ISC/11-1/PBFWG/01

Exploration of a Hybrid Stock Synthesis-VPA Model of Pacific Bluefin Tuna to Distinguish Between Trends in Recruitment and Changes in Fishing Intensity on Young Fish

Alec D. MacCall[1] and Steven L. H. Teo[2]

[1] Southwest Fisheries Science Center, Fishery Ecology Division, Santa Cruz, CA, USA
[2] Southwest Fisheries Science Center, Fishery Resources Division, La Jolla, CA, USA

Jan. 2011

Working document submitted to the ISC Pacific bluefin tuna Working Group, International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), 6-13 January 2011, Shimizu-ku, Shizuoka, Japan. **Document not to be cited without author's permission.**

Abstract

We develop a new "hybrid" SS3-VPA assessment, where ages 0 and 1 are disconnected from the selectivity curves for older fish. Catches of ages 0 and 1 are removed from the landings, and are combined to form two "artificial" single-age fisheries with unit selectivity so that fishing mortality rates can be estimated directly. Applying this model to Pacific bluefin tuna indicates that recent (since 1990) fishing mortality rates may be higher than are estimated by the 2010 base model, and that current biomass is substantially lower by nearly half. Expected recruitment is constant (steepness h=1) in both models, and estimated recruitments from the two models are nearly identical.

Introduction

Recent stock assessments of Pacific bluefin tuna (PBF) indicate an increase in recruitment during the past decade. However, there is a risk that this pattern could be an artifact resulting from use of constant selectivity curves that fail to recognize a shift in fishery selectivity. It is possible that fleets have increased their targeting of young fish, especially ages 0 and 1. Although the Stock Synthesis model could be re-configured to reflect this alternative selectivity hypothesis, it could be difficult to judge which hypothesis is more likely.

Under a VPA approach, as was used in much earlier PBF assessments, there is no selectivity curve for the younger fish—F values on young fish are free to take any value that explains the time series of catches from the cohort. In VPA, only the older fish age classes are linked by a selectivity assumption. Thus we would expect that a VPA approach would be less likely to misinterpret a shift toward targeting of young fish as an increase in recruitment.

In this paper we explore a new approach to modeling in Stock Synthesis where a conventional SS3 model applies to fish aged 2 and older, and fish aged 0 and 1 are disconnected from the selectivity model so that the youngest ages conform to a VPA solution of estimated fishing mortality rates. This is accomplished by creating two "artificial" fisheries on ages 0 and 1 respectively. Catches of these two age groups are removed from the existing fisheries, and are re-assigned to the two artificial fisheries. By assigning unit selectivity to each age, SS3 calculates F directly.

We also introduce a new diagnostic tool for use in SS3 models—use of principal components analysis on portions of the correlation matrix. The resulting patterns summarize major multidimensional axes of variability or uncertainty in the assessment.

Materials and Methods

Hybrid Stock Synthesis – VPA Model

The hybrid model uses Stock Synthesis 3 (v3.11c) and is a modification of the PBF base model developed in 2010 (Ichinokawa et al. 2010). The modifications are partially based on discussions with R. Methot (Pers. Comm.), and result in an SS3 assessment with VPA-like properties for the youngest age groups.

Data File for Hybrid Model: The data used in this study were modified from the SS3 data file (Data08_20100611ver7.SS) used in the PBFWG update of bluefin tuna status in July 2010 (Ichinokawa et al. 2010). The main modifications are: 1) catches of age-0 and age-1 fish were removed from all fisheries except Japan longline and Taiwan longline fisheries, which catch negligible numbers of age-0 and age-1 fish, 2) two "artificial" fisheries (AGE0 and AGE1) were created to contain the age-0 and age-1 catches that were removed, and 3) length compositions in bins for age-0 and age-1 fish were set to zero for all fisheries except the two longline fisheries. The CPUE indices in the data file were not modified. Besides the original data file, all of the auxiliary data (e.g., catch-at-age, length-at-age) used in this study were extracted or derived from quantities in the original SS3 report file from Ichinokawa et al. (2010).

