. “True” age-specific selectivity ogives for the six fleets (solid lines: female; dashed lines: male).

the ‘free’ parameters of the model includes two components (the data available for assessment purposes and the constraints based on a priori assumptions) (Appendix B).

The values for the parameters of the length–weight relationship, growth and maturity are obtained from Sun et al. (2002) and Wang et al. (2003) (Tables 1 and 2). However, the values for the parameters related to natural mortality (M), the steepness of

the stock–recruitment relationship (h), and the extent of variation in recruitment (σ_v) cannot be determined from auxiliary information, nor can they be estimated reliably by fitting the model to the data (results not shown) and must therefore be pre-specified. In this study, M is taken to be 0.25 year^{-1} based on Pauly’s empirical equation (Pauly, 1980), h is assumed to be 0.9 (Anon, 1997; Punt et al., 2001), and σ_v is assumed to be 0.4 (Punt et al., 2001). The impact of the assessment model making erroneous assumptions regarding

Table 1

The values assumed for the parameters of the relationships between length and weight, length and age, and maturity and length

| Parameter | Females | Males |
|--|-------------------------|-------------------------|
| Asymptotic size, L_∞^s (cm) | 300.656 | 213.052 |
| Growth parameter, k^s (year^{-1}) | 0.04 | 0.086 |
| Age-at-zero-length, a_0^s (year) | −0.75 | −0.626 |
| Shape parameter, m^s | −0.785 | −0.768 |
| Length–weight, A^s | 1.3528×10^{-6} | 1.3528×10^{-6} |
| Length–weight, B^s | 3.4297 | 3.4297 |
| Length-at-50%-maturity, L_m (cm) | 168.16 | – |
| Maturity slope, r_m | −0.1392 | – |
| Maximum age, λ (year) | 15 | 15 |

Table 2

The values assumed for the standard deviation of the length-at-age

| Age | Female | Male |
|-----|--------|-------|
| 0 | 25.00 | 25.00 |
| 1 | 18.63 | 15.92 |
| 2 | 10.79 | 9.59 |
| 3 | 10.27 | 10.82 |
| 4 | 12.26 | 10.28 |
| 5 | 13.39 | 11.46 |
| 6 | 9.79 | 11.53 |
| 7 | 14.41 | 13.94 |
| 8 | 12.64 | 11.47 |
| 9 | 15.45 | 11.21 |
| ≥10 | 10.00 | 10.00 |

Table 3

The parameters of the population dynamics model not known from auxiliary information (the number of years was 49 and number of fleets was 6)

| Parameter | No. of parameters |
|---|-------------------|
| Estimated | |
| Recruitment at unfished equilibrium, R_0 | 1 |
| Process error, v_t | 1 per year |
| Selectivity | |
| Length-at-50%-selectivity, L_{50}^f | 1 per fleet |
| Length-at-95%-selectivity, L_{95}^f | 1 per fleet |
| Length-at-modal-selectivity, L_{mu}^f | JPSE fleet only |
| Standard deviation of selectivity, L_{sd}^f | JPSE fleet only |
| Weight assigned to dome-shaped ogive, φ^f | JPSE fleet only |
| Pre-specified | |
| Natural mortality, M | 1 |
| Steepness, h | 1 |
| Variation in recruitment, σ_v | 1 |

the values for these parameters is examined in the tests of sensitivity.

2.5. Summarizing estimator performance

The operating model includes a large number of outputs. In this study, focus is placed on the ability to estimate the following key outputs:

- (1) S_0 , the spawning stock biomass at unfished equilibrium;
- (2) S_{2000} , the spawning stock biomass at the start of the last year of the assessment period;
- (3) F_{2000} , the fleet-aggregated fishing intensity (see Section 7 of [Appendix A.7](#) for the definition of this quantity) during the last year of the assessment period;
- (4) MSY, the maximum sustainable yield (see [Shepherd and Pope \(2002\)](#) for details on how MSY is calculated given the estimates for growth, maturity and the stock–recruitment relationship);
- (5) S_{MSY} , the spawning stock biomass at which MSY is achieved;
- (6) F_{MSY} , the fleet-aggregated fishing intensity at which MSY is achieved;
- (7) S_{2000}/S_0 , the ratio of the spawning stock biomass at the start of the last year of the assessment period to S_0 ;
- (8) S_{2000}/S_{MSY} , the ratio of the spawning stock biomass at the start of the last year of the assessment period to that at which MSY is achieved;

- (9) F_{2000}/F_{MSY} , the ratio of F_{2000} to the fleet-aggregated fishing intensity at which MSY is achieved.

Quantities (1) and (2) relate to the ability to estimate absolute population size (both at present and in the past), while quantity (3) evaluates the ability to estimate the current exploitation rate. A fleet-aggregated measure of exploitation rate is used in quantity (3) to reduce the volume of results. Quantities (4)–(6) relate to the ability to estimate the yield, spawning stock biomass and exploitation rate at which MSY is achieved. Unlike quantities (1)–(6), quantities (7)–(9) relate to relative measures of spawning stock biomass and exploitation rate. Past evaluations of the performance of stock assessment methods (e.g. [Punt et al., 2002](#)) suggest that relative measures should be estimated better than absolute measures.

Estimation ability is quantified for each Monte Carlo replicate and quantity of interest by the relative error:

$$E_i^j = \frac{\hat{Q}_i^j - Q^j}{Q^j} \times 100 \quad (3)$$

where E_i^j is the relative error (%) for simulation i and quantity j , \hat{Q}_i^j the value for simulation i and quantity j provided by the estimation model, and Q^j the ‘true’ value for quantity j .

The distributions of median relative errors for each scenario and quantity of interest are summarized by the median relative error (MRE) and inter-quartile range (IQR).

3. Results

3.1. Impacts of individual factors

The first set of results is based on simulations in which only one source of data (catches, catch-rates and length–frequency) is subject to error.

3.1.1. Impact of errors when measuring catch

[Fig. 4](#) shows the distributions for the relative errors for the nine quantities of management interest for each of the three levels of observation error for the catch data. The sizes of the relative errors increase with extent of error in catch. As expected, the estimates

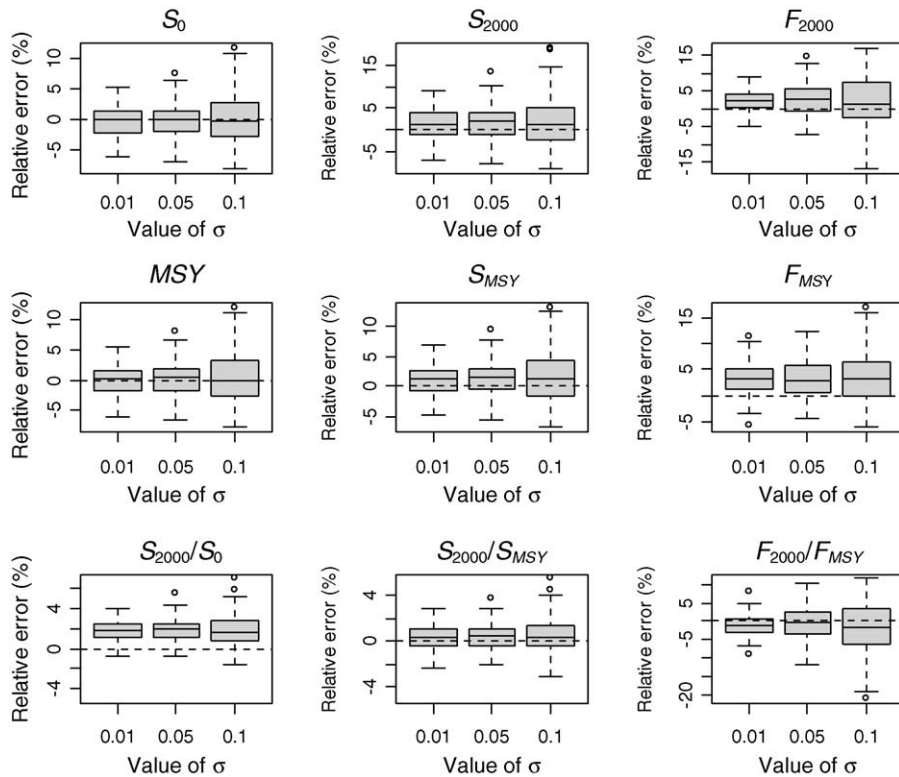


Fig. 4. Box plots of the relative errors for the quantities of management interest corresponding to various levels of observation error for the catch data. The centerline represents the median and the box represents the quartiles. The whiskers extend 1.5 times the inter-quartile range.

