Attempt of stock assessment of the Indian Ocean swordfish by production models based on the Bayesian averaging method

Toshihide Kitakado[1*] and Tom Nishida[2]

[1]Tokyo University of Marine Science and Technology, 5-7, Konan 4, Minato, Tokyo, 108-8477, Japan
[2]National Research Institute of Far Seas Fisheries, 5-7-1, Orido, Shimizu, Shizuoka, 424-8633, Japan
*Corresponding author's email address: kitakado@kaiyodai.ac.jp

Abstract

Advanced age-structured models and synthesis analyses are getting common in the fishery stock assessment, but these are sometimes sensitive to assumptions made. Meanwhile, simple and traditional production models such as Schaefer and Fox models require a few assumptions and therefore they inherently tend to be robust but might be less informative though. This is an important feature when not so much information/data are available for a target stock, in which cases even selection of a better production model has uncertainty to some extent. In this paper, an attempt of model averaging is introduced to address the model uncertainty. Among possible procedures of such an attempt, Bayesian model averaging is employed. The posterior probabilities of models, which are used as weights to respective models, are derived through a Sampling and Importance Resampling (SIR) method. The results based on the Japanese CPUE series showed that the whole Indian stock is in a state of healthy; the total biomass ratio is significantly greater than 1 and the F ratio is clearly less than 1. It should be noted that the results are still preliminary and more evaluation might be necessary.

Note. During the meeting, the analysis was updated based on revised and additional CPUE series. The results are given in the Addendum.

Introduction

In fishery stock assessment, it is now common to use age-structured (or length-based) models and/or stock synthesis approaches like SS3. In these approaches, in addition to some reliable abundance indices, extra information and assumptions are required. Adding such information and assumptions could improve the estimation performance and hence the fishery resource management, but it might not be the case if the assumptions are inappropriate or results are quite sensitive to alternative assumptions. In this sense, traditional but less complicated production models have not lost their reputes although there is also uncertainty in the choice of appropriate models.

In this paper, an attempt of model averaging is introduced to address the model uncertainty. Before this exercise, some initial considerations on statistical analysis will be made through likelihood inference. And then, Bayesian model averaging is employed although there might be room to improve.

Data

The following data set were employed in this exercise.

1. Swordfish catch statistics 1950-2009
   (“New Estimated Catch” in “NC_ALL.zip” downloaded from the IOTC website).
2. A series of standardized CPUE for 1980-2009 estimated with a GLM analysis (scaled to make the average 1.0) and their associated standard errors (from Nishida et al. 2011).

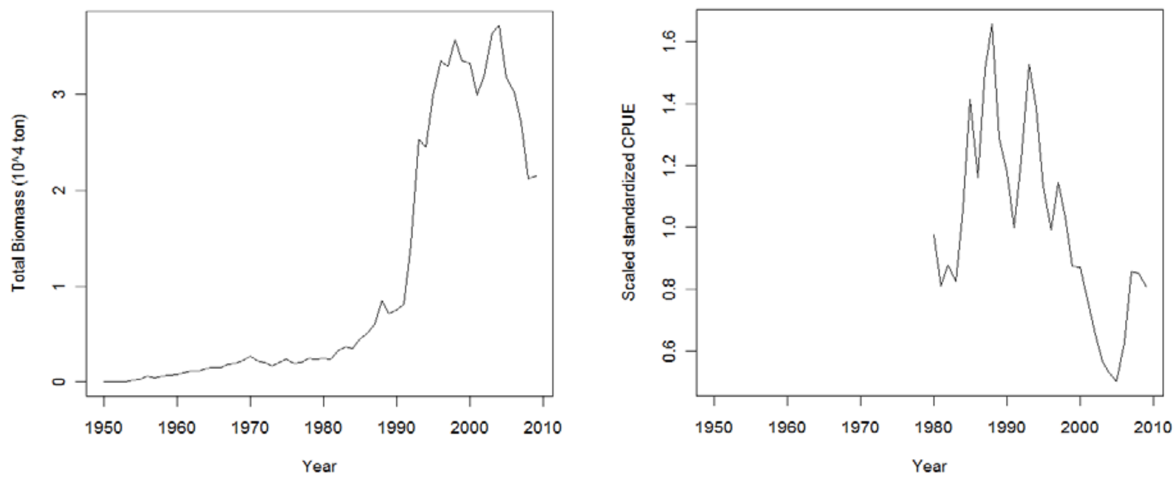


Figure 1. Catch statistics and a series of standardized CPUE series for swordfish.

Basic population dynamics

In this paper, we assume only the two different production models, Schaefer and Fox models. Here, let P_t and C_t denote the population biomass at the beginning of year *t* and catch in year *t*, respectively. Then, あ the following basic recursive formula is supposed to hold.

$$P_t = P_{t-1} + P_{t-1} f(P_{t-1}) - C_{t-1}. \qquad (1)$$

Schaefer and Fox models respectively have $f(P_t) = r(1 - P_t / K)$ and $f(P_t) = r(\log K - \log P_t)$. With the depletion rate, which is a ratio of P_t relative to the carrying capacity *K* and denoted as $D_t = P_t / K$, the formula above is re-expressed as

$$\begin{aligned} &\text{(Schafer)} \quad D_t = D_{t-1} + rD_{t-1}(1 - D_{t-1}) - C_{t-1} / K, \\ &\text{(Fox)} \qquad D_t = D_{t-1} - rD_{t-1} \log D_{t-1} - C_{t-1} / K. \end{aligned} \qquad (2)$$

The deterministic recursive formula is sometimes extended to a model with the process error like below:

$$P_t = \left\{ P_{t-1} + P_{t-1} f(P_{t-1}) - C_{t-1} \right\} e^{u_t}, \quad u_t \sim N(0, \tau^2) \qquad (3)$$

where u_t is so-called a process error term, which accounts for stochasticity in the dynamics. The process error can be further extended to have auto-correlation, which can be sometimes occurred due to serial correlation of environmental conditions.

Initial considerations through "likelihood inference"

1. Observation errors other than sampling errors in standardized CPUE estimates (data)

The standardized CPUEs are employed as data for the estimation of production model. These standardized CPUEs, of course, have uncertainty and therefore the standard errors are associated with the estimates. Let I_t and cv_t respectively be estimated standardized CPUE index and its coefficient of variation in year *t*. Let $\tilde{I}_t$ be the true abundance index in year *t*.

$$I_t = \tilde{I}_t \, \mathrm{e}^{u_t}, \quad u_t \sim N(0, cv_t^2) \qquad (4)$$

However, there should be another stochasticity expressing the model error,

$$\tilde{I}_t = qKD_t e^{v_t}, \quad v_t \sim N(0, cv_{add}^2) \qquad (5)$$

and therefore

$$\log I_t \sim N(\log qKD_t, \, cv_t^2 + cv_{add}^2). \qquad (6)$$

Comparison of results in the runs "sc0" and "sc0ad", which difference is in presence/absence of the additional variance, clearly showed that the incorporation of the additional cv dramatically decreases the AIC value although the point estimates were almost same (see Table 1). The difference in the fitness was reflected in the values of standard errors (assessed by the inverse of Hessian matrix). This simple exercise justifies the assumption of additional variance.