Age-0 and age-1 catches (mt) were estimated by multiplying the number of age-0 and age-1 fish caught by each fishery for each season and year, with the retained weight of age-0 and age-1 fish for their respective fishery and season. The number of age-0 and age-1 fish, and the retained weight were extracted from the 'CATCH_AT_AGE' and 'BIOLOGY_AT_AGE' components of the SS3 report file. The catches of age-0 and age-1 fish from all fisheries, except for Japan longline and Taiwan longline, were summed and placed into the "AGE0" and "AGE1" fisheries, respectively; and were removed from their respective original fisheries.

Length bin boundaries for age-0 and age-1 fish were estimated by cohort slicing for all fisheries, except Japan longline and Taiwan longline, and length frequencies in those bins were set to zero. The seasonal length bin boundaries, l_{ij}, between adjacent age classes i and j were estimated by,

$$l_{ij} = \mu_i + [\sigma_i / (\sigma_i + \sigma_j)] * (\mu_j - \mu_i)$$

where, μ and σ are the mean and standard deviation (SD) of each age class by season, and $j = i + 1$. Length frequencies in bins that were smaller than l_{12} were set to zero for each season. The means and SDs of seasonal age class lengths were extracted from the 'BIOLOGY_AT_AGE' component of the report file. Length composition data for the Japan pole and line fishery were eliminated due to having only two lines of data remaining after the modifications.

In addition, the initial conditions for this study are slightly different from the original assessment. Initial catches in the original data file were defined for the Japan longline, Tuna purse seine, and Japan troll fisheries. Since we were unable to determine how the initial catches in the original data file were derived, we estimated the initial catches of age-0 and age-1 fish from reported initial catch-at-age and retained weight. Similar to the procedure described above, we removed the estimated initial age-0 and age-1 catches from the Tuna purse seine and Japan troll fisheries and put them into the initial catches for AGE0 and AGE1 fisheries. The removed catch was larger than the original initial catch for the Japan troll fishery so we set the initial catch for that fishery to zero.

SS3 Control File for Hybrid Model: The SS3 control file used in this study was modified from the SS3 control file (Control08.SS) used in the PBFWG update of the status of bluefin tuna during the workshop in Nanaimo, Canada in July 2010 (Ichinokawa et al. 2010). The modifications are: 1) age selectivities were set to unit value for AGE0 and AGE1 fisheries respectively and length selectivities for AGE0 and AGE1 fisheries were turned off, 2) length selectivities for CPUE indices were not mirrored to their respective fisheries but were instead set to fixed parameters extracted from the original report file, and 3) the age selectivities for age-0 and age-1 were set to zero for all fisheries except Japan longline and Taiwan longline. All other parameters remained the same as in the original control file.

Principal Components Analysis

Principal components analysis (PCA) is a method of analyzing a correlation matrix to describe major axes of variability or uncertainty. The SS3 program produces a matrix of correlation coefficients (ss3.cor) which can be used directly as input to a PCA. The first line of the "cor" file is not part of the matrix and has to be deleted. Table 1 is a script written in the R language, and shows how the relevant parts of the "cor" file are extracted and input to the PCA. The output of the PCA is a list of "principal components" in descending order of the fraction of total variance they explain. The "loadings" of each principal component are coefficients that describe the general pattern of fluctuation associated with that component.

