**Diagnoses for stock synthesis model on yellowfin tuna in the Indian Ocean**

Takayuki Matsumoto[1], Hiroki Yokoi[1], Keisuke Satoh[1] and Toshihide Kitakado[2]

[1]*National Research Institute of Far Seas Fisheries (NRIFSF), Japan Fisheries Research and Education Agency, 5-7-1, Orido, Shimizu, Shizuoka, 424-8633, Japan*
[2]*Tokyo University of Marine Science and Technology, 5-7, Konan 4, Minato, Tokyo, 108-8477, Japan.*

**Abstract**

Diagnoses of stock synthesis model SS3 were conducted using that for Indian Ocean yellowfin tuna stock assessment in 2016, based on the methods used in this year's ICCAT bigeye tuna assessment. Jitter analysis, residual analysis, retrospective analysis, R0 likelihood profile and age structured production model (ASPM) analysis were conducted. According to the results of these diagnoses, the model seems to be comparatively good and robust. The diagnoses are useful, and so hopefully will be applied to other stocks as well.

## 1. Introduction

Nowadays Stock Synthesis model SS3 (Methot 2009; Methot and Wetzel 2013) is often and commonly used for stock assessment of tuna and tuna-like species, and is also used for many stocks in the Indian Ocean. SS3 model enables many options, for example, weighting of several components (size data, CPUE, and so on) for input data or biological parameters. It is desirable to select the models with best performance as reference (base) models. To do so, it is necessary to check the performance of the models.

Recently, Yokoi et al. (2018) conducted several kinds of diagnoses for SS3 model on this year's ICCAT Atlantic Ocean bigeye tuna stock assessment. They used several methods including cutting edge ones, and reported that in the process of exploring reference case, the performance of the models got better mainly by downweighting size data, and that stock assessment was regarded as successful.

In this study, the methods by Yokoi et al. (2018) were applied to SS3 model on IOTC Indian Ocean yellowfin tuna stock assessment in 2016 (Langley, 2016; IOTC, 2016).

## 2. Materials and methods

Model diagnoses were conducted based on Stock Synthesis (SS3), which is one of the integrated stock assessment models. We used input files for the reference model adopted at Indian Ocean yellowfin tuna stock assessment in 2016 (Langley, 2016). The model used four areas and one CPUE index in each area. The detail of each diagnosis is as follows. Diagnoses were conducted using R program code or by manually changing SS3 input files.

### *2.1 Jitter analysis*

Jitter analysis is a diagnosis that measures the stability of the model, and it examines whether the estimation result changes by repeating the execution several times by randomly changing the initial value (maximum likelihood estimation value) of the parameter. The model is considered as stable if (negative) log likelihood values are similar among runs. In this study, 50 times of runs were conducted and negative log likelihood values were compared.

### *2.2 Residual analysis*

The standard deviation of the normalized residuals (SDNR) and root mean square error (RSME) were calculated and examined. SDNR is a measure of the fit to the data that is independent of the number of data points. A relatively good model fit will be characterized by smaller residuals (i.e. close to zero) and a SDNR close to 1. Francis (2011) notes that it is also necessary to conduct a visual examination between observed and predicted values to be sure that the fit is good even when SDNR values are not much greater than 1. CPUE fit based on SDNR was also created and examined. RSME is measure of the differences between estimated and observed values. If the RSME is small, fit is good.

### *2.3 Retrospective analysis*

Retrospective analysis is a diagnosis that measures the stability of a model by the estimation result changes with data removed from recent years. In an unstable model, the estimation result fluctuates each time the data is removed.

### *2.4 R0 likelihood profile*

In SS3, the virgin biomass (B0) is estimated from the virgin recruitment (R0) using the stock recruitment relationship; that is, R0 determines the scale of the resource amount. The R0 likelihood component profiles diagnosis fixing the virgin recruitment at different values and plotting the negative log-likelihood values for each data component against this parameter (Maundur and Piner 2015). If the model is good, the minima of negative log-likelihood has approximately the same R0 value among data components. Different minima among data components indicates possible conflict in the data sources about scale of resource amount. The R0 likelihood component profile does not indicate which data is correct. However, it is assumed that CPUE will indicate correct R0 value because CPUE seems to have greater information on scale of resource amount than the other data components (Francis 2011). In this study, the difference of log-likelihood value from the minimum one within the range of R0 examined was plotted and considered.