2. Depletion level in the initial year considered

According to the catch history of the Indian Ocean swordfish, the amounts of catch during and before

1950's seems negligible, which implies a rationale to assume the total biomass being the carrying capacity in 1950, the initial year of the catch record.

Three different assumptions are made in run sets "sc0ad", "sc1ad" and "sc2ad" for the Schaefer production model (see Table 1). Only "sc2ad" include a specific parameter expressing an initial depletion in the initial year, so it has an additional parameter for that. The results in "sc0ad" and "sc1ad" showed the determination of the initial year gave a little impact on estimates and AIC as far as assuming the population stayed at its carrying capacity (see Table 1 and Figure 2). This result is consistent with the fact of less catch before 1980 although the depletion level in 1980 should be less than 1. However, in the best-fitted model "sc2ad", the depletion level in 1980 was estimated at about the half of carrying capacity because the model explained an increasing trend in CPUE in early 1980's although it does not make sense because the stock was far from over-exploitation.

One possible reason for such CPUE trend is due to some difficulty in the original catch and effort data. In addition, change in catchability and/or carrying capacity around those years is other possible reason. This problem warrants further investigation, but in this paper we will try to capture the dynamics from 1950, when the stock was considered virgin. It should be noted that the standard error for the estimate of carrying capacity is quite large.

Table 1. Summaries of estimation results for justifying the use of additional cv and also investigating the sensitivity to the assumption of initial depletion.

Run	sc0		sc0ad		sc1ad		sc2ad	
Production	SC		SC		SC		SC	
Additional CV	None		Yes		Yes		Yes	
Initial year	1950		1950		1980		1980	
Initial depletion	B1950 = K		B1950 = K		B1980 = K		B1980 < K	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
r	0.348	0.027	0.373	0.282	0.407	0.307	0.393	0.110
K	33.679	2.108	31.75	19.55	29.67	18.80	28.56	6.20
q	0.036	0.002	0.039	0.025	0.041	0.028	0.048	0.011
D1950	1.000							
D1980	0.980	0.000	0.981	0.002	1.000		0.561	0.085
D2009	0.580	0.008	0.581	0.074	0.589	0.075	0.511	0.058
cv_add	0.000		0.214	0.028	0.216	0.028	0.179	0.023
loglike	-1545.51		31.12		30.84		36.41	
#parameters	3		4		4		5	
AIC	3097.02		-54.23		-53.68		-62.83	

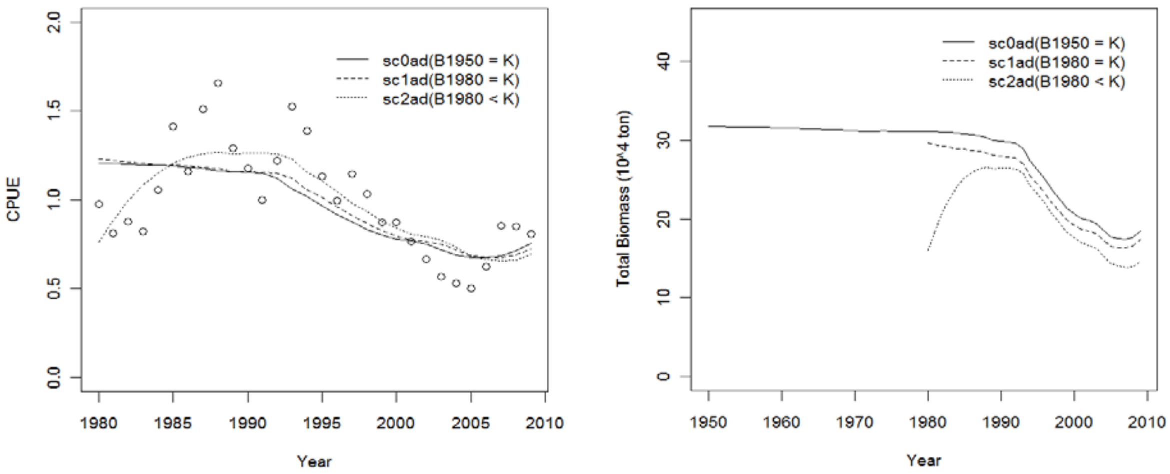


Figure 2. Fitness of the CPUE data to the models and estimated population trajectories for the three different assumptions for initial depletion.

3. *Variants in the production model*

As introduced earlier, the two alternative production models are employed in this study. Under the likelihood framework, a well-known model selection criterion is usually used. The Table 2 and Figure 3 show that analyses based on these two models provided different consequences in the trajectories and future projection although the AIC values between the two models are rather similar. This fact urges us to consider some model uncertainty (even model selection uncertainty) and motivates to attempt the "model averaging" for the stock assessment.

Table 2. Estimation of results based on two different production models.

Run	sc0ad		fx0ad	
Production	SC		FX	
Additional CV	Yes		Yes	
Initial year	1950		1950	
Initial depletion	B1950 = K		B1950 = K	
	Estimate	SE	Estimate	SE
r	0.373	0.282	0.234	0.149
K	31.75	19.55	38.03	18.99
q	0.039	0.025	0.033	0.017
D1950				
D1980	0.981	0.002	0.975	0.003
D2009	0.581	0.074	0.570	0.069
cv_add	0.214	0.028	0.213	0.028
loglike	31.12		31.24	
#parameters	4		4	
AIC	-54.23		-54.48	