Results

Hybrid SS3-VPA model

Stock-Recruitment Relationship and Initial Conditions: The initial conditions are roughly equivalent for both models, but also are somewhat questionable in general. Initial (year 1952) spawning stock biomass (SSB) in the hybrid model is 22200 tons, compared with 25000 in the base model. The constant expected recruitment of 13213 in the hybrid model is slightly larger than 12145 in the base model, or about 9% more productive. Both models generate very large recruitments in 1947 (recdev >2), but early large recdevs are a frequent artifact reflecting the model's use of early recruitment values to adjust its initial conditions. (This hypothesis can be tested by beginning the model earlier or later. If it is an artifact, the year of the large recruitment event will shift accordingly.) Initial recruitments in the hybrid model are better balanced with subsequent expected recruitments: In the hybrid model expected recruitment drops 7% from 1951 to 1952, but in the base model they rise 61%. This difference is probably not important after the first decade of the model.

Spawning Stock Biomass: The histories of estimated SSB for the two models are shown in Figure 1. The time trajectories are similar (the hybrid model tends to be a little lower) until the early 1990s where they diverge. The hybrid model gives an ending 2007 SSB that is slightly over half (56%) of that estimated by the base model.

Recruitment: The history of estimated recruitments for the two models is shown in Figure 2. Values are similar for both models, with the hybrid model giving slightly larger recruitments in before the mid-1980s, but being nearly identical since then. Both models agree that recruitment has increased since the early 1990s. Clearly, the difference in recent SSB values is not due to a difference in estimated recruitment.

Fishing Mortality Rate: Histories of F are shown for young fish (ages 0-3) in Figure 3, and for combined intermediate (ages 4-6) and older fish (ages 7-10) in Figure 4. VPA-based values of F for age 0 fish increased relative to the base model since about 1980, and recently have been higher that the base model estimates by about 0.2. Though they are freely estimated by the VPA approach, F values for age 1 are nearly identical to the base model. Values of F for ages 2 and older are constrained by selectivity curves, but the hybrid model nonetheless gives higher F values for age 2, with a strong trend of increasing relative F since the late 1980s. *It is possible that this is an artifact of the cohort-slicing used to separate the age 0 and age 1 fish from the older fish in the length frequencies.* Still the difference in estimated F on age 2 fish is nearly 0.4 in recent years. Ages 3 and older all show increased F in recent years. The increased estimated F at ages 0 and 2 is enough to account for the recent decline in relative estimated biomass. The increase in F at older ages is consistent with the given catch and lower estimated biomass. In summary, the hybrid model is in agreement with the base model regarding stock productivity, but indicates a higher fishing mortality rate and lower spawning stock biomass than the base model for the past two decades.

Principal Components Analysis

The PBF assessment is unusual in that the estimated recruitments are highly independent. They show no coherent patterns of variability, and do not benefit from use of PCA (PC1 explained only 5% of the total variability). This may be the result of using a steepness h=1 so that uncertainty in the SRR does not appear in the assessment. This is also evidence that the base assessment does not contain an artifact of increasing recruitment due to possible recent shifts in fleet selectivity toward younger fish.

The first three principal components of spawning stock biomass from the base model and hybrid model are compared in Figure 5. The respective patterns of PC loadings are not strictly comparable between the two assessment models, but share enough similarity to be plotted together. The largest component in the base model (PC1) shows positive loadings over the entire range of years following the initial conditions of the model. This pattern is characteristic of general uncertainty in the overall scaling of the biomass estimates (all are higher or lower together), and accounts for a fairly large portion (41%) of the total uncertainty. The corresponding component from the hybrid model shows a similar pattern since the 1970s, but the early years vary inversely, which may reflect the somewhat different initial conditions used in the hybrid model. For the hybrid model, the largest component accounts for less of the overall uncertainty (34%). Because the recruitments do not share this pattern of loadings, the uncertainty in biomass scaling is not the result of recruitment uncertainty, but is more likely to be associated with uncertainty in estimated fishing mortality rates.