are almost identical to the true values when $\sigma_C = 0.01$ (MREs between -1.1 and 3.1% and IQRs between 1.4 and 4.9% for all quantities). Estimation error remains good (MREs between -0.5 and 2.8% and IQRs between 1.4 and 6.0%) for a more realistic extent of error in the catches ($\sigma_C = 0.05$). The ability to determine the quantities of management interest deteriorates somewhat when $\sigma_C = 0.1$. The estimates of relative biomass (i.e. S_{2000}/S_0 and S_{2000}/S_{MSY}) are close to being unbiased (MREs of 1.6 and 0.3% , respectively). However, noteworthy bias and imprecision is evident for some of the other quantities when $\sigma_C = 0.1$. For example, F_{MSY} is positively biased (MRE = 3.2%) while F_{2000}/F_{MSY} is imprecise (IQR = 9.2%).

The results in Fig. 4 are based on the assumption that the estimates of catch are unbiased. Given the lack of management actions for swordfish in the North Pacific, there is no obvious reason why the catches by the nations considered in the assessment should be biased. However, catches by nations other than the Japan

and the US may lead to the catches used in the assessment being negatively biased estimates of the total removals. The impact of a negative bias of 10% in total catch (over all years) was examined and found to have little impact on the estimates of S_{2000}/S_0 , S_{2000}/S_{MSY} , and F_{2000}/F_{MSY} (MREs of 0.3 , 0.4 and 1.0% , respectively). In contrast, the estimates of S_0 , S_{2000} , S_{MSY} and MSY were negatively biased (MREs $\sim -9.0\%$) when catches were negatively biased and those of F_{2000} and F_{MSY} were positively biased (MREs between 5.3 and 6.3%).

3.1.2. Impact of errors in the catch-rate data

Fig. 5 shows the distributions for the relative errors for the nine quantities of management interest for each of the four levels of observation error for the catch-rate data. The impact of a low level of observation error in the catch-rate data ($\sigma_I = 0.01$) is minor, but larger than the lowest level of variation in catches (MREs between -3.4 and 10.6% and IQRs between 1.4 and 6.2% for

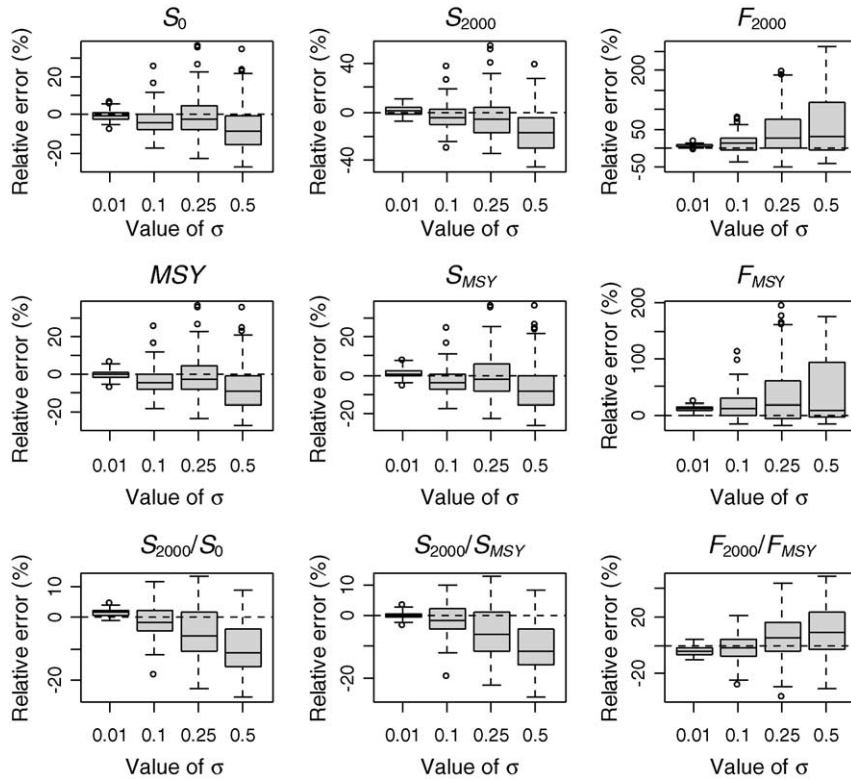


Fig. 5. Box plots of the relative errors for the quantities of management interest corresponding to various levels of observation error associated with the catch-rate data. The centerline represents the median and the box represents the quartiles. The whiskers extend 1.5 times the inter-quartile range.

all quantities). The accuracy and precision of the estimates deteriorates when σ_1 is increased from 0.01 to 0.1 (the MREs now lie between -5.1 and 12.5% and IQRs between 6.7 and 30.6%) and the pattern of greater bias and lower precision with higher values for σ_1 is clearly evident in Fig. 5. As a consequence, the estimates can be very biased (MREs between -17.0 and 29.3%) and highly imprecise (IQRs between 11.4 and 118.5%) for the highest level of observation error considered in Fig. 5. It is perhaps of some concern that the relative measures are biased when there is large uncertainty about the relationship between catch-rate and abundance. However, the extent of bias is less for the ratios than for their constituent parts (e.g. F_{2000}/F_{MSY} is less biased than either F_{2000} or F_{MSY}) confirming the expectation that the estimates of ratios of quantities are more reliable than the estimates of absolute values.

3.1.3. Impact of errors in the length–frequency data

The estimates of the management quantities except S_{2000}/S_0 and S_{2000}/S_{MSY} are biased (MREs between -11.9 and 116.4%) when the length–frequency sample size is quite small ($n=10$). However, the MREs are close to zero (-2.0 – 3.8%) and IQRs are small (2.5 – 11.3%) for all quantities for sample sizes larger than 50 (Fig. 6). There is little discernable change in performance once the sample size reaches 100.