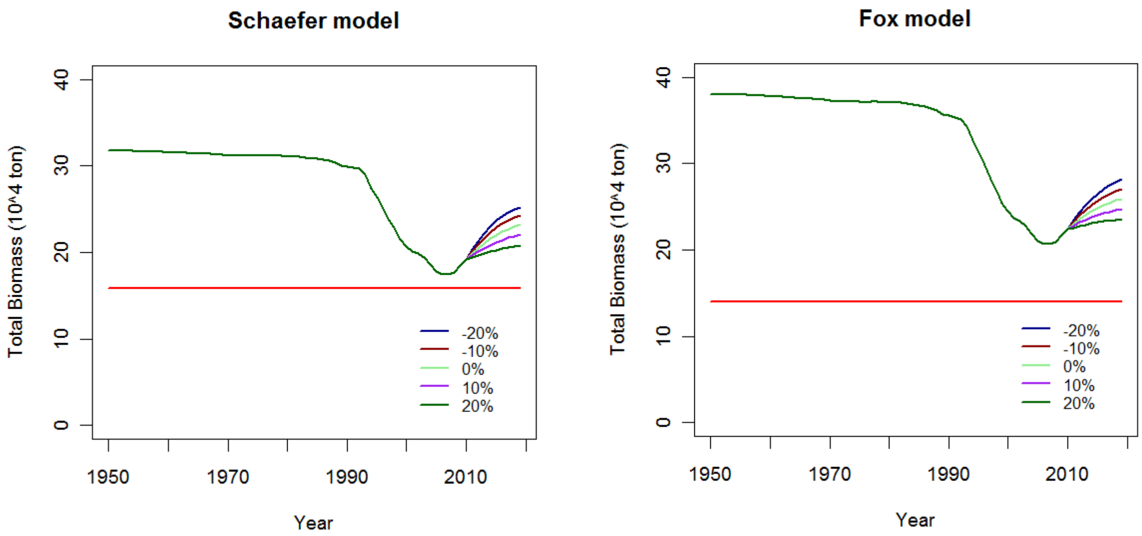


Figure 3. Estimated trajectories of the total biomass (1950-2009) and future projection (2010-2019) by five catch scenarios (-20%, -10%, 0%, 10% and 20% changes from the 2009 catch level) under the two different production models.

4. *Incorporation of process errors*

As shown in Equation (3), the process errors are further potential stochasticity to be considered. The errors are random quantities and therefore those should be integrated out from the likelihood according to their distribution assumption. Here, the maximization of the integrated likelihood was conducted via ADMB-RE (Skaug and Fournier 2006).

The results are summarized in Table 3 and Figure 4. The results indicated that the models with the process errors dramatically improved the fitting. More interestingly and importantly,

1) the standard errors for the population parameters were decreased by incorporation of process errors because the models with only the additional variance assigned the deviations to the parameter’s standard errors while the those with both the process errors could share the deviation with the observation errors;
2) the depletion in 1980 was a somewhat reasonable level (not to be so close to 1.0) in process error models;
3) the process error models provide too well fits (it might be difficult to distinguish the additional cv and the extent of process error in this case) and high variation in the population size.;

4) the process error models give chances of overshoot in the population trajectories, which means that the depletion level can exceed the carrying capacity;
5) incorporation of auto-correlation into the process error distribution made the convergence difficult and also brought quite different biomass levels (we will not pursue this model any further in this paper).

Here, the model with only the additional cv will be employed because of highly stochastic nature in population size for the process error models and difficulty in separation of the additional cv and process error.

Table 3. Comparison of results based on different error assumptions.

Run	sc0		sc0ad		sc0pr		sc0adpr		sc0adau	
Additional CV	None		Yes		None		Yes		Yes	
Process error	None		None		Yes		Yes		Auto-correlated	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
r	0.348	0.027	0.373	0.282	0.427	0.187	0.414	0.127	0.617	0.252
K	33.679	2.108	31.75	19.55	34.54	17.98	35.61	15.15	25.64	13.09
q	0.036	0.002	0.039	0.025	0.035	0.020	0.034	0.016	0.047	0.027
D1950	1.000		1.000		1.000		1.000		1.000	
D1980	0.980	0.000	0.981	0.002	0.803	0.114	0.805	0.096	0.806	0.105
D2009	0.580	0.008	0.581	0.074	0.667	0.420	0.670	0.425	0.671	0.496
cv_add			0.214	0.028			0.018	0.038	0.000	0.000
tau (process)					0.152	0.020	0.152	0.020	0.145	0.019
auto-correlation									0.396	0.248
loglike	-1545.51		31.12		94.71		94.74		96.10	
#parameters	3		4		4		5		6	
AIC	3097.02		-54.23		-181.41		-179.49		-180.20	

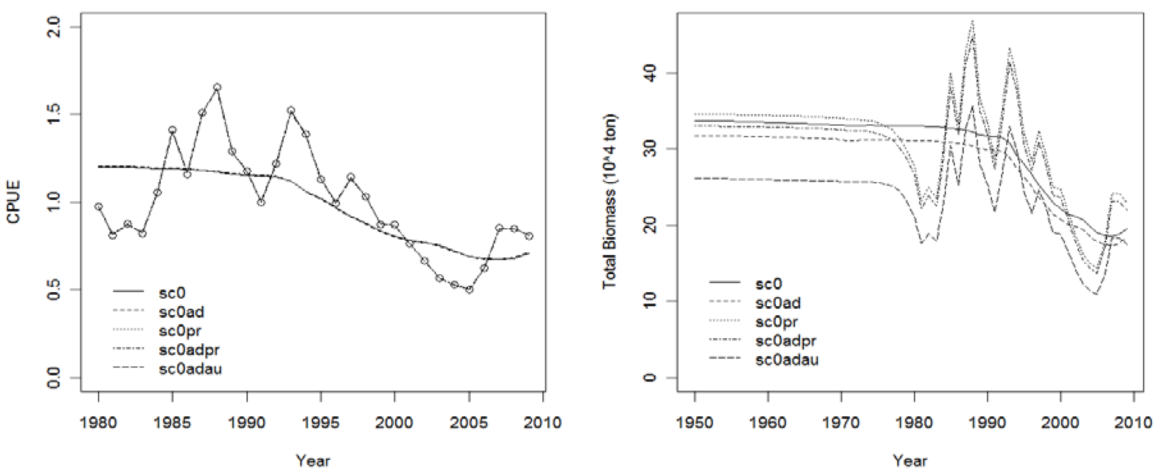


Figure 4. Estimated CPUEs and Population trajectories for the five alternative models with different error assumptions.

Bayesian estimation and attempt of Bayesian model averaging

Bayesian inference is especially promising when the information of data is poor but some auxiliary or past information on parameters are available. Bayesian inference is also useful to construct flexible models, express hierarchical structures, avoid difficult optimization by using simulation procedures such as a Markov Chain Monte Carlo (MCMC) and Sampling and Importance Resampling (SIR), and let the data tell probabilistically about parameters, important quantities and validity of models etc through their posterior distributions.

When it comes to getting posterior distribution of models, things are not straightforwards even when using MCMC because computation of marginal likelihood based on MCMC outputs tends to be unstable. To overcome this difficulty, several approaches have been developed (e.g. Chib's method and a reversible jump method).