The second component (PC2) for the base model (18%) seems to correspond most closely to the third component (PC3) for the hybrid model (14%), and indicates an inverse relationship between pre-1980 and post-1980 biomass estimates. The combination of the first and second components accounts for about half of the total uncertainty, and can be interpreted as a tendency for biomass estimates of the first half of the time period to be partially independent from those of the second half of the time period (this could be shown by a factor analysis rotation of the axes, such as "varimax" rotation). As usual, subsequent components explain smaller portions of the overall variability, and show more complicated patterns that have no simple interpretation.

Conclusions

Our interpretation of these results is subject to a caveat: There is a possibility that the hybrid model exhibits an artifact due to the data treatment, including cohort slicing, or some detail of the SS3 code and resulting model. The hybrid model has not previously been attempted for any SS3 assessment, and merits close examination to assure that it is valid.

Although use of selectivity curves can provide statistical efficiency, and are necessarily used when using SS3 conventionally, selectivity curves can create a bias when the assumed patterns of fishing mortality rates are violated systematically. If the selectivity curves in the base model correctly reflect the pattern of fishing mortality rate, there should be no substantial difference between the base model and the hybrid model. Because the hybrid model produces higher estimates of fishing mortality rates, especially on age 0, it appears that the base model may contain such a bias. The consequence of this bias is a tendency for the base model to overestimate bluefin tuna biomass and to underestimate F for recent years.

References

Ichinokawa, M., M. Kai, and Y. Takeuchi. 2010. Stock assessment of Pacific bluefin tuna with updated fishery data until 2007. ISC/10-1/PBF/01

Table 1. R script for applying principal components analysis to SS3.cor file.

```
## identify portions of correlation matrix to be examined
## in pbf base model start1 to end1 is biomass, start2 to end2 is recruitment
start1<-89
end1<-146
start2<-147
end2<-203

## read correlation matrix from SS3 model
## important pre-processing:delete first line from SS3.cor
ss3cor<-read.table("ss3.cor", header=T, fill=T)

## extract values and std devs of models outputs--should be same as in ss3.std
SSB<-ss3cor[start1:end1,3]
REC<-ss3cor[start2:end2,3]
SSBstd<-ss3cor[start1:end1,4]
RECstd<-ss3cor[start2:end2,4]

## extract portions of correlation matrix
## if biomass or recruitment to be done alone use (delete comment marks)
rowkeep<-(start1:end1)
##or
##rowkeep<-(start2:end2)
##if biomass and recruitment are to be done together use
##rowkeep<-c(start1:end1,start2:end2)
colkeep<-rowkeep+4
CORMAT<-ss3cor[rowkeep,colkeep]

## replace any na values with -0
CORMAT[is.na(CORMAT)]<-0

## do principal components analysis
summary(pc.cr<-princomp(x=CORMAT, cor=T))
print(loadings(pc.cr), digits=5, cutoff=0)

## results can be plotted in R or in spreadsheet
```