3.2. Impact of factors in combination

Distributions of relative error are shown for three quantities of interest for three levels for the extent of observation error for the catch-rate data ($\sigma_1 = 0.1, 0.25$ and 0.5) and three length–frequency sample sizes ($n = 50, 100$ and 250) in Fig. 7. The extent of observation error

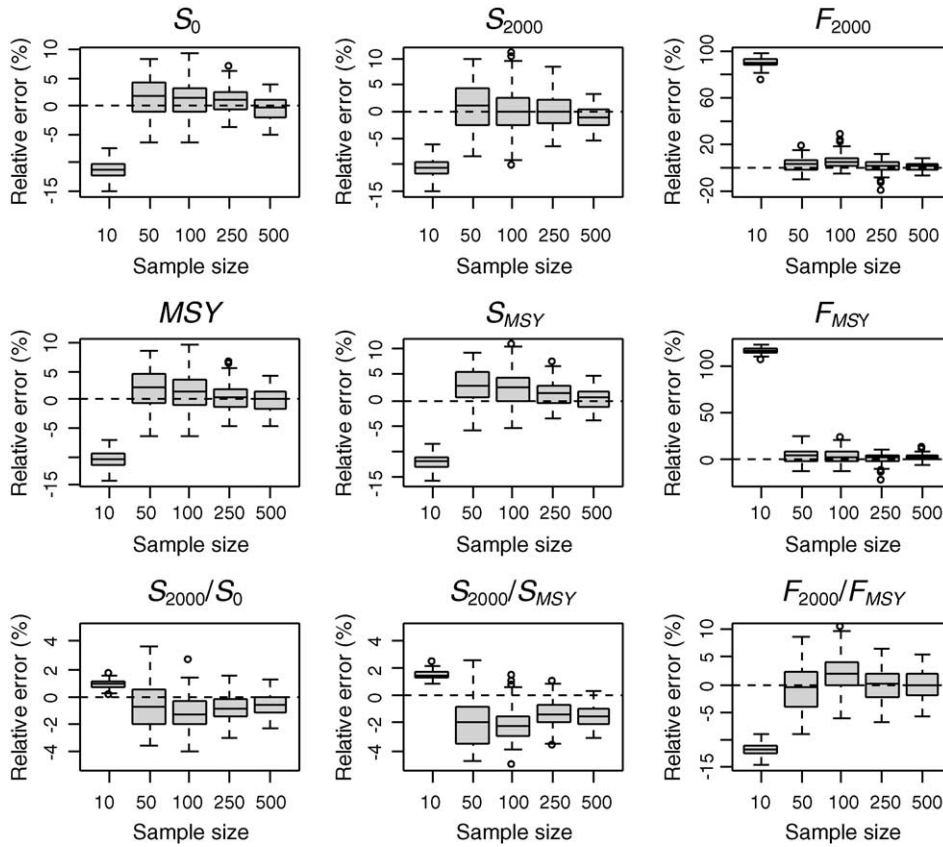


Fig. 6. Box plots of the relative errors for the quantities of management interest corresponding to various sample sizes for the length–frequency data. The centerline represents the median and the box represents the quartiles. The whiskers extend 1.5 times the inter-quartile range.

for the catch data is the same for all of the simulations in Fig. 7 ($\sigma_C = 0.1$). Fig. 7 is restricted to the three quantities which performed ‘best’ for the simulations in which only one data source is subject to error. Relative biomass is generally determined better than relative exploitation rate (MREs between -14.1 and -11.6% ($n = 50$ and 100) and of -3.7 and -4.8% ($n = 250$) for relative biomass and of 31.0 , 35.8 , and 7.8% ($n = 50$, 100 and 250 , respectively) for relative exploitation rate). As before, less catch-rate observation error and larger length–frequency sample sizes lead to improved performance, although the impact of, for example, increasing the sample size from 100 to 250 is much more substantial in Fig. 7 than was the case in Fig. 6.

3.3. Impact of fixing parameters to the wrong values

The results in Figs. 4–7 are based on the assumption that natural mortality, stock–recruitment steepness, and the extent of variation in recruitment are known exactly. This is clearly unrealistic so Fig. 8 explores the implications of the estimation model assuming incorrect values for these parameters ($M = 0.2$ and 0.3 year^{-1} ; $h = 0.6, 0.8$ and 0.95 ; $\sigma_v = 0.2$ and 0.6). The values for these parameters in the operating model were set equal to their base values (i.e. $M = 0.25$, $h = 0.9$, $\sigma_v = 0.4$). Results are shown in Fig. 8 for the same three quantities as in Fig. 7 and for $\sigma_C = 0.1$, $\sigma_1 = 0.25$, and $n = 100$.

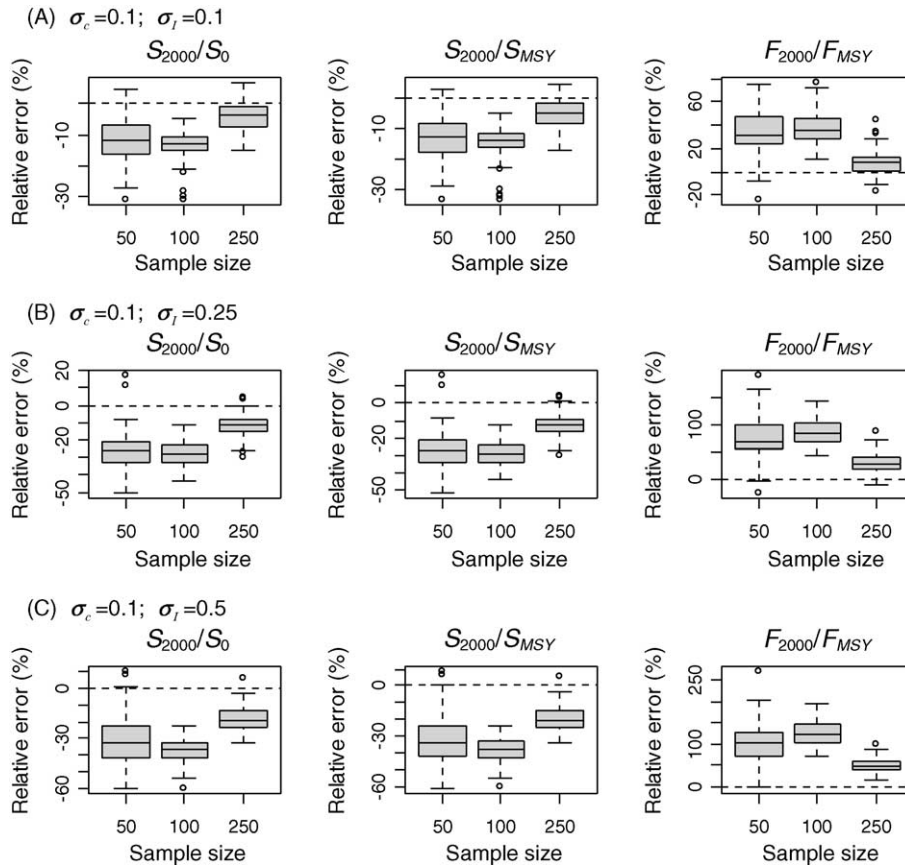


Fig. 7. Box plots of the relative errors for three quantities of management interest for various length–frequency sample sizes and catch-rate observation error standard deviations. The centerline represents the median and the box represents the quartiles. The whiskers extend 1.5 times the inter-quartile range.

The results in Fig. 8 are somewhat counter-intuitive because the best performance does not occur when the values for h , M and σ_v assumed when conducting the assessment equal their ‘true’ values. For example, the estimates were more accurate (less biased) when $M = 0.3 \text{ year}^{-1}$ and $\sigma_v = 0.2$. The results for h show no consistent pattern; setting $h = 0.6$ leads to more biased estimates of S_{2000}/S_{MSY} and F_{2000}/F_{MSY} but less biased estimates of S_{2000}/S_0 . The impact of the values assumed for h and M on bias substantially exceeds that of the impact of the value assumed for σ_v . The sensitivity of the results to changing the values for the weights assigned to the various data sources (see Appendix B) was explored, but found to be fairly minor.