### *2.5 ASPM analysis*

This approach was reported by Maunder and Piner (2015), Felipe et al. (2017) and Minte-Vera et al. (2017). This diagnosis consists of comparing the results of 'ASPM' setting to those from original integrated analysis. ASPM setting is constructed with the selectivity fixed at the values estimated in the base case of SS3. If the ASPM cannot explain the CPUE, either the stock is recruitment driven, resources are not affected by catch because they are many enough, the model is incorrect, or the CPUE is not proportional to abundance (Maunder and Piner 2015). In addition, alternative models of ASPM (i.e. ASPM with recruitment estimated and ASPM with the recruitment set equal to the values from the base case: named as "aspm_fix" and "aspm_est", respectively) are also compared to examine the influence of recruitment.

## 3. Results

### *3.1 Jitter analysis*

**Fig. 1** shows the results of jitter analysis. Negative log likelihood values have some difference among different runs (maximum >100), indicating that the initial value of the original model was not very good and so the model was not very robust.

### *3.2 Residual analysis*

**Fig. 2** shows CPUE fit based on SDNR analysis with the values of SDNR and RMSE. Generally CPUE fit was good. The values of SDNR and RMSE were not so much different among scenarios. SDNR values were around 1.0, indicating that the values were similar to observation error and so the model was good. The values of RMSE seem small enough to regard that CPUE fit was good.

### *3.3 Retrospective analysis*

Fig. 3 shows the results of retrospective analysis (change in B/Bmsy and F/Fmsy). These show that the trends of B/Bmsy and F/Fmsy were similar to those for base model except for near the terminal year, indicating that the model was robust.

### *3.4 R0 likelihood profile*

Fig. 4 shows the R0 likelihood profile. Although the graphs are not smooth, minimum values of each components are at or close to R0=11.80, indicating that there is little conflict among components and so the model is good.

### *3.5 ASPM diagnosis*

**Fig.** 5 shows the results of comparison of CPUE fit based on ASPM diagnosis. These show that overall trend of CPUE fit was similar between base case and ASPM diagnosis. In all the scenarios, "aspm_fix" was almost the same as base case, indicating that length data have information on only selectivity curves and estimation of recruitment in the model, in other words the size composition data have no much information about the stock abundance. Large difference between base case and ASPM was observed for fleet 28 and 29 around 2000. This is probably because of insufficient size data during that period.

**Fig. 6** shows the result of ASPM diagnosis (trend of spawning stock biomass). This shows that overall trend by "aspm" was similar to that from the base case with small differences, indicating that the information on the selectivity curves based on size data was not so much lost and only the information on the estimation of recruitment was deleted.

**4. Discussion**

We consider that the diagnostic tools used in this study are very useful to check the performance of SS3 models. These methods seem to be useful. The same methods are expected to be applied to analyses for other stocks in the Indian and other oceans. This time we used SS3 files for the past stock assessment. However, in the near future hopefully these diagnoses will be conducted before selecting base case and making management advice.

**5. References**


Felipe, C., Punt, A. E., Chang, Y. J., Maunder, M. N., Piner, K. R. (2017) Can diagnostic tests help identify model misspecification in integrated stock assessment? Fisheries Research. 192: 28-40.

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

IOTC. (2016) Report of the 18th Session of the IOTC Working Party on Tropical Tunas. IOTC–2016–WPTT18–R[E]. pp 126.

Langley, A. (2016) An update of the 2015 Indian Ocean Yellowfin Tuna stock assessment for 2016. IOTC-2016-WPTT18-27. pp 14.

Maunder, M. N. and Piner, K. R. (2015) Contemporary fisheries stock assessment: many issues still remain. ICES Journal of Marine Science. 72 (1): 7-18.

Methot, R. D. (2012) User Manual for Stock Synthesis: model version 3.24f. Oct 3, 2012. NOAA Fisheries Service, Seattle, WA, USA.

