

**A method for analyzing the abundance and mortality of Atlantic and Gulf King Mackerel  
when the two stocks are presumed to intermix**

Clay E. Porch

Southeast Fisheries Science Center  
Miami Laboratory  
Sustainable Resource Division  
75 Virginia Beach Drive  
Miami, FL 33149-1099

November, 2003

Sustainable Fisheries Division Contribution: SFD-2003-0030

Past assessments of the king mackerel populations in U.S. waters have assumed that there are two distinct migratory groups, one living predominantly in the Atlantic and the other predominantly in the Gulf of Mexico. The abundance and mortality of each group has been estimated separately by use of a tuned VPA (Fadap, Restrepo 1996), which implies the belief that the two groups seldom intermix. Recently, concern has been expressed that the level of intermixing may not be negligible and that the current estimates of stock status could be biased.

One way to address the possible implications of intermixing is through the use of box-transfer models where the interchange between the two groups can be modeled explicitly. Porch et al. (2001) have implemented two types of box-transfer models in the form of two-stock VPAs. The software for this, known as VPA-2BOX, is available from the official web site of the International Commission for the Conservation of Atlantic Tuna (<http://www.iccat.es>) or directly from the author. This document provided a brief overview of program VPA-2BOX and how it might be applied to Gulf and Atlantic King Mackerel. For a more detailed accounting of the methodology the reader is referred to the User's guide (Porch, 2003)

## THE MODELS

The program VPA-2BOX includes two types of box-transfer models: Diffusion and overlap (see Tables 1 and 2 below). The former assumes a fraction  $T_{jk}$  of the population in zone  $j$  transfers to zone  $k$  whereas the latter assumes a fraction  $\tilde{T}_{sk}$  of stock  $s$  transfers to zone  $k$ . Essentially, the diffusion model ascribes membership to a stock based on the current location of an animal with the tacit assumption that immigrants adopt the movement and spawning behavior patterns of the endemic population. The overlap model, on the other hand, ascribes membership to a stock at birth, with the tacit assumption that the two stocks have overlapping ranges but otherwise seldom interact. Both models reduce to single-stock VPA's when the transfer fractions are set to zero, otherwise  $T$  and  $\tilde{T}$  are not directly comparable inasmuch as the diffusion model allows fish born in one zone to accumulate in the other. For example, if the transfer fractions for both stocks were both 10%, the diffusion model (without mortality) would calculate that 10% of the fish born in zone 1 will be living in zone 2 after one year, 18% after two years and 24.4% after three years.

The equations describing the catch and population dynamics include a large number of variables representing the transfer coefficients, natural mortality rates, fishing mortality rates, catches and population abundances. Of these, it is often only the catches that are actually observed. The VPA approach to this problem is to develop an explicit recursion that determines the historical abundance and fishing mortality rate of each cohort from the observed catches and prescribed values for  $M$ ,  $T$ , and the fishing mortality rate on the last age observed for each cohort ( $F_{Ay}$  or  $F_{aY}$ , where  $a$  is age,  $y$  is year,  $A$  is the last age, and  $Y$  is the last year).

Successful application of the VPA approach depends on three assumptions: (1) the catch equation correctly describes the dynamics of the fishery, (2) the observed catches are known with negligible error, and (3) reasonable values of  $M$ ,  $T$ ,  $F_{Ay}$  and  $F_{aY}$  can be derived. The first assumption is common to all modeling exercises and is discussed further in the section on Model Selection. The second assumption determines whether or not the VPA approach is suitable; if the catches are not well known, then a statistical catch-at-age model may be more appropriate.

The third assumption has often been addressed by fixing  $M$ ,  $T$ ,  $F_{Ay}$  and  $F_{aY}$  to several values that seem plausible and then comparing the resulting solutions. The approach taken by ADAPT and related

methods differs from this in that at least some of these parameters are estimated internally by use of auxiliary data such as indices of abundance or tag recoveries (a process sometimes referred to as ‘tuning’ the VPA). The tuning procedure occurs in two phases. The first phase is the explicit recursion, which determines the historical abundance and mortality rates from the latest estimates for  $M$ ,  $T$ ,  $F_{Ay}$  and  $F_{aY}$ . The second phase uses the abundance and mortality rate estimates from the first step to predict the values of the data. New values of  $M$ ,  $T$ ,  $F_{Ay}$  and  $F_{aY}$  are then selected so as to minimize the discrepancy between the model predictions and the observed values of the data as measured by an appropriate statistical model. These values are in turn used to reinitiate phase 1 and the process continues in iterative fashion until certain stopping criteria are met.

### Estimation from abundance indices

Program VPA-2BOX allows parameters to be estimated by minimizing the negative log-likelihood function of the observed and predicted values for various indices of abundance (Table 3). This introduces several new variables that need to be accounted for-- the index standard error  $\mathbf{s}$ , catchability  $q$ , and relative vulnerability to the gear  $v$  (which implicitly includes factors such as gear selectivity and the fraction of the population available to be caught). There are many options available in VPA-2BOX pertaining to how these variables may be treated and the reader is referred to the user’s manual for a detailed accounting.

### Estimation from tag recoveries

Program VPA-2box assumes the temporal and spatial distribution of tag recoveries from a given group (cohort) of releases are multinomial-distributed (Schweigert and Schwarz, 1993; Anganuzzi et al., 1994) as described in Table 4. This tag attrition model includes terms for the reporting rate ( $\mathbf{r}$ ), immediate loss of tags owing to misapplication or tagging-induced mortality ( $d$ ), and chronic loss of tags owing to shedding or fouling ( $\mathbf{I}$ ). The first two terms always occur together as the product  $(1-d)\mathbf{r}$  and cannot be distinguished using tag recovery data alone; therefore only one of the terms may be estimated and the other must be fixed. The chronic tag loss rate  $\mathbf{I}$  is similarly confounded with the natural mortality rate  $M$ , but independent estimation becomes possible when abundance indices are available in addition to the tagging data.