3.4. Impact of sex factor of the model

The estimation model considered in Figs. 4–8 is based on a sex-structured population dynamics model. A set of simulations were therefore conducted in which the estimation model was based on sex-aggregated rather than sex-specific growth curves (length-at-age was set equal to the average of that for males and females). The values for the parameters M , h and σ_v were set to their base values (i.e. assumed to be known exactly) while the data were generated based on $\sigma_c = 0.1$, $\sigma_l = 0.25$, and $n = 100$. Fig. 9 compares the relative error distributions for all nine management-related quantities for the sex-specific and sex-aggregated estimation models. The estimates of S_0 , S_{2000} , MSY and

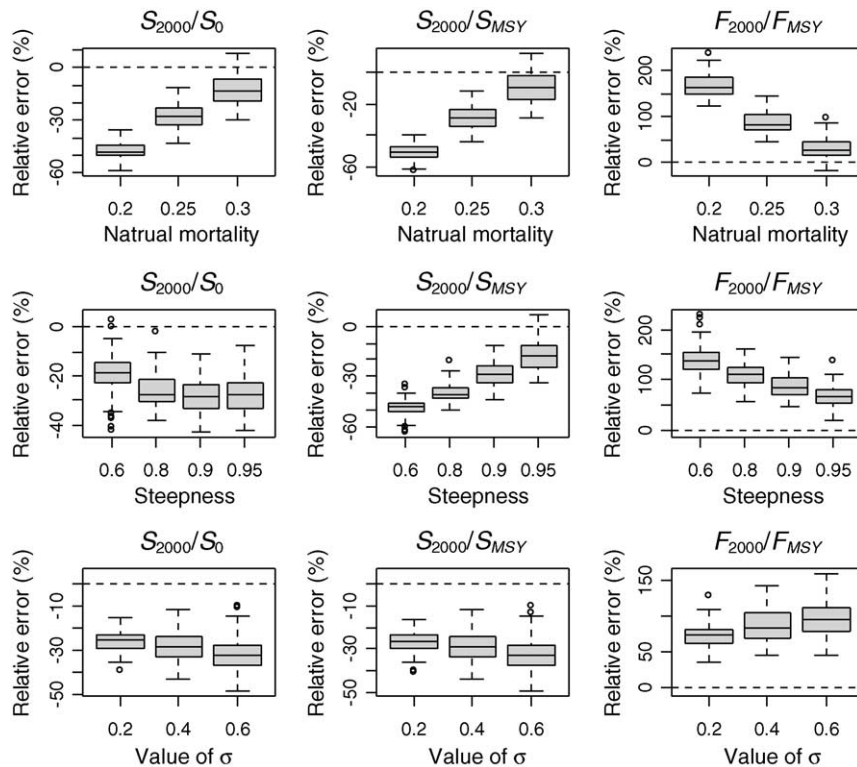


Fig. 8. Box plots of the relative errors for three quantities of management interest for analyses in which the values for natural mortality, stock–recruitment steepness, and the extent of variation in recruitment differ from those in operating model. The centerline represents the median and the box represents the quartiles. The whiskers extend 1.5 times the inter-quartile range. $\sigma_c = 0.1$, $\sigma_I = 0.25$, and $n = 100$ for all of the analyses in this figure.

S_{MSY} are substantially more biased (MREs of -70.2 , -80.7 , -53.8 and -75.5% , respectively) when based on the sex-aggregated estimation model. The estimates of F_{2000} , F_{MSY} and S_{2000}/S_{MSY} are also more biased (and S_{2000}/S_{MSY} less biased) when sexual dimorphism is ignored when conducting the assessment.

4. Discussion

4.1. Impact of the factors considered

Of the three data sources considered (catches, catch-rates, and length–frequency), uncertainty about the relationship between catch-rate and abundance has the greatest impact on performance. This is reflected not only in reduced precision but also in increased bias (Fig. 5). Although the biases are relatively small for $\sigma_I = 0.1$, they can be fairly substantial (e.g. $\sim 30\%$ for

F_{2000}) when σ_I is larger than 0.25. Fitting the age- and sex-structured model to the catch-rate indices for North Pacific swordfish suggests that σ_I may, in fact, be 0.1 or less (Wang, 2004).

For swordfish in the North Pacific, catches should be known (at least for recent years) with a coefficient of variation of 10%. The impact of increased length–frequency sample sizes is fairly small when there is no uncertainty associated with the catches or catch-rates. In contrast, larger (effective sample sizes > 250) sample sizes do improve estimation performance when there are several sources of uncertainty (Fig. 7). Unfortunately, the range of effective sample sizes inferred by fitting the age- and sex-structured model to the data for North Pacific swordfish is 1–513 (average 83). This suggests that the actual length–frequency sample sizes are not quite large enough for some fleets and there may be value in increasing the number of fish measured to improve

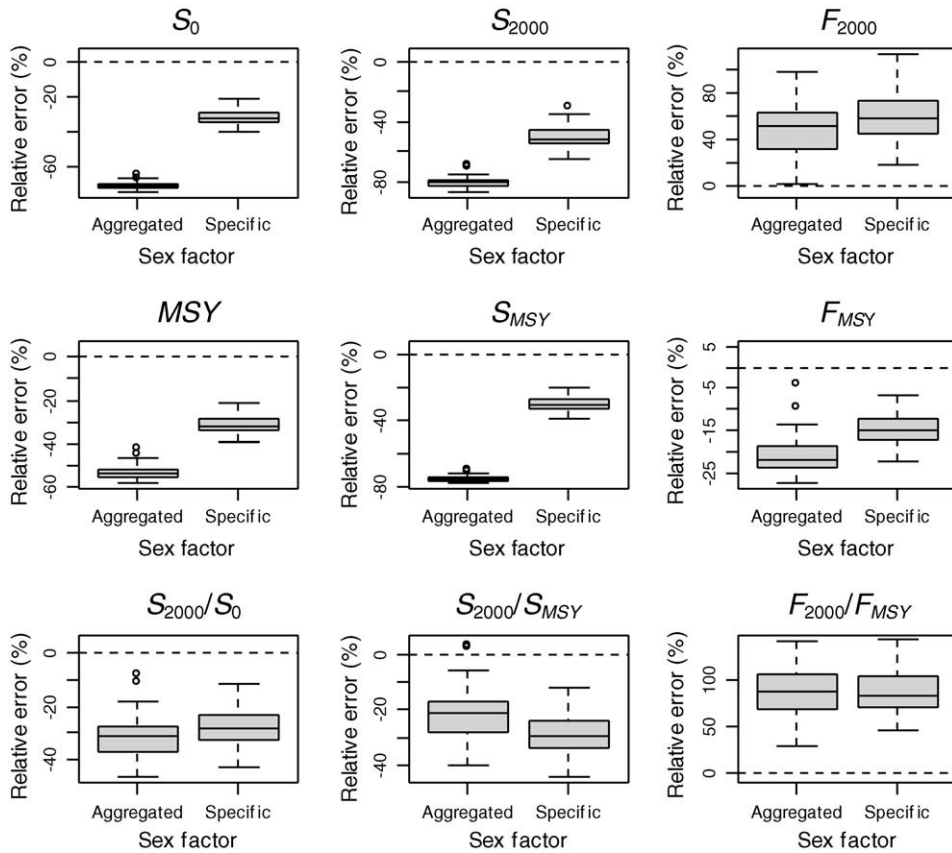


Fig. 9. Box plots of the relative errors for the quantities of management interest for the sex-specific and sex-aggregated estimation models. The centerline represents the median and the box represents the quartiles. The whiskers extend 1.5 times the inter-quartile range. $\sigma_c = 0.1$, $\sigma_I = 0.25$, and $n = 100$ for all of the analyses in this figure.

the accuracy and imprecision of the estimates from the model.