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

Minte-Vera, C. V., Maunder, M. N., Aires-da-Silva A. M., Satoh K., Uosaki K. (2017) Get the biology right, or use size-composition data at your own risk. Fisheries Research. 192: 114-125.

Walter J., Yokoi, H., Satoh, K., Matsumoto, T., Ijurco, A. U., Oritz, M., Sharma, R., Mosquiera, I. and Schirripa, M. (2018) Atlantic bigeye tuna stock assessment in stock synthesis. ICCAT SCRS-2018-111, 67p.

Yokoi, H., Satoh, K., Matsumoto, T. and Kitakado, T. (2018) Model diagnosis for stock synthesis on bigeye tuna in the Atlantic Ocean. ICCAT SCRS-2018-140, 38p.

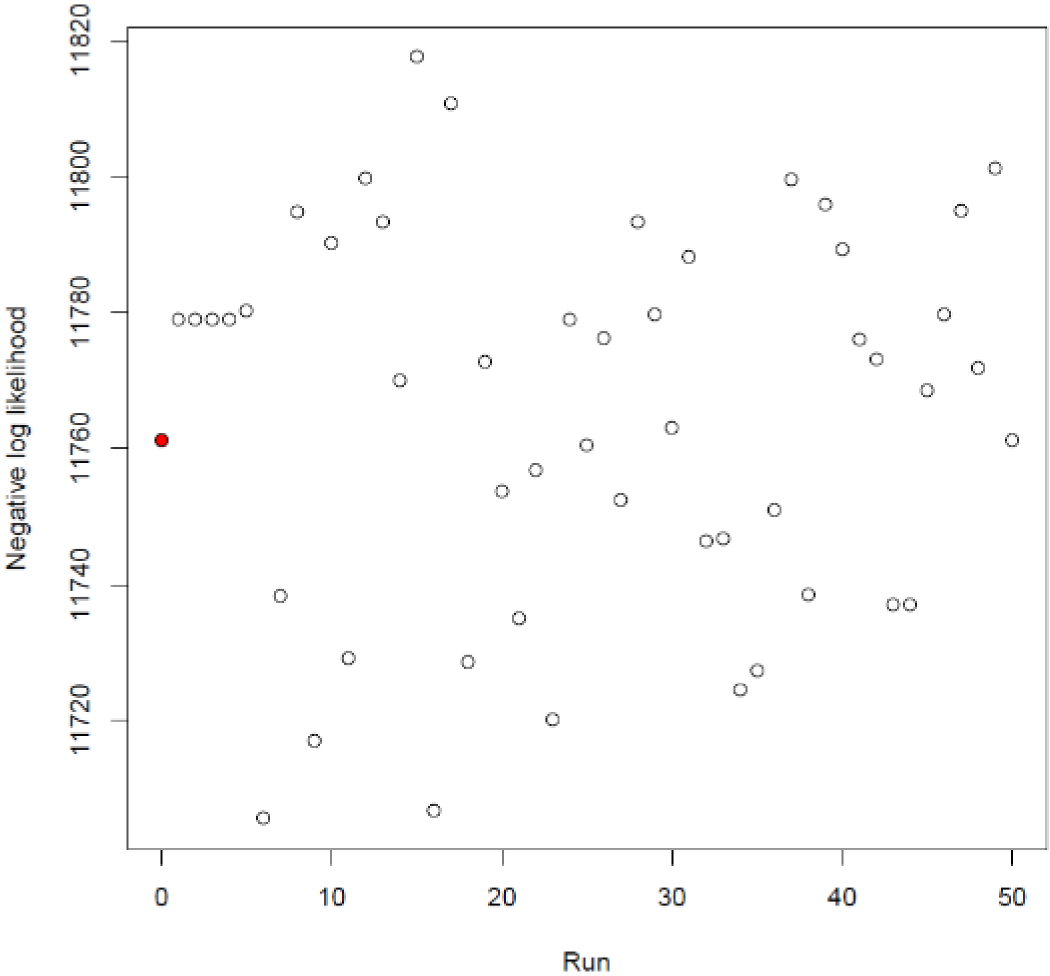


**Fig. 1.** Results of jitter analysis (comparison of negative log likelihood values). Red circle indicates the result of base model.