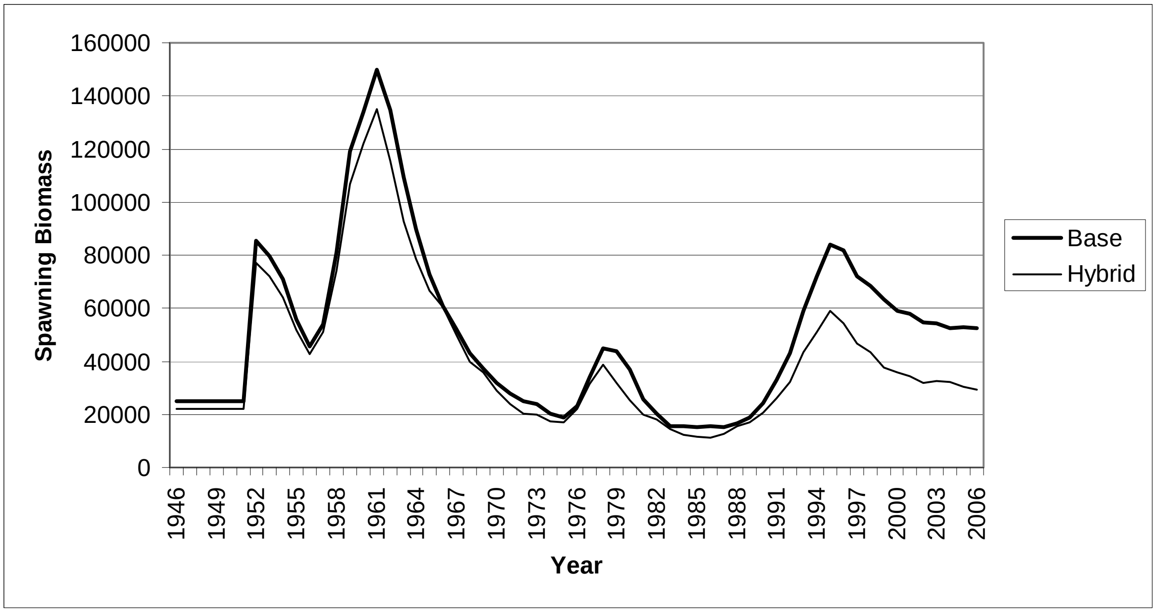

Figure 1. Comparison of estimated SSB.