The results are clearly sensitive to the values assumed for M and h (Fig. 8). Moreover, the best performance did not occur when the values assumed for these parameters were the true values. This is surprising, but reflects the complicated nature of the estimation method. In principle, estimates for M , h and σ_v could be obtained from research (such as tagging studies to determine M —Hampton, 2000). However, in the short- to medium-term, it seems likely that the values for these parameters will have to be based on the results of studies for other stocks of swordfish. Unfortunately, the data for all swordfish stocks are sparse (Punt et al., 2001) and it may be inappropriate anyway to infer the values for the biological parameters for a stock in one ocean basin from those for stocks in other ocean basins.

The estimates of spawning stock biomass, MSY , and (to a lesser extent) fishing intensity are substantially more biased when the assessment ignores sexual dimorphism (Fig. 9). Therefore, if control rules based on, for example, F_{MSY} or MSY are to be used to manage swordfish in the North Pacific Ocean, an assessment approach which allows for sexual dimorphism (such as that outlined in Appendices A and B) needs to be applied to reduce uncertainty. It is perhaps somewhat reassuring from a conservation viewpoint that the estimates of the biomass-related management quantities are negatively biased while those of the fishing mortality-related quantities are positively biased when the assessment model ignores sex-structure (Fig. 9). However, the extent to which this result is general and applies to other stocks of large pelagic fish is unclear, although this could be

evaluated using simulations such as those of this paper.

4.2. Most robust quantities of interest

The accuracy and precision of the various management-related quantities differed markedly. The quantities S_{2000}/S_{MSY} and S_{2000}/S_0 were least biased of the nine quantities for the simulations in which observation error was added to the catches and the length–frequency data (Figs. 4 and 6). The accuracy and precision of all nine quantities deteriorated substantially when the catch-rates were subject to observation error (Fig. 5). Nevertheless, these two quantities remained the most robustly determined even when all of the data inputs were subject to error. This result is consistent with the conclusions of studies that focused on biomass dynamics and age-structured models (e.g. Punt, 1995, 1997; Punt et al., 2002).

Somewhat surprisingly, the ratio of current fishing intensity to that corresponding to MSY is rather poorly determined (Figs. 7–9) even though it is a ratio. The reasons for this are unclear, but may relate to the greater dependence of fishing intensity on selectivity (see Eqs. (A.15) and (A.16)).

4.3. General discussion

The impact of noisy catch-rate data can be offset by larger length–frequency sample sizes (Fig. 7). There is a trade-off between cost and precision when deciding on samples sizes for length–frequency and age–composition data (e.g. Chen, 1996). Nevertheless, attempts should be made to develop a sampling program that aims to achieve a random sample of at least 250 animals from the catch of each sex for each fleet. However, about 60% of the effective sample sizes computed by fitting the model to the actual length–frequency data for North Pacific swordfish are not only less than 250 but also less than 83 (the average effective sample size). It is not possible at present to relate effective sample size to the actual number of fish measured. However, given the non-random nature of fishing operations, collecting 250 animals randomly from the catch (e.g. using observers) could necessitate the collection of a large number of length measurements for some fleets. The level of 250 animals as being sufficient should, however, not be taken as an

absolute limit because the simulations ignored the possibility of time-varying selectivity. If selectivity varies over time, much larger sample sizes will be needed to achieve the same level of precision.

A further assumption underlying the analyses is that the growth curve (by sex) is known without error. The results of the assessments will be biased if the growth curve is in error or if growth is, in fact, density-dependent and the assessment assumes it to be time-invariant.

This study highlights the importance of obtaining an index of abundance that is both precise and related in a known (in this study, linear) way to abundance. The catch-rate indices used when fitting the operating model were based on the outcomes from a general linear model. An alternative approach is to integrate the catch-effort standardization with fitting the population dynamics model (Maunder, 2001; Maunder and Langley, 2004). This approach has several advantages over the conventional two-step process of standardizing the catch and effort data first and then fitting the population model, although it is computationally highly intensive. To date, applications of this integrated approach have been restricted to age-structured and biomass dynamics models. Extension of the approach to models based on length–frequency data is needed although testing such an assessment approach by means of Monte Carlo simulation may be computationally prohibitive at present.

Kleiber and Yokawa (2002, 2004) attempted to assess the stock of swordfish in the North Pacific Ocean using MULTIFAN-CL. However, the results of this study suggest that estimates of some of the quantities of interest to management (including estimates of current and historical spawning stock biomass) will be substantially biased and that an assessment approach takes sexual dimorphism into account should lead to less biased estimates of these quantities.

Acknowledgements

We thank Drs. Yuji Uozumi and Kotaro Yokawa of the National Research Institute of Far Seas Fisheries, Shimizu, Japan, and Dr. Pierre Kleiber of the Pacific Islands Fisheries Science Center, NMFS, Honolulu, Hawaii, USA, for providing the Japanese and the Hawaii-based longline swordfish catch, effort and length–frequency data, respectively. We also thank two

anonymous reviewers for their comments on an earlier version of this manuscript. This study was in part supported financially by the National Science Council, Taiwan through grant NSC91-2313-B-002-330 to Chi-Lu Sun.

Appendix A. Population dynamics model

A.1. Basic sex-specific population dynamics

The population dynamics are governed by an age- and sex-structured model with an annual time-step. Natural mortality occurs continuously throughout the year and the fishery occurs instantaneously in the middle of the year, i.e.:

$$N_{t,a}^s = \begin{cases} R_t^s & \text{if } a = 0 \\ (N_{t-1,a-1}^s e^{-M/2} - C_{t-1,a-1}^s) e^{-M/2} & \text{if } 0 < a < \lambda \\ [(N_{t-1,\lambda-1}^s + N_{t-1,\lambda}^s) e^{-M/2} - (C_{t-1,\lambda-1}^s + C_{t-1,\lambda}^s)] e^{-M/2} & \text{if } a = \lambda \end{cases} \quad (\text{A.1})$$

where $N_{t,a}^s$ is the number of fish of age a and sex s (m/f) at the start of year t , R_t^s the recruitment (at age 0) of fish of sex s at the start of year t , $C_{t,a}^s$ the catch in number of fish of age a and sex s during year t :

$$C_{t,a}^s = \sum_f C_{t,a}^{s,f} \quad (\text{A.2})$$

$C_{t,a}^{s,f}$ is the catch in number of fish of age a and sex s by fleet f during year t , M the instantaneous rate of natural mortality on fish of both sexes, and λ the maximum age (treated as a plus-group).

A.2. Catches

The number of fish of age a and sex s caught by fleet f during year t can be calculated using the equation, i.e.:

$$C_{t,a}^{s,f} = F_{t,a}^{s,f} N_{t,a}^s e^{-M/2} \quad (\text{A.3})$$

where $F_{t,a}^{s,f}$ is the exploitation rate on fish of age a and sex s during year t by fleet f :

$$F_{t,a}^{s,f} = s_a^{s,f} F_t^f \quad (\text{A.4})$$

F_t^f is the exploitation rate by fleet f during year t on fully selected animals:

$$F_t^f = \frac{C_t^{f,\text{obs}}}{\sum_a \sum_s s_a^{s,f} N_{t,a}^s e^{-M/2}} \quad (\text{A.5})$$

$C_t^{f,\text{obs}}$ is the observed total catch (in number) by fleet f during year t , and $s_a^{s,f}$ the selectivity for fleet f on fish of age a and sex s .