Bayesian estimation

In this paper, we employ the SIR for the model without assuming the prcoess errors. Also, we only consider flat priors for all the parameters as a non-informative prior. It should be remarked that choice of appropriate prior distributions is one of key issues (including better parameterization of models). In the absence of any positive reasons, such non-information prior is used although the definition of it is sometimes ambiguous.

For a proposal distribution in the initial sampling in the SIR, a multivariate normal distribution with the maximum likelihood (ML) estimates (in log-scale) as the mean and their associated errors as the covariance is assumed. This might be a good approximation of the posterior distribution. The initial samples are randomly sampled with replacement according to the weight, which consider the Jacobian for the log-transformation. We set at 200,000 and 10,000 for the initial and secondary sample sizes, respectively.

Estimation results with the future projection are given in Figure 5. The two different models provided comparable but a little different results as in the case of ML estimation. For example, the median biomass in late 1990's was below the MSY level for the Schaefer model while that for the Fox model was above the MSY level. Also, the credible interval for the Fox model is wider than that for the Schaefer model. These two different outcomes are averaged in the next subsection.

(a) Schaefer model (b) Fox model

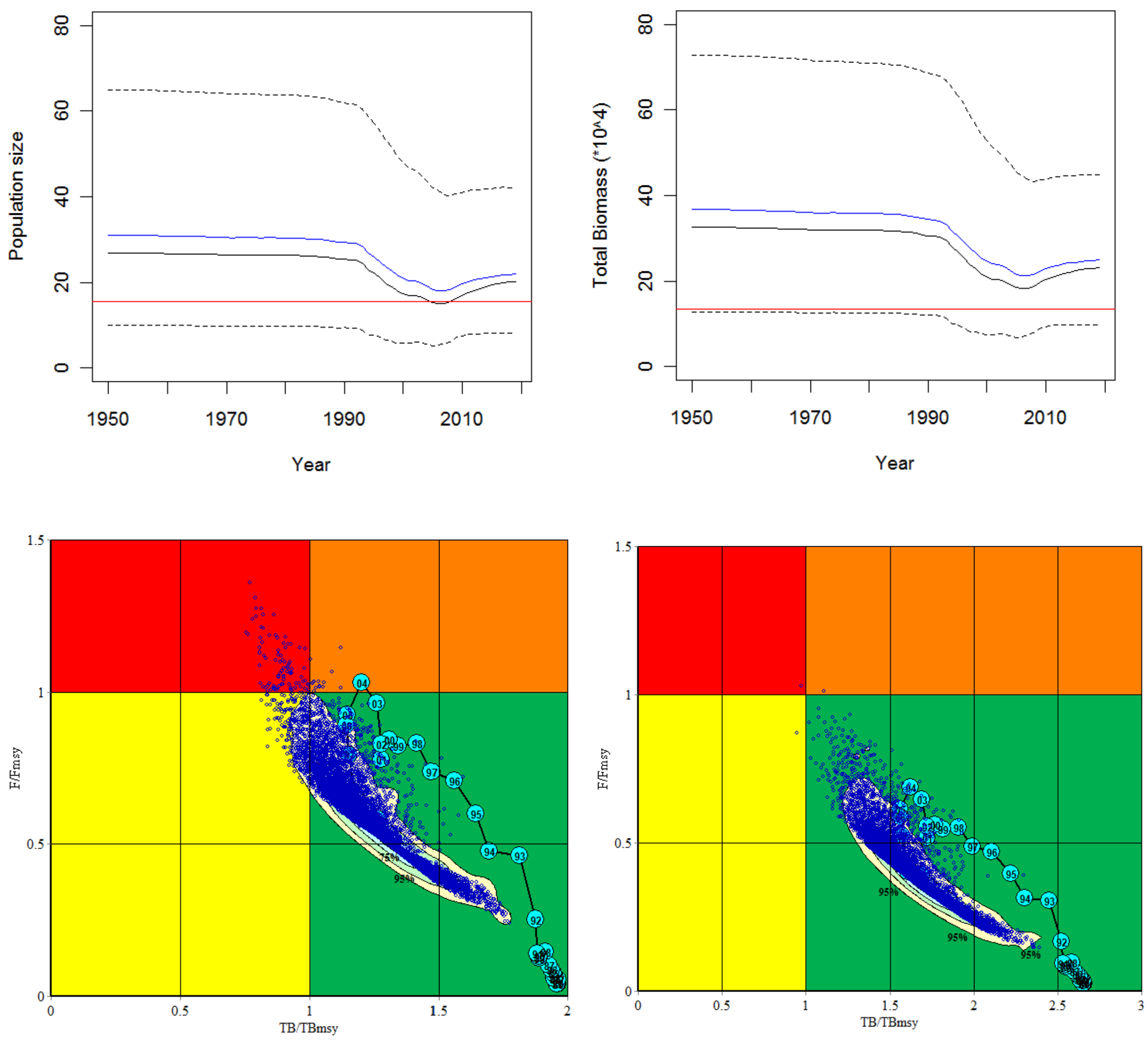


Figure 5. Results of Bayesian inference. (a) Schaefer model; (b) Fox model. Biomass levels in 2010-2019 were computed under the assumption of the same catch level as in 2009. The solid and shaded lines show the median and 90% CI and the blue solid line shows the posterior mean.

Model averaging

The posterior probability of each model (Schaefer or Fox) is evaluated by the Importance Sampling (IS) given below:

$$
\begin{aligned}
& f(Data \mid Model) \\
&= \int f(Data, \theta \mid Model)\pi(\theta \mid Model)\, d\theta \\
&= \int \frac{f(Data, \theta \mid Model)\pi(\theta \mid Model)}{g(\theta)} g(\theta) d\theta \\
&\approx \sum \frac{f(Data, \theta_i^* \mid Model)\pi(\theta_i^* \mid Model)}{g(\theta_i^*)}, \quad \theta_i^* \sim g(\theta)
\end{aligned} \tag{7}
$$

For a proposal distribution of the IS, $g(\theta)$, we use the asymptotic distribution of the ML estimates of in log-scale and assess the marginal distribution with importance weights with consideration of parameter transformation. The sample size was set at 1,000,000. Then, the posterior model probabilities are estimated as

$$
\begin{aligned}
& P(Model_{SC} \mid Data) = \frac{f(Data \mid Model_{SC})}{f(Data \mid Model_{SC}) + f(Data \mid Model_{FX})} \\
& P(Model_{FX} \mid Data) = 1 - P(Model_{SC} \mid Data)
\end{aligned} \tag{8}
$$

The estimated posterior probabilities of the Schaefer and Fox models were 0.488 and 0.512, which were almost even. This result is fairly consistent with so-called Akaike weights, 0.469 and 0.531, which is defined based on AIC values. Summaries for the averaging exercise are shown in Table 4 and Figures 6 and 7.