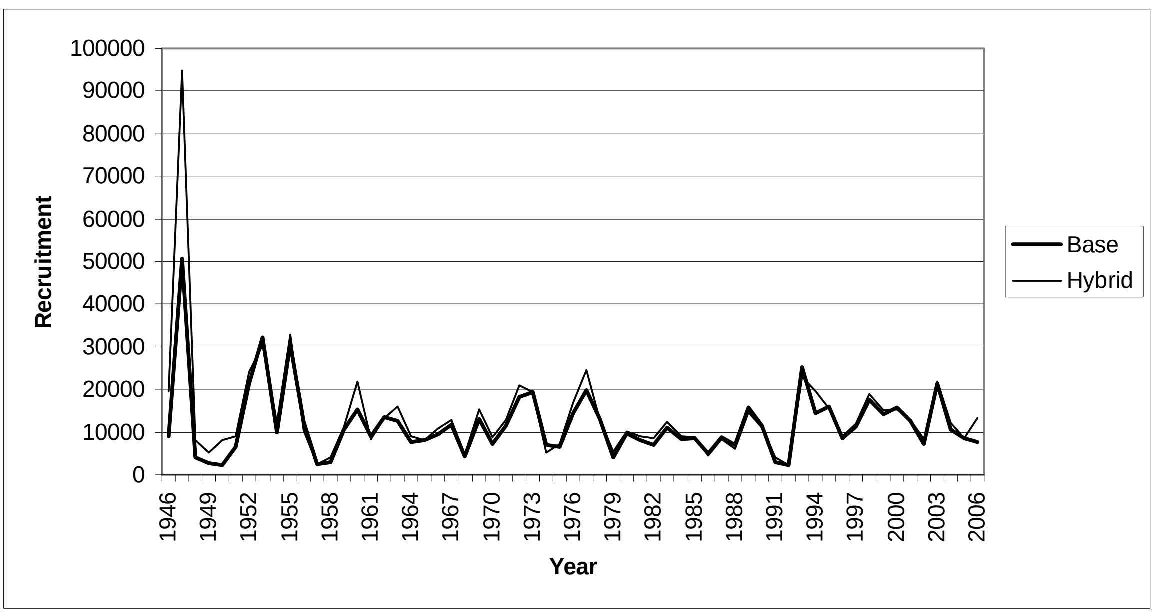

Figure 2. Comparison of estimated recruitments**.**

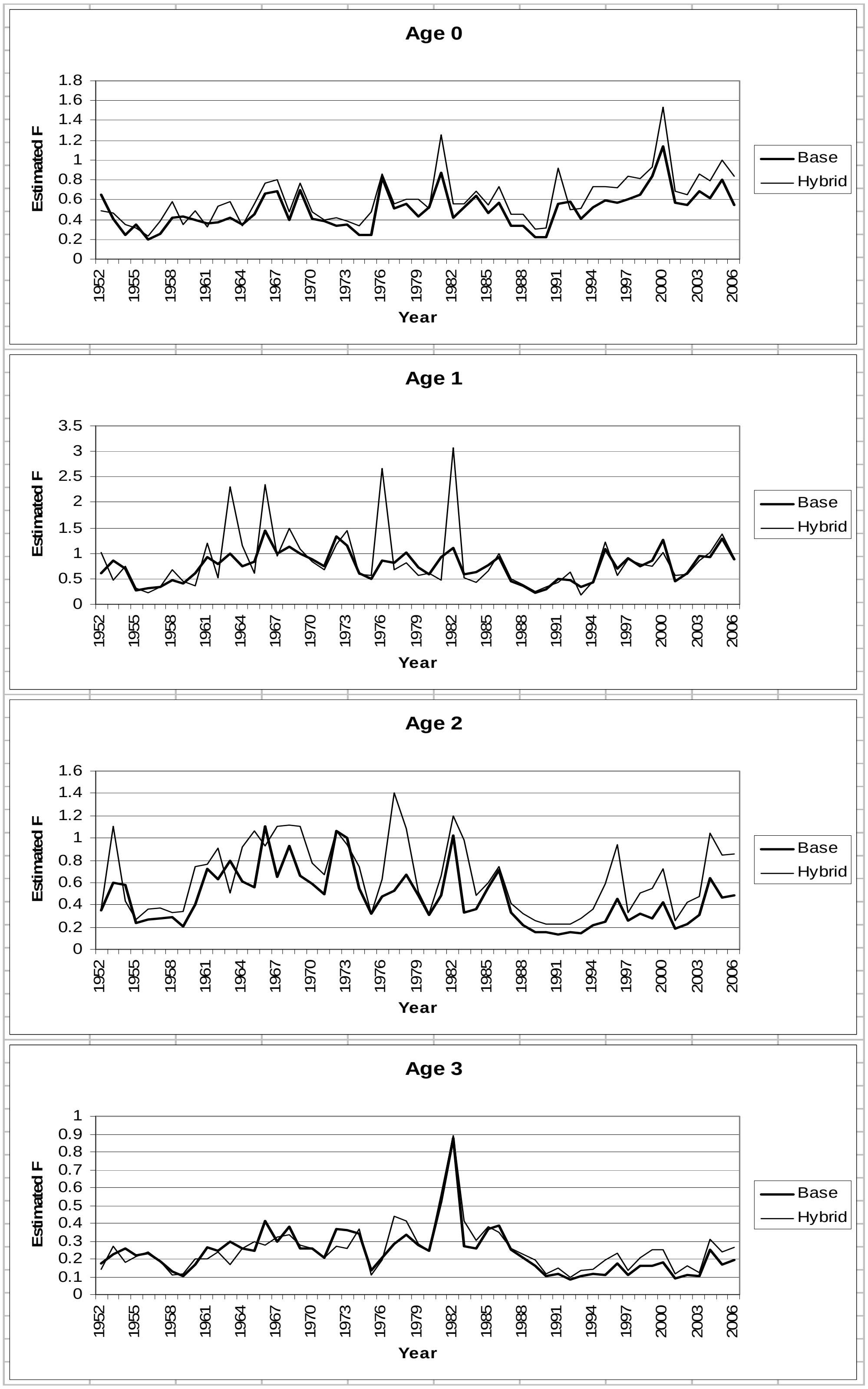

Figure 3. Comparison of F at age for younger ages.