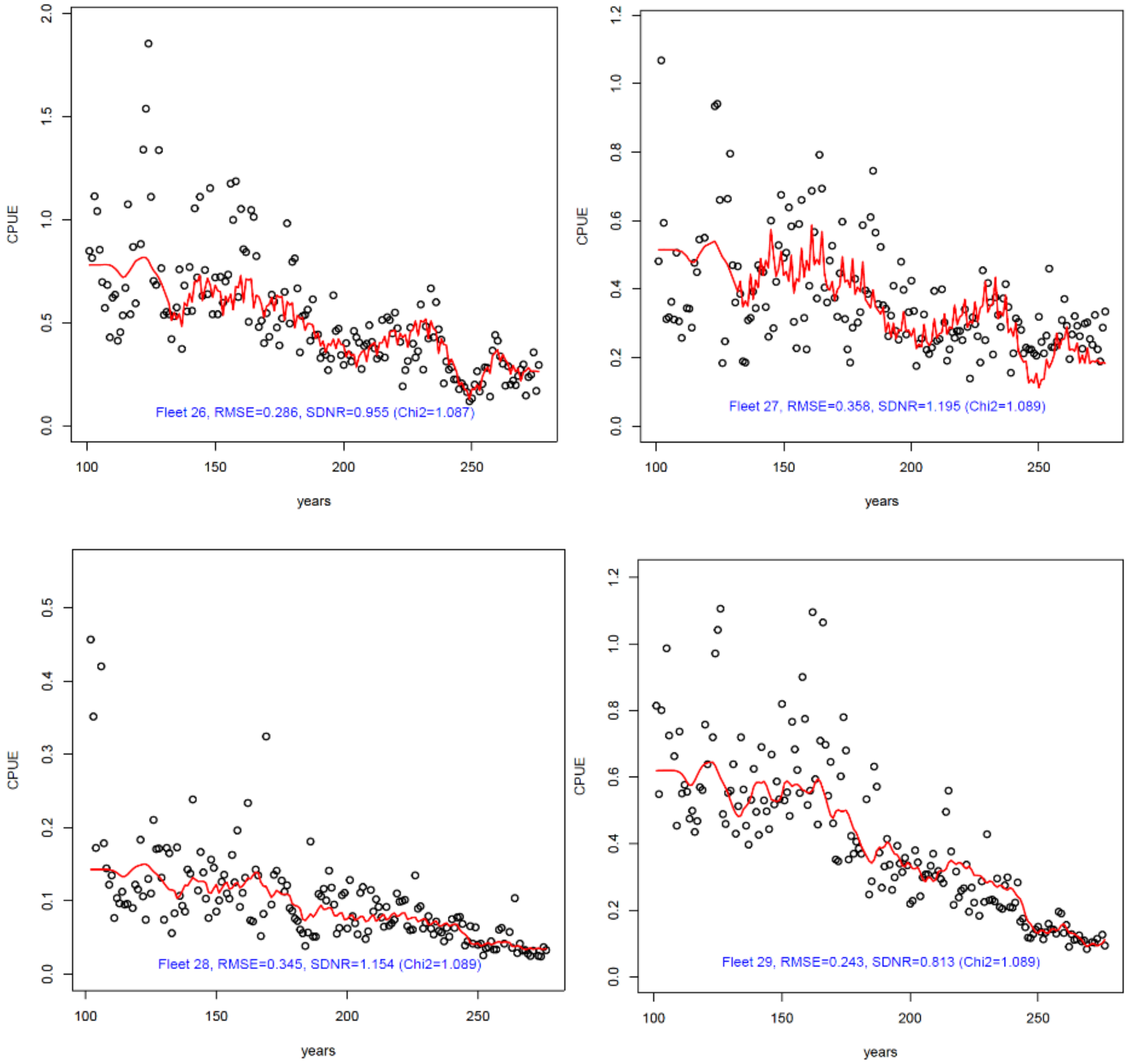


**Fig. 2.** Results of SDNR analysis including fit to each CPUE.

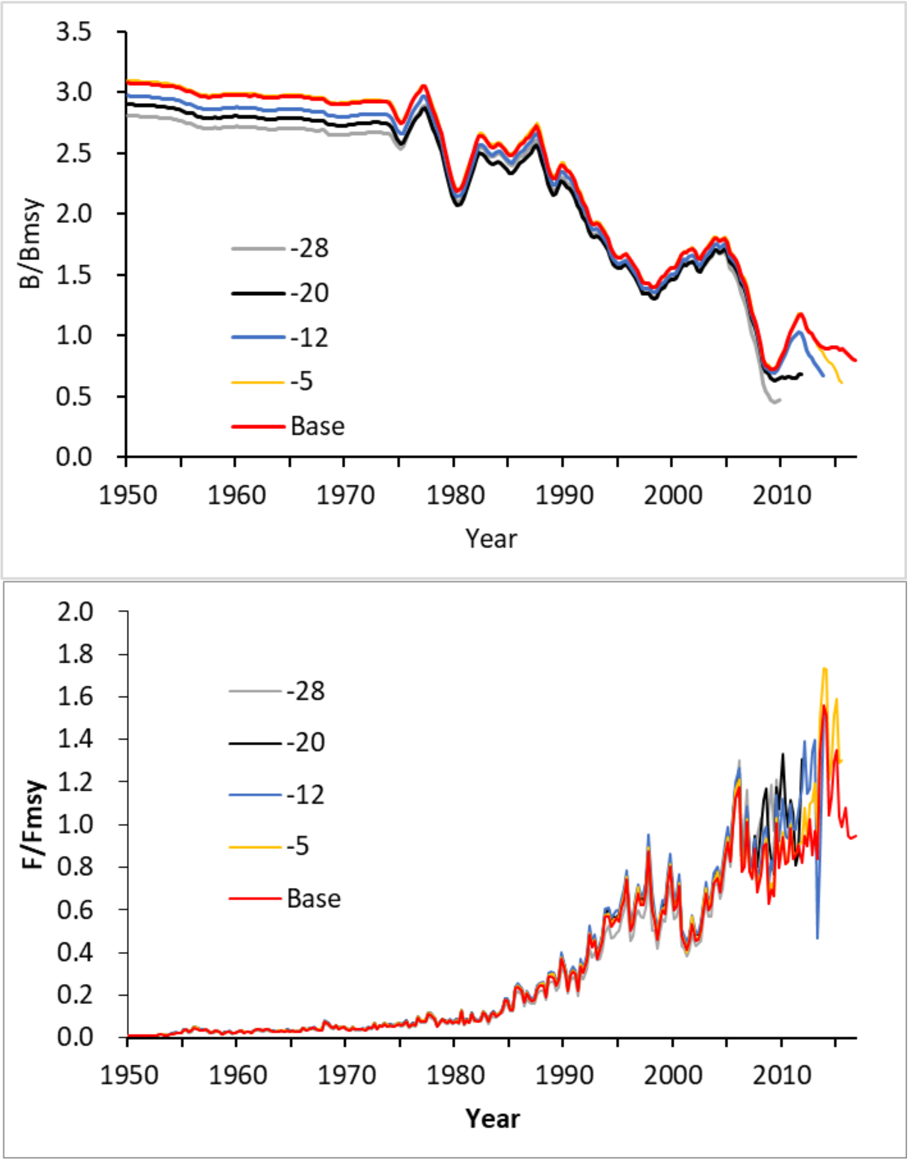


**Fig. 3.** Results of retrospective analysis (upper: B/Bmsy, lower: F/Fmsy). The legend shows number of time steps removed from terminal year.