Table 4. Comparison of model specific and model averaging results.

Model	Schaefer		Fox		Averaging	
	Estimate	SE	Estimate	SE	Estimate	SE
r	0.580	0.404	0.372	0.290	0.474	0.248
K	30.97	18.23	36.70	20.76	33.90	14.06
MSY	3.093	0.445	3.534	0.727	3.319	0.433
Bmsy	15.48	9.12	13.50	7.64	14.47	6.010
Fmsy	0.290	0.202	0.372	0.290	0.332	0.179
B2009	19.05	11.50	22.31	13.60	20.72	9.08
F2009	0.148	0.075	0.123	0.062	0.135	0.049
B2009/Bmsy	1.264	0.175	1.689	0.234	1.420	0.270
F2009/Fmsy	0.581	0.161	0.391	0.131	0.484	0.104

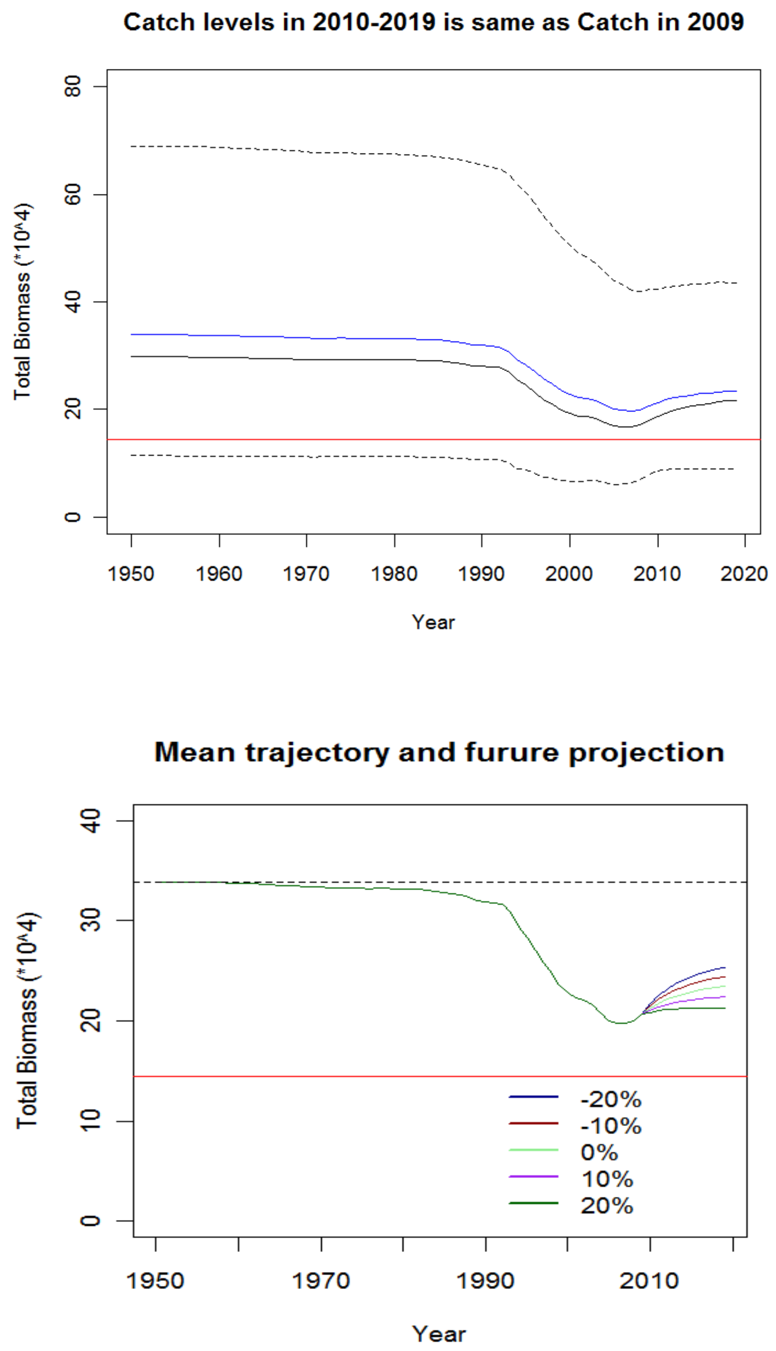


Figure 6. Results of Bayesian model averaging. The top panel shows the population trajectory and future projection given the same catch level with 2009 (solid and shaded lines show the median and 90% credible interval and the red horizontal line shows the MSY level). The bottom panel shows mean future projections under the different catch level (the red and shaded lines are respectively for the MSY level and carrying capacity).