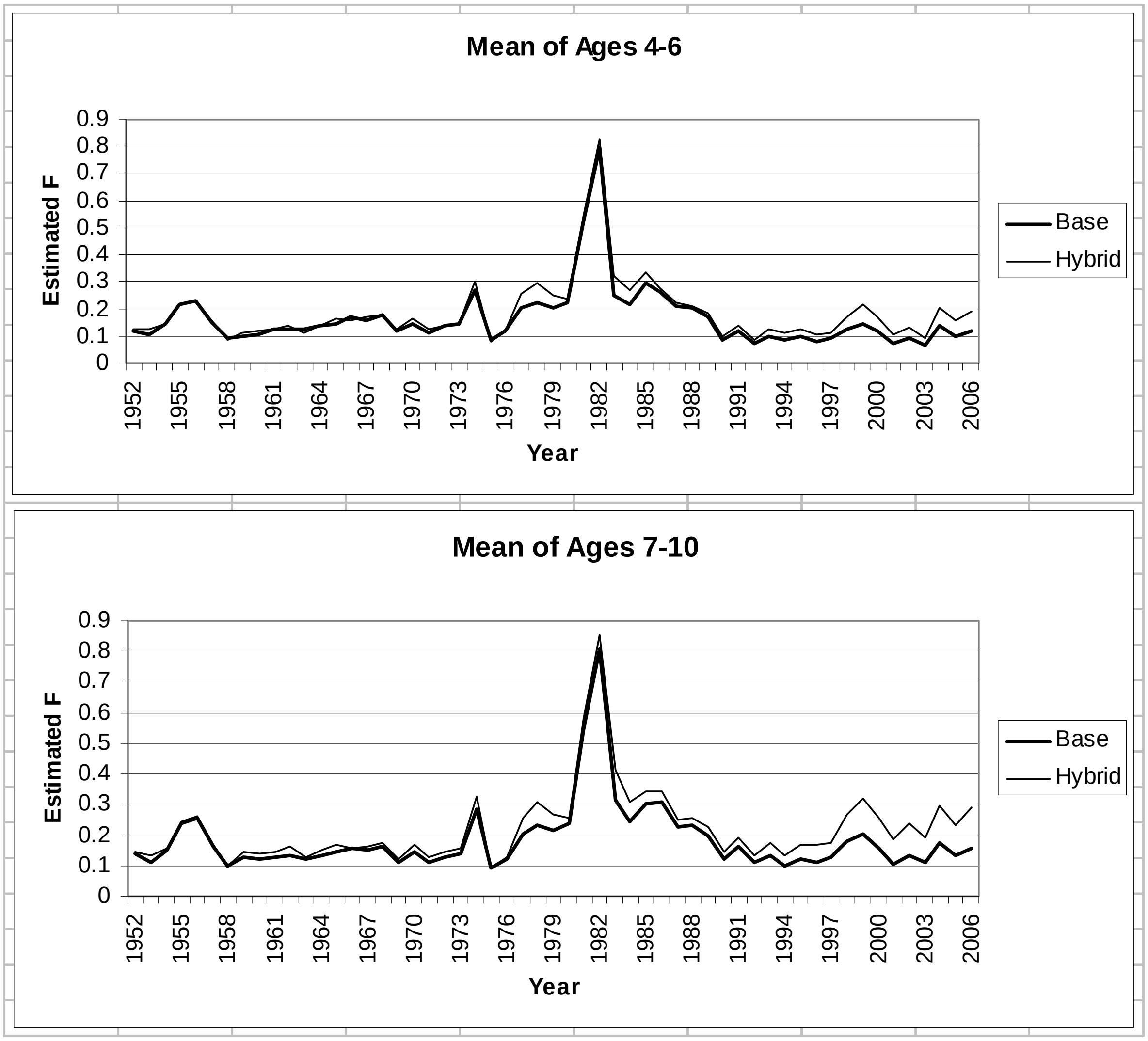

Figure 4. Comparison of mean F at age for older ages.