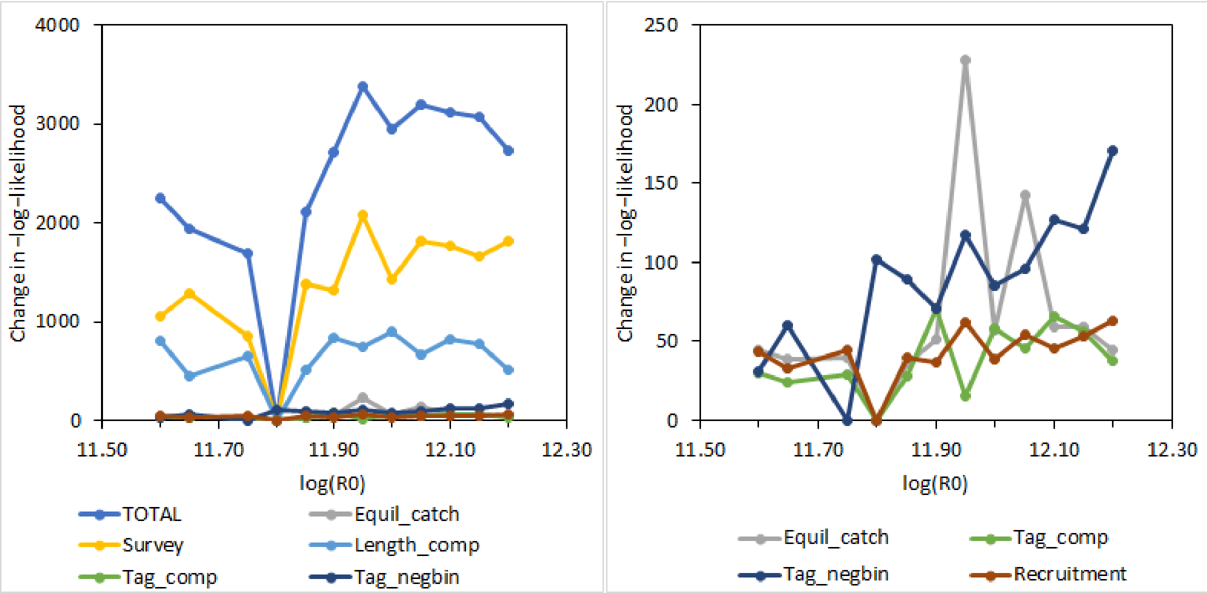


**Fig. 4.** Results of R0 likelihood profile. Left: all the components, right: the components whose change in log-likelihood is smaller.