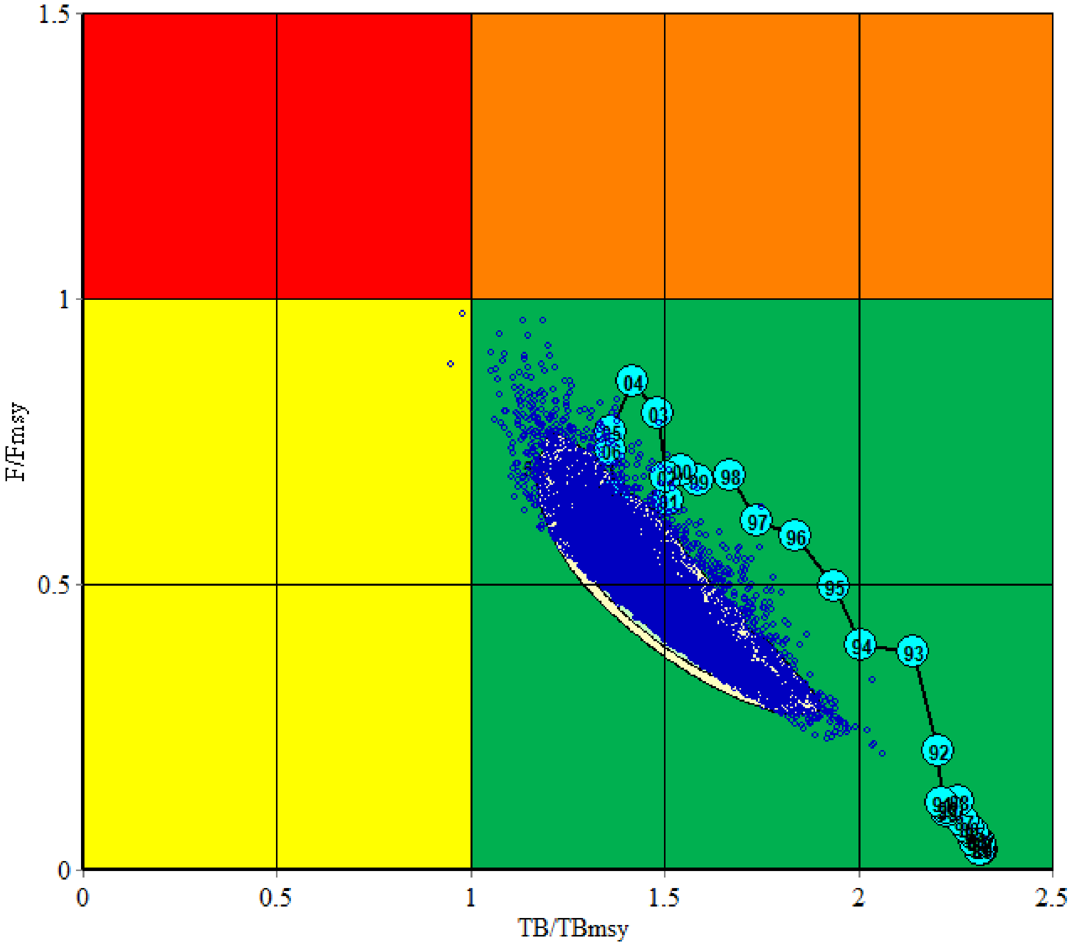


Figure 7. KOBE I for the results of Bayesian model averaging.

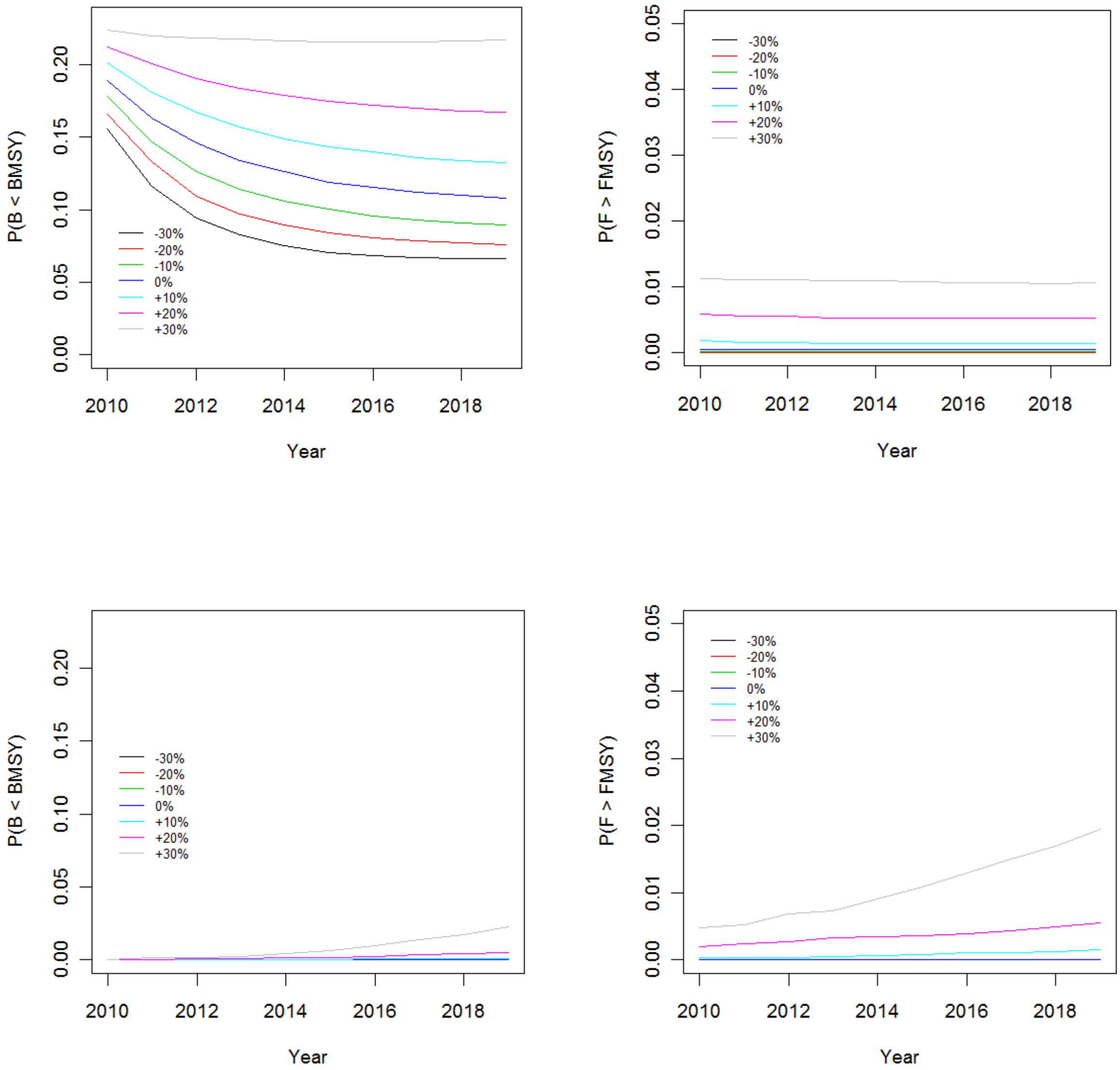


Figure 8. Future projection of probabilities for KOBE II, P(B<BMSY) and P(F>FMSY), basd on the results of Bayesian model averaging. The values of probabilities change according to their definition. The upper two figures are for the case of fixed BMSY and FMSY (as the posterior means) while the bottom ones are computed as the average of chances B<BMSY and F>FMSY in posterior outputs (this might be better) .

Discussion

This paper dealt with an issue of model uncertainty through a model averaging approach. As shown in this paper, the approach has potential to take the model uncertainty into account when considering the management advice.

However, we admit that our work might be preliminary because 1) we consider averaging of only "two models"; 2) impact of choice of priors are not discussed; 3) only Japanese CPUE series are used; etc.

1) The assumption of initial depletion can also be considered in averaging if necessary. And, in theory, further models can also be taken into account.

2) Assumption of prior distributions may be influential when information of the data is really poor. We avoid any discussion about this, but the resultant estimates are not so different from the ML estimation (except for the intrinsic rate of increase in the Schaefer model), and therefore the flat priors are not so problematic.

3) The anaylysis conducted here can be extend to use other CPUE series. It should be noted that, even when using only the CPUE series from Japanese fisheries, the range of years should be carefully considered because of smaller CPUE values in early 1980's than in late 1980's. Consideration of change in catchability in the assessment is one possible ways (change point is chosen using AIC), but it introduces an additional parameter. Another way is just to delete the CPUE data in early 1980's (but only as a sensitivity test).