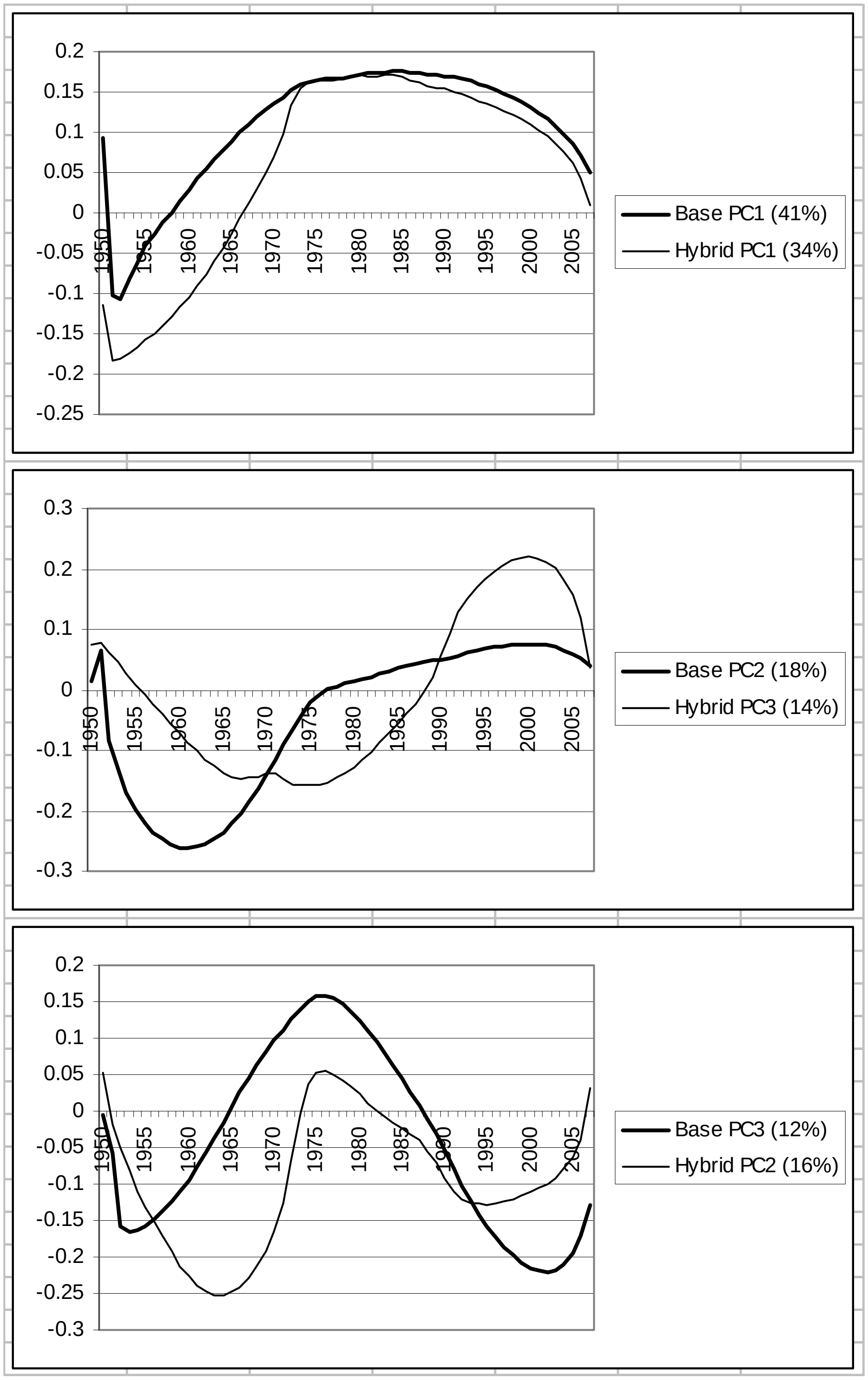

Figure 5. First three principal components of spawning stock biomass from base and hybrid models.