A.3. Growth and maturity

The relationship between length and weight is given by the equation:

$$w_a^s = A^s (L_a^s)^{B^s} \quad (\text{A.6})$$

$$\begin{cases} \text{if } a = 0 \\ \text{if } 0 < a < \lambda \\ \text{if } a = \lambda \end{cases} \quad (\text{A.1})$$

where A^s , B^s are the parameters of the length–weight relationship for animals of sex s , and L_a^s the expected length of an animal of sex s and age a :

$$L_a^s = L_\infty^s (1 - e^{-k^s(1-m^s)(a-a_0^s)})^{1/(1-m^s)} \quad (\text{A.7})$$

L_∞^s , k^s , a_0^s , m^s are the parameters of the generalized von Bertalanffy growth equation.

Maturity as a function of age is modeled by means of a logistic curve, i.e.:

$$\phi_a = \frac{1}{1 + \exp[r_m(L_a^f - L_m)]} \quad (\text{A.8})$$

where ϕ_a is the fraction of females of age a that are mature, L_m the length-at-50%-maturity for females, and r_m the maturity slope parameter.

A.4. Recruitment

Recruitment (number of animals of age 0) is assumed to be related to the spawning stock biomass by means of the [Beverton and Holt \(1957\)](#) stock–recruitment relationship, parameterized in terms of the ‘steepness’ of the stock–recruitment relationship, h ([Francis, 1992](#)) and the recruitment at unfished

pre-exploitation equilibrium, R_0 , i.e.:

$$R_t^s = 0.5 \frac{4hR_0S_t/S_0}{(1-h) + (5h-1)S_t/S_0} e^{v_t - \sigma_v^2/2} \quad (\text{A.9})$$

where S_0 is the spawning stock biomass at unfished pre-exploitation equilibrium, S_t the spawning stock biomass at the start of year t , i.e.:

$$S_t = \sum_{a=1}^{\lambda} \phi_a w_a^f N_{t,a}^f \quad (\text{A.10})$$

v_t is a normally distributed process error, $v_t \sim N(0, \sigma_v^2)$, and σ_v^2 the (assumed) variance of the process error in recruitment.

The sex-ratio at age 0 is assumed to be 1:1 in the absence of evidence to the contrary.

A.5. Selectivity

Selectivity as a function of age is determined using a general functional form which combines a logistic and a dome-shaped component:

$$s_a^{s,f} = \frac{s_a^{1,s,f} + \varphi^f s_a^{2,s,f}}{\max_{a'}(s_{a'}^{1,s,f} + \varphi^f s_{a'}^{2,s,f})} \quad (\text{A.11})$$

where $s_a^{1,s,f}$ is the logistic component:

$$s_a^{1,s,f} = \left[1 + \exp \left(-\ln 19 \frac{L_a^s - L_{50}^f}{L_{95}^f - L_{50}^f} \right) \right]^{-1} \quad (\text{A.12})$$

L_{50}^f is the length-at-50%-selectivity for fleet f for the logistic component of the selectivity ogive, L_{95}^f the length-at-95%-selectivity for fleet f for the logistic component of the selectivity ogive, $s_a^{2,s,f}$ the dome-shaped component:

$$s_a^{2,s,f} = \frac{1}{\sqrt{2\pi} L_{sd}^f} \exp \left[\frac{-(L_a^s - L_{mu}^f)^2}{2(L_{sd}^f)^2} \right] \quad (\text{A.13})$$

L_{mu}^f is the mode of the dome-shaped component of the selectivity ogive for fleet f , L_{sd}^f the standard deviation of the dome-shaped component of the selectivity ogive for fleet f , and φ^f , for fleet f , the weight assigned to the dome-shaped component of the selectivity function.

A.6. Initial conditions

The initial conditions correspond to a population at its unfished pre-exploitation equilibrium at the start of the first year for which catches are available, i.e.:

$$N_{t_1,a}^s = \begin{cases} 0.5R_0 & \text{if } a = 0 \\ N_{t_1,a-1}^s e^{-M} & \text{if } 0 < a < \lambda \\ N_{t_1,\lambda-1}^s e^{-M}/(1 - e^{-M}) & \text{if } a = \lambda \end{cases} \quad (\text{A.14})$$

where t_1 is the first year for which catch data are available, i.e. $t_1 = 1952$.

A.7. Fleet-aggregated fishing intensity

A measure of overall fishing intensity for the most recent 11 years (when all six fleets took catches of swordfish) is calculated by dividing the total catch by a fleet-averaged exploitable biomass:

$$F_t = \frac{\sum_f C_t^{f,obs}}{\sum_a \sum_s s_a^s N_{t,a}^s e^{-M^s/2}} \quad (\text{A.15})$$

F_t is the fleet-aggregated fishing intensity during year t , and s_a^s the ‘fleet-aggregated’ selectivity for fish of age a and sex s , calculated by weighting the fleet-specific selectivity patterns by the fleet-specific exploitation rates:

$$s_a^s = \frac{\sum_f \left(\sum_t F_t^f \right) s_a^{s,f}}{\sum_{f'} \sum_{t'} F_{t'}^{f'}} \quad (\text{A.16})$$

where the summations over year relate to the last 11 years (1990–2000).

Appendix B. Contributions to the objective function

B.1. The sex-ratio data

The sex-ratio data are included in the likelihood function by assuming that the observed estimates of the sex-ratio of the catch are normally distributed about the model predictions. The contribution of the sex-ratio data to the negative of the logarithm of the likelihood function (ignoring constants independent of the model

parameters) is therefore:

$$L_2 = \sum_t \sum_f \rho^s (Q_t^{f, \text{obs}} - Q_t^f)^2 \quad (\text{B.1})$$

where $Q_t^{f, \text{obs}}$ is the observed sex-ratio for fleet f during year t :

$$Q_t^{f, \text{obs}} = \frac{C_t^{f, f, \text{obs}}}{\sum_{s=1}^2 C_t^{s, f, \text{obs}}} \quad (\text{B.2})$$

$C_t^{s, f, \text{obs}}$ is the observed catch of fish of sex s by fleet f during year t , Q_t^f the model-predicted sex-ratio for fleet f during year t :

$$Q_t^f = \frac{\sum_a C_{t,a}^{f, f}}{\sum_{a'} \sum_{s=1}^2 C_{t,a'}^{s, f}} \quad (\text{B.3})$$

ρ^s is the weight assigned to the sex-ratio data (set to 0.1).