Finally, regarding the model averaging, a frequentist method, bootstrapped weight, is an alternative possible approach if we need to avoid the discussion about prior choice although a better bias correction should be invented. Even in Bayesian framework, a reversible jump is another possible way. More comprehensive analyses will be conducted by us in the near future.

Acknowledgement

The leading author (T. Kitakado) is very grateful to Dr. Alejandro Anganuzzi (IOTC) and the secretariat of IOTC for their providing funding support for my participation. The authors also thank Mr. Miguel Herrera (IOTC) to provide the updated nominal catch record.

References

Nishida, T. and Kitakado, T. (2011) Estimation of the Abundance Index (AI) of swordfish (*Xiphias gladius*) in the Indian Ocean (IO) based on the fine scale catch and effort data of the Japanese tuna longline fisheries (1980-2010). A paper IOTC-2011-WPB09-14.

Skaug, H.J. and Fournier, D. (2006) Automatic Approximation of the Marginal Likelihood in non-Gaussian Hierarchical Models; Computational Statistics and Data Analysis. 51: 699-709.

ADDENDUM

The analysis was for the putative stock in SW region and for the whole stock in the Indian Ocean. The CPUE series employed in these studies are given below.

SW region only:

1) from Japan (1980-2009, spatial effects with 5 degree bands)
2) from Taiwan (1995-2009, spatial effects with 5 degree bands)
3) from Spain (2001-2009, Run 5 in WPB09-23)
4) from Re Union (1994-2000)

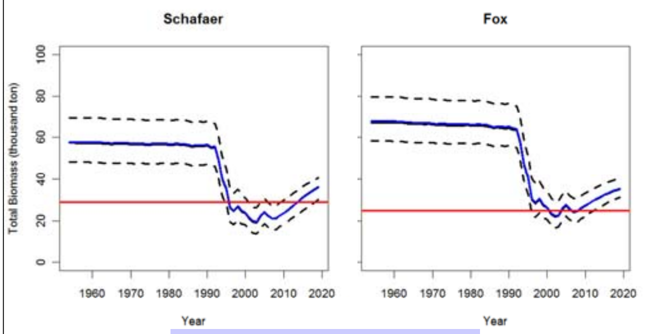
Whole Indian Ocean

1) from Japan (1980-2009, spatial effects with 5 degree bands)
2) from Taiwan (1995-2009, spatial effects with 5 degree bands)
3) from Spain (2001-2009, Run 1 in WPB09-23)

The variance of model error was assumed to be constant across the CPUE series above although different catchability coffeficients were incorporated. The results are given in the following slides.

Results for SW Region

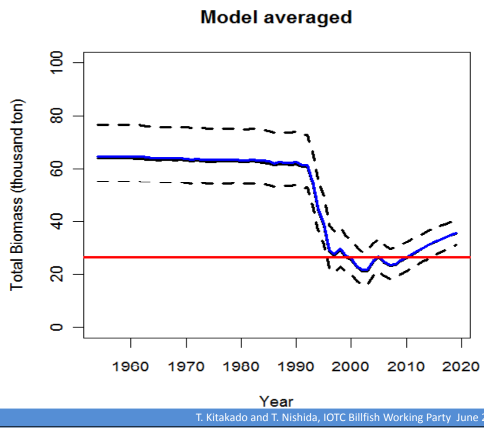
Results for SW (1) Model specific



$$\omega_{Schaefer} = P(Model_{SC} \mid Data) = 0.308$$

$$\omega_{Fox} = P(Model_{FX} \mid Data) = 0.692$$

Results for SW (2) Model averaging

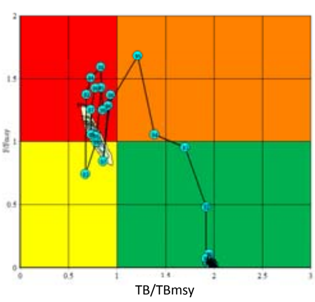


Results for SW (3) Reference points

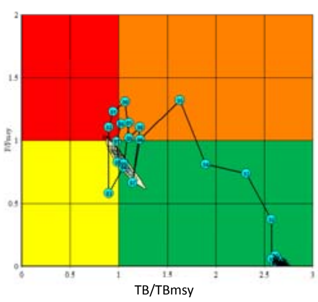
Management Quantity	**SW Region Only**
Most recent catch estimate	
Mean catch over last 5 years	
MSY (1000 t)	7.91 (SE=0.199)
Current Data Period	Catch:1954-2009 CPUE:1980-2009
F(Current)/F(MSY)	0.884 (SE=0.071)
B(Current)/B(MSY)	0.942 (SE=0.071)
SB(Current)/SB(MSY)	NA
B(Current)/B(0)	0.375 (SE=0.028)
SB(Current)/SB(0)	NA
B(Current)/B(Current, F=0)	what?
SB(Current)/SB(Current, F=0)	NA

Results for SW (4) KOBE (by model)

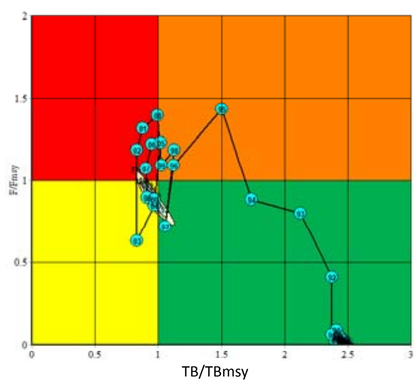
(a)Schaefer model



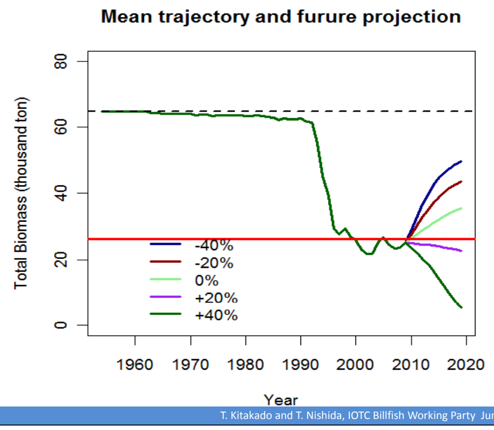
(b) Fox model



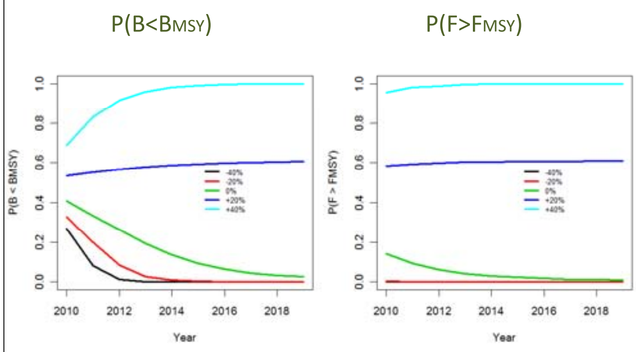
Results for SW (5) KOBE (averaged)