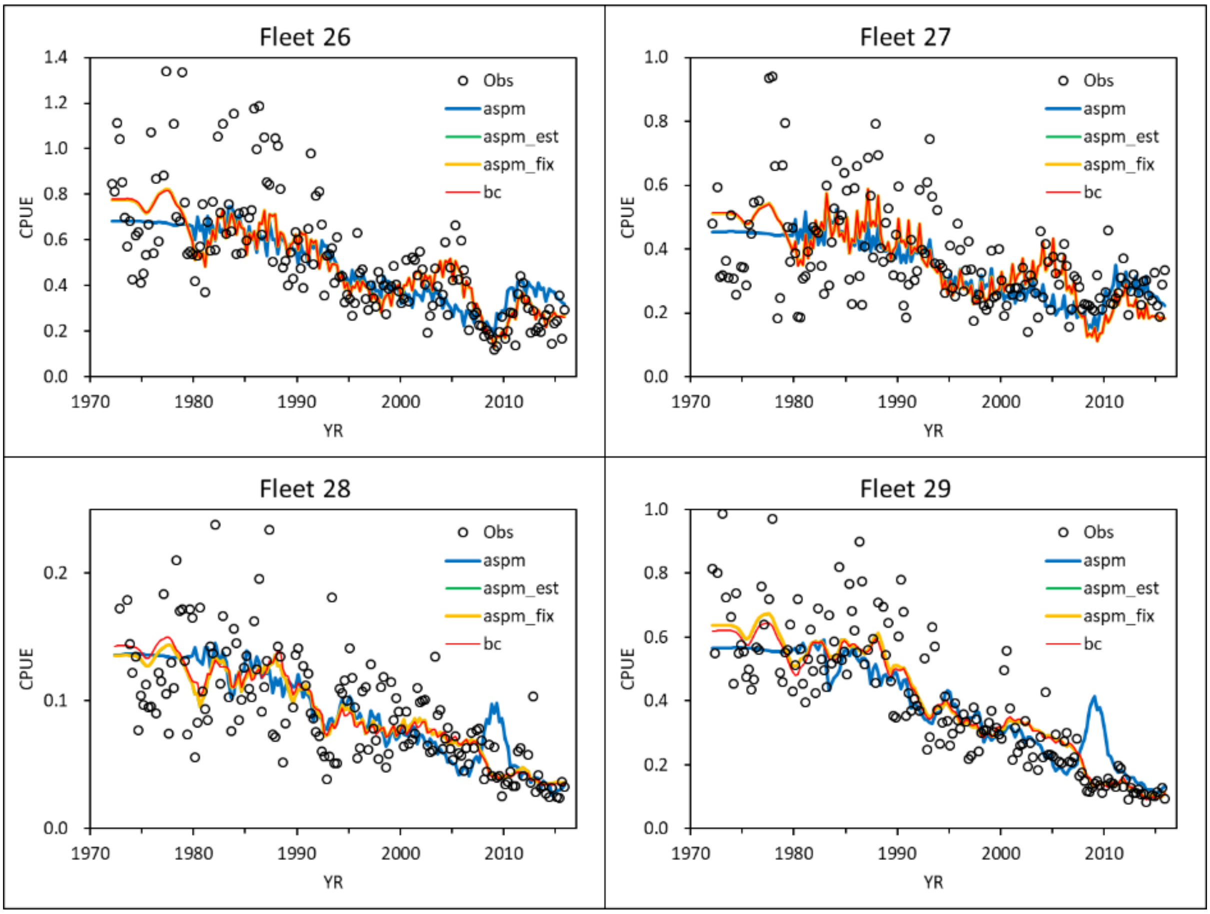


**Fig. 5**. Comparison of the trend of CPUE fit based on ASPM diagnosis.

bc: base case (integrated model), aspm: using the selectivity parameters from integrated model, the recruitment deviation is not used, aspm_est: using the selectivity parameters from integrated model, but estimated recruitment deviation, aspm_fix: using the selectivity parameters and recruitment deviation from integrated model.

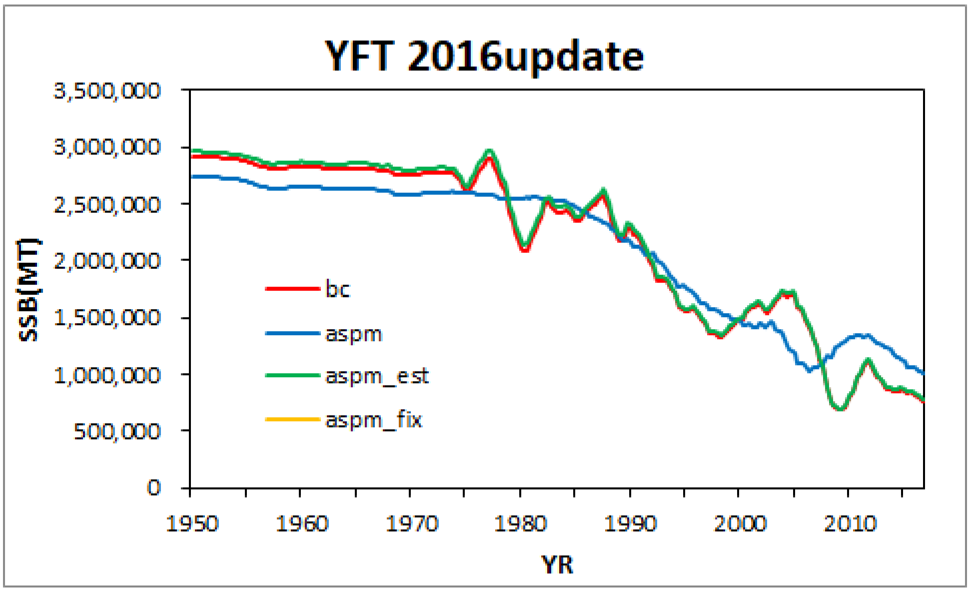


**Fig. 6**. Results of ASPM diagnosis (comparison of SSB trend). For the legend, see that in Fig. 5.