B.2. The length–frequency data

The likelihood function assumed for the length–frequency data is the robust normal formulation of Fournier et al. (1990), adjusted so that the variance of the predictions is based on the observed rather than the model-predicted fractions (B. Ernst, Univ. Washington, pers. commun.). Sex-specific and sex-aggregated length–frequency data are available. Denoting the sex-aggregated data (i.e. the length–frequency data for the fleets JPNW, JPSW and JPNE for all years and for the fleet JPSE before 1987) as sex $s=0$, the contribution of the length–frequency data to the negative of the log-likelihood function (ignoring constants independent of the model parameters) is:

$$-\ln L_3 = 0.5 \sum_t \sum_l \sum_{s=0}^2 \sum_f \left\{ \rho_t^f \left[\frac{(P_{t,l}^{s, f, \text{obs}} - P_{t,l}^{s, f})^2}{2 (P_{t,l}^{s, f, \text{obs}} [1 - P_{t,l}^{s, f, \text{obs}}] + \frac{1}{N_l}) (\tau_t^{s, f})^{-2}} + 0.01 \right] \right\} \quad (\text{B.4})$$

where $P_{t,l}^{s, f, \text{obs}}$ is the observed fraction that fish in length-class l made up of the catch (in numbers) by fleet f of animals of sex s during year t , $P_{t,l}^{s, f}$ the model-estimate of the fraction that fish in length-class l made up of the catch (in numbers) by fleet f of animals of sex

s during year t :

$$P_{t,l}^{s, f} = \frac{C_{t,l}^{s, f}}{\sum_{l'} C_{t,l'}^{s, f}} \quad (\text{B.5})$$

$C_{t,l}^{s, f}$ is the model-estimate of the fraction that fish in length-class l made up of the catch (in numbers) by fleet f of animals of sex s during year t :

$$C_{t,l}^{s, f} = \begin{cases} \sum_{a=1}^{\lambda} \Lambda_{a,l}^s C_{t,a}^{s, f} & \text{if } s = m/f \\ \sum_s \sum_{a=1}^{\lambda} \Lambda_{a,l}^s C_{t,a}^{s, f} & \text{otherwise} \end{cases} \quad (\text{B.6})$$

N_l is the number of length-classes, $\tau_t^{s, f}$ the ‘effective’ sample size for year t , fleet f and sex s (the maximum of the number of animals measured and 50), $\Lambda_{a,l}^s$ the fraction of animals of age a and sex s that are in length-class l :

$$\Lambda_{a,l}^s = \int_{L_l - \Delta L}^{L_l + \Delta L} \frac{1}{\sqrt{2\pi}\sigma_a^s} \exp \left[-\frac{(L - L_a^s)^2}{2(\sigma_a^s)^2} \right] dL \quad (\text{B.7})$$

σ_a^s is the standard deviation of the length of a fish of sex s and age a (Table 2), L_l the mid-point of length-class l , ΔL the width of half of a length-class (the length–frequency data were aggregated to 5 cm length-classes owing to the way the original data were collected), ρ_t^f the weight assigned to the length–frequency data for year t and fleet f :

$$\rho_t^f = \rho^L \frac{C_t^{f, \text{obs}}}{(\sum_t C_t^{f, \text{obs}}) / n^f} \quad (\text{B.8})$$

ρ^L is the overall (pre-specified) weight assigned to the length–frequency data (set to 0.01), and n^f the number of years for which length–frequency data are available.

B.3. The catch-rate data

The catch-rate-based indices of abundance are assumed to be log-normally distributed about the corresponding model predictions. Ignoring constants independent of the model parameters, the contribution of the catch-rate data to the negative of the log-likelihood function is:

$$-\ln L_4 = \sum_f \sum_t \left\{ \ln \tilde{\sigma}^f + \frac{(\ln I_t^{f,obs} - \ln I_t^f)^2}{2(\tilde{\sigma}^f)^2} \right\} \quad (\text{B.9})$$

where $I_t^{f,obs}$ is the observed catch-rate index for year t and fleet f , I_t^f the model-estimate corresponding to $I_t^{f,obs}$:

$$I_t^f = q^f B_t^f \quad (\text{B.10})$$

q^f the catchability coefficient for fleet f , B_t^f the number of exploitable animals during year t corresponding to fleet f :

$$B_t^f = \sum_s \sum_{a=1}^{\lambda} s_a^{s,f} N_{t,a}^s e^{-M/2} - C_{t,a}^{s,f} / 2 \quad (\text{B.11})$$

$\tilde{\sigma}^f$ is the standard deviation of I_t^f , assumed to be time-invariant.

B.4. Constraint contribution to the objective function

A constraint is placed on the deviations about the stock–recruitment relationship:

$$\frac{1}{2\sigma_v^2} \sum_t v_t^2 \quad (\text{B.12})$$

References

- Anon, 1997. Report of the ICCAT Swordfish Stock Assessment Session. ICCAT Coll. Vol. Sci. Pap. 46 (3), 165–291.
- Anon, 1999. Report of the swordfish working group meeting. In: Proceedings of the Second Meeting of the Interim Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), ISC2/99/PLEN/11, Honolulu, Hawaii, USA, January 15–23, p. 22 pp.
- Anon, 2002. Report of the swordfish working group meeting. In: Proceedings of the Third Meeting of the Interim Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), ISC3/PLEN/09, Nagasaki, Japan, January 28–30, p. 21 pp.
- Bartoo, N.W., Coan, Jr., A.L., 1989. An assessment of the Pacific swordfish resource. In: Stroud, R.H. (Ed.), Planning the Future of Billfishes: Research and Management in the 90s and Beyond. Part 1: Fishery and Stock Synopses, Data Needs and Management. Proceedings of the Second International Billfish Symposium, Kailua-Kona, Hawaii, August 1–5, 1988. Natl. Coalition Mar. Conserv., Savannah, GA, pp. 137–153.
- Beverton, F.J.H., Holt, S.J., 1957. On the dynamics of exploited fish population. Fisheries Investigations from the Ministry of Agriculture Fisheries and Food of Great Britain, Series II, vol. 19, 533 pp.
- Chen, Y., 1996. A Monte Carlo study on impacts of the size of subsample catch on estimation of fish stock parameters. Fish. Res. 26, 207–223.
- Cope, J.M., Piner, K., Mente-Vera, C.V., Punt, A.E., 2004. Status and Future Prospects for the Cabezon (*Scorpaenichthys marmoratus*) as Assessed in 2003. Pacific Fishery Management Council, Portland, OR, 157 pp.
- DeMartini, E.E., Uchiyama, J.H., Williams, H.A., 2000. Sexual maturity, sex ratio, and size composition of swordfish, *Xiphias gladius*, caught by the Hawaii-based pelagic longline fishery. Fish. Bull. 98, 489–506.
- Di Nardo, G.T., Kwok, W., 1998. Spatiotemporal dynamics of swordfish, *Xiphias gladius*, landings in the Hawaii-based North Pacific pelagic longline fishery. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 51–64.
- FAO, 1995. Precautionary approach to fisheries. Part 1: Guidelines on the precautionary approach to capture fisheries and species introductions. FAO. Fish. Tech. Pap. 350/1.
- Fournier, D.A., Sibert, J.R., Majkowski, J., Hampton, J., 1990. MULTIFAN: a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). Can. J. Fish. Aquat. Sci. 47, 301–317.
- Fournier, D.A., Hampton, J., Sibert, J.R., 1998. Multifan-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Can. J. Fish. Aquat. Sci. 55, 2105–2116.
- Francis, R.I.C.C., 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. 49, 922–930.
- Gavaris, S., 1980. Use of a multiplicative model to estimate catch rate and effort from commercial data. Can. J. Fish. Aquat. Sci. 37, 2272–2275.
- Gavaris, S., 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29, 12 pp.
- Hampton, J., 2000. Natural mortality in tropical tunas: size really does matter. Can. J. Aquat. Sci. 55, 1002–1010.
- Hanan, D.A., Holts, D.B., Coan, A.L., 1993. The California drift gill net fishery for sharks and swordfish, 1981–82 through 1990–91. Fish Bull. Ser. 175, 95.

- Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment and Management: Choice, Dynamics and Uncertainty. Chapman and Hall, New York, 570 pp.
- Hinton, M.G., Deriso, R.B., 1998. Distribution and stock assessment of swordfish, *Xiphias gladius*, in the eastern Pacific Ocean from catch and effort data standardized on biological and environmental parameters. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 161–180.
- Holts, D., Sosa-Nishizaki, O., 1998. Swordfish, *Xiphias gladius*, fisheries of the eastern North Pacific Ocean. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 65–76.
- Ito, R.Y., Dollar, R.A., Kawamoto, K.E., 1998. The Hawaii-based longline fishery for swordfish, *Xiphias gladius*. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 77–88.
- Kimura, D.K., 1981. Standardized measures of relative abundance based on modelling log(c.p.u.e.), and the application of Pacific Ocean perch (*Sebastes alutus*). J. Cons. Int. Explor. Mer. 39, 211–218.
- Kleiber, P., Bartoo, N., 1998. Standardizing swordfish, *Xiphias gladius*, longline catch per unit of effort using general additive models. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 181–194.
- Kleiber, P., Yokawa, K., 2002. MULTIFAN-CL assessment of swordfish in the North Pacific. In: Proceedings of the 15th Meeting of the Standing Committee on Tuna and Billfish, SCTB15 Working Paper BBRG-3, p. 15.
- Kleiber, P., Yokawa, K., 2004. MULTIFAN-CL assessment of swordfish in the North Pacific. In: Proceedings of the Third Swordfish Working Group of the Interim Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), ISC/04/SWO-WG/07, p. 32.
- Maunder, M.N., 2001. A general framework for integrating the standardization of catch per unit of effort into stock assessment models. Can. J. Fish. Aquat. Sci. 58, 795–803.
- Maunder, M.N., Langley, A.D., 2004. Integrating the standardization of catch-per-unit-of-effort into stock assessment models: testing a population dynamics model and using multiple data types. Fish. Res. 70, 389–395.
- Methot, R.D., 2000. Technical description of the stock synthesis assessment program. U.S. Department of Commerce, NOAA Tech. Memm., NMFS-NWFSC-43, pp. 1–56.
- Nakamura, I., 1985. FAO Species Catalogue: vol. 5. Billfishes of the World. An annotated and illustrated catalogue of marlins, sailfishes, spearfishes and swordfishes known to date. FAO Fish. Synop. vol. 125, No. 5, 65 pp.
- Nakano, H., 1998. Stock status of Pacific swordfish, *Xiphias gladius*, inferred from CPUE of the Japanese longline fleet standardized using general linear models. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 195–218.
- Palko, B.J., Beardsley, G.L., Richards, W.J., 1981. Synopsis of the biology of the swordfish, *Xiphias gladius* Linnaeus. U.S. Dep. Commer., NOAA Tech. Rep. NMFS CIRC-441, 21 pp.
- Patterson, K.R., Kirkwood, G.P., 1995. Comparative performance of ADAPT and Laurec–Shepherd methods for estimating fish population parameters and in stock management. ICES J. Mar. Sci. 52, 183–196.
- Pauly, D., 1980. On the interrelationships between natural mortality, growth parameters and mean environmental temperatures in 175 fish stocks. J. Cons. Int. Mer. 39, 175–192.
- Pope, J.G., Shepherd, J.G., 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Int. Explor. Mer. 42, 129–151.
- Punt, A.E., 1995. The performance of a production-model management procedure. Fish. Res. 21, 349–374.
- Punt, A.E., 1997. The performance of VPA-based management. Fish. Res. 29, 217–243.
- Punt, A.E., Campbell, R.A., Smith, A.D.M., 2001. Evaluating empirical indicators and reference points for fisheries management: application to the broadbill swordfish fishery off eastern Australia. Mar. Freshwater Res. 52, 819–832.
- Punt, A.E., Smith, A.D.M., Cui, G., 2002. Evaluation of management tools for Australia's South East Fishery. 2. How well can management quantities be estimated? Mar. Freshwater Res. 53, 631–644.
- Quinn II, T.J., Turnbull, C.T., Fu, C., 1998. A length-based population model for hard-to-age invertebrate populations. In: Funk, F., Quinn II, T.J., Heifetz, J.N., Ianelli, J.E., Powers, J.F., Schweigert, P.J., Sullivan, Zhang, C.I. (Eds.), Fishery Stock Assessment Models. University of Alaska Sea Grant College Program Report No. AK-SG-98-01, Fairbanks, pp. 531–556.
- Sakagawa, G.T., 1989. Trends in fisheries for swordfish in the Pacific Ocean. In: Stroud, R.H. (Ed.), Planning the Future of Billfishes: Research and Management in the 90s and Beyond. Part 1: Fishery and Stock Synopses, Data Needs and Management. Proceedings of the Second International Billfish Symposium, Kailua-Kona, Hawaii, August 1–5, 1988. Natl. Coalition Mar. Conserv., Savannah, GA, pp. 61–79.
- Sampson, D.B., Yin, Y., 1998. A Monte Carlo evaluation of the stock synthesis assessment program. In: Funk, F.T.J., Quinn II, T.J., Heifetz, J.N., Ianelli, J.E., Powers, J.F., Schweigert, P.J., Sullivan, Zhang, C.I. (Eds.), Fishery Stock Assessment Models. University of Alaska Sea Grant College Program Report No. AK-SG-98-01, Fairbanks, pp. 315–338.
- Shepherd, J.G., Pope, J.G., 2002. Dynamic pool models II. Short term and long term forecasts of catch and biomass. In: Hart, P.J., Reynolds, J.D. (Eds.), Handbook of Fish Biology and Fisheries, vol. 2. Blackwell Publishing, pp. 164–188.
- Sigler, M.F., 1999. Estimation of sablefish, *Anoplopoma fimbria*, abundance off Alaska with an age-structured population model. Fish. Bull. 97, 591–603.
- Skillman, R.A., 1998. Central Pacific swordfish, *Xiphias gladius*, fishery development, biology, and research. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 101–124.
- Sun, C.L., Wang, S.P., Yeh, S.Z., 2002. Age and growth of the swordfish (*Xiphias gladius* L.) in the waters around Taiwan determined from anal-fin rays. Fish. Bull. 100, 822–835.
- Sosa-Nishizaki, O., Shimizu, M., 1991. Spatial and temporal CPUE trends and stock unit inferred from them for the Pacific swordfish caught by the Japanese tuna longline fishery. Bull. Far. Seas. Fish Res. Lab. 28, 75–89.
- United Nations, 1995. Structure and Process of the 1993–1995 United Nations Conference on Straddling Fish Stocks and Highly Migratory Fish Stocks. FAO Fish. Circ., vol. 898.

- Uosaki, K., 1998. Standardized CPUE of North Pacific swordfish, *Xiphias gladius*, in the Japanese large-mesh driftnet fishery. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 142, pp. 125–132.
- Wang, S.P., Sun, C.L., Yeh, S.Z., 2003. Sex ratios and sexual maturity of swordfish (*Xiphias gladius* L.) in the waters of Taiwan. Zool. Stud. 42, 529–539.
- Wang, S.P., 2004. Sex-specific age-structured model for stock assessment of swordfish (*Xiphias gladius*) in the North Pacific Ocean. PhD Thesis. National Taiwan University, 170 pp.
- Yamanaka, H., 1958. Variability of hooked-rates for broadbill swordfish in the North Pacific. Rep. Nankai Reg. Fish. Res. Lab. 7, 37–48.
- Zheng, J., Kruse, G.H., Fair, L., 1998. Using multiple data sets to assess red king crab in Norton Sound, Alaska: length-based stock synthesis. In: Funk, F., Quinn II, T.J., Heifetz, J.N., Ianelli, J.E., Powers, J.F., Schweigert, P.J., Sullivan, Zhang, C.I. (Eds.), Fishery Stock Assessment Models. University of Alaska Sea Grant College Program Report No. AK-SG-98-01, Fairbanks, pp. 591–612.