Results for SW (7) Future catch



Results for SW (8) Risk Assessment

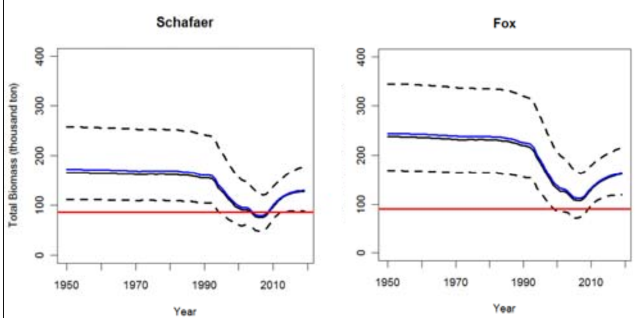


Results for SW (9) KOBE II matrix

	Constant Catch Level (relative to 2009)				
Probability	**60%**	**80%**	**100%**	**120%**	**140%**
B(2012) <B(MSY)	0.011	0.083	0.264	0.567	0.915
F(2012) >F(MSY)	0.000	0.000	0.060	0.597	0.989
B(2019) <B(MSY)	0.000	0.000	0.025	0.606	1.000
F(2019) >F(MSY)	0.000	0.000	0.010	0.609	1.000

Results for Aggregated Indian Ocean

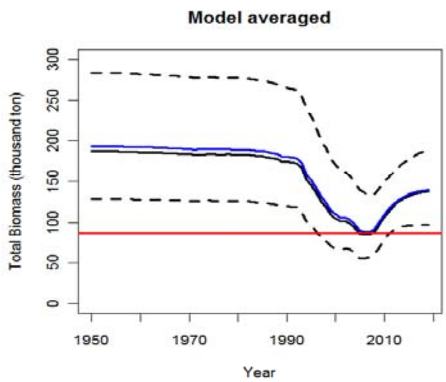
Results for IO (1) Model specific



$$\omega_{Schaefer} = P(Model_{SC} \mid Data) = 0.699$$

$$\omega_{Fox} = P(Model_{FX} \mid Data) = 0.301$$

Results for IO (2) Model averaging

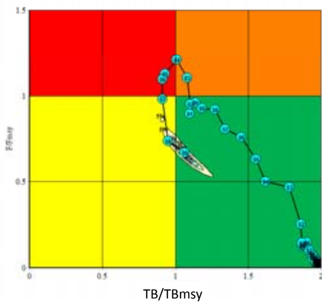


Results for IO (3) Reference points

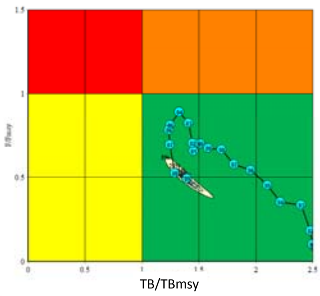
Management Quantity	**Aggregate Indian Ocean**	**SW Region Only**
Most recent catch estimate		
Mean catch over last 5 years		
MSY (1000 t)	30.77 (SE=0.880)	7.91 (SE=0.199)
Current Data Period	Catch:1950-2009 CPUE:1980-2009	Catch:1954-2009 CPUE:1980-2009
F(Current)/F(MSY)	0.615 (SE=0.053)	0.884 (SE=0.071)
B(Current)/B(MSY)	1.073 (SE=0.090)	0.942 (SE=0.071)
SB(Current)/SB(MSY)	NA	NA
B(Current)/B(0)	0.481 (SE=0.043)	0.375 (SE=0.028)
SB(Current)/SB(0)	NA	NA
B(Current)/B(Current, F=0)		
SB(Current)/SB(Current, F=0)	NA	NA

Results for IO (4) KOBE (by model)

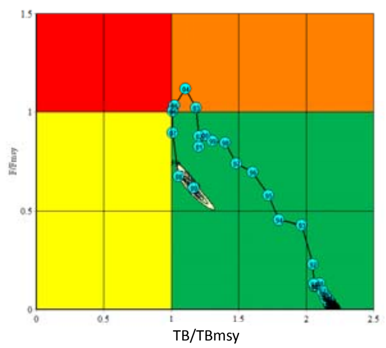
(a)Schaefer model



(b) Fox model

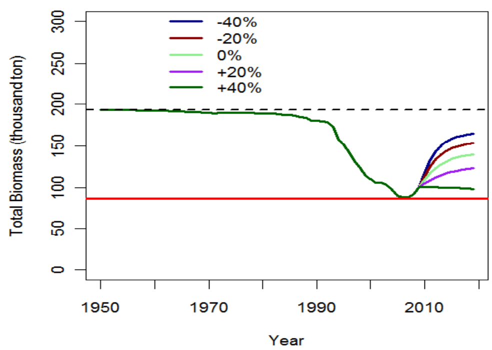


Results for IO (5) KOBE (averaged)



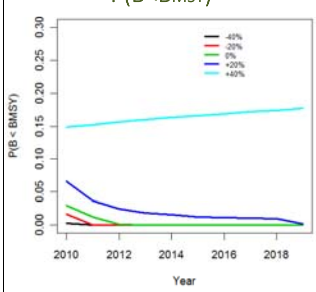
Results for IO (6) Future catch

Mean trajectory and furure projection

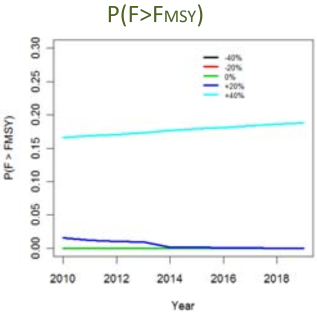


Results for IO (7) Risk Assessment

P(B<BMSY)



P(F>FMSY)



Results for IO (8) KOBE II matrix

	Constant Catch Level (relative to 2009)				
Probability	**60%**	**80%**	**100%**	**120%**	**140%**
B(2012) <B(MSY)	0.000	0.000	0.001	0.024	0.157
F(2012) >F(MSY)	0.000	0.000	0.000	0.011	0.171
B(2019) <B(MSY)	0.000	0.000	0.000	0.002	0.178
F(2019) >F(MSY)	0.000	0.000	0.000	0.011	0.188