The model in Table 4 also allows for the possibility that the effective fishing mortality on the tagged population may differ from that on the untagged population by use of pre-mixing adjustment factors  $\mathbf{g}_t$ . The values of  $\mathbf{g}_t$  can be difficult to distinguish precisely from  $F$  when only tag recovery data are available (Hoenig et al., 1998), but the task is made easier when catch and abundance data are included because  $F$  is largely determined from them.

Another necessary adjustment has to do with the fact that the VPA accounting is by year, but tag releases tend to occur at various times within a year-- sometimes before and sometimes after the periods of most intense fishing. Consider, for example, a group of fish that were tagged and released at the end of the 7<sup>th</sup> month after the major fishery was over. In that case the tagged fish would be exposed to five months of tag loss and natural mortality, but none would be recaptured. If no adjustments are made to account for this, the model would erroneously interpret the lack of recaptures for that year as no fishing for the year, when in fact the fishing pressure could have been very high. The fishing fraction parameter  $\mathbf{f}$  in Table 4 adjusts the fishing mortality rate parameter

by the fraction of the total fishing pressure exerted during the year that was experienced by tags. This requires information on the seasonality of the fisheries beyond that which is normally required by VPA (see discussion on line 34 of the control file). Of course this sort of problem only occurs for the year when the fish were first tagged. In subsequent years the tagged population will face the full cycle of fishing such that  $f=1$ .

Finally, it usually will not be possible to distinguish members of the two stocks during the tagging process. In the case of the diffusion model this is unimportant because the two stocks are defined by management zone (rather than by origin)--fish born in zone 1 and tagged in the zone 2 are considered to be members of the zone 2 stock and vice versa. In the case of the overlap model, where the two stocks are defined by origin, fish from each stock are assumed to be tagged in proportion to their relative abundance in each management zone, i.e.,

$$R_{st} = R_t \frac{T_{skay} N_{say}}{\sum_s T_{skay} N_{say}} \quad (2.17)$$

where the  $N$  and  $T$  values are the estimates from the VPA.

The application of the two box models to King Mackerel is straightforward. The past approach of conducting two independent VPAs (one for each migratory group) is identical to conducting a two-stock VPA with zero mixing. The fact that the geographic boundary delineating the two stocks shifts with time poses no special problems for the two-stock VPA. Both the overlap or diffusion models will essentially assume that the transfer rates across the moving boundary are temporally invariant (although the model does allow the transfer coefficients to vary with age). They also assume that the animals that transfer across the boundary have the same probability of being caught as the endemic population. This condition may be violated, for example, when mixing occurs within a limited area and the fishing pressure in that area is different from the overall average (in which case a three-area, two-stock model might be more appropriate).

Past experience with box-transfer models has shown that indices of abundance are generally insufficient for estimating the transfer coefficients. Thus, tagging data are required (the model as currently configured does not use information from biological tags such as microconstituents or otolith shape). In the case of king mackerel, the type of tag used has changed during the course of the time series and it is likely that tag-shedding rate and tag-induced mortality have changed as well. Inasmuch as the model does not allow for temporal changes in these factors, it is suggested that the analyses be restricted to the tag type with the most recoveries. For the same reason, it is not necessary to have recoveries for every year in the time series; generally 5 or more consecutive years will do.

It is important to realize that the use of tagging data affects the estimation of the mortality rates as well as the transfer rates. Accordingly, one must be careful to distinguish the effect of using the tagging data from the effect of incorporating transfer coefficients. For this reason it is often useful to run the two-stock VPA with the tagging data assuming zero transfer and then without the tagging data assuming values for the transfer coefficients similar to the values estimated with the tagging data. If it is found

that the estimates obtained with the tagging data are very different from those obtained with the indices of abundance alone, then consideration will need to be given to downweighting one or the other in accordance with the degree to which they are likely to reflect the dynamics of the actual stocks.

- Anganuzzi, A., Hilborn, R. and Skalski, J. R. 1994. Estimation of size selectivity and movement rates from mark-recovery data. *Can. J. Fish. Aquat. Sci.* 51: 734-742
- Porch, C. E. 2003. VPA-2BOX Version 3.01 User's guide. Sustainable Fisheries Division Contribution SFD-01/02-151. NOAA Fisheries, Southeast Fisheries Science Center, 75 Virginia Beach Dr. Miami, Florida 33149 USA
- Porch, C. E., S. C. Turner and J. E. Powers. 2001. Virtual population analyses of Atlantic bluefin tuna with alternative models of trans-Atlantic migration: 1970-1997. *ICCAT Coll. Vol. Sci. Pap.* 52:1022-1045.
- Restrepo, V.R. 1996. FADAPT 3.0 A Guide. University of Miami, Cooperative Unit for Fisheries Research and Education (CUFER).
- Schweigert, J. F., and Schwarz, C. J. 1993. Estimating migration rates for Pacific herring (*Clupea pallasii*) using tag-recovery data. *Can. J. fish. Aquat. Sci.* 50: 1530-1540.

Table 1. Standard catch equation assumed in program VPA-2BOX

Catch equation	Description
$C_{kay} = \tilde{N}_{kay} \frac{F_{kay}(1 - e^{-Z_{kay}})}{Z_{kay}}$	Catch of age $a$ in year $y$ from all stocks in management zone $k$
$Z_{kay} = F_{kay} + M_{kay}$	Total mortality rate in zone $k$
$F_{kay}$	Fishing mortality rate in zone $k$
$M_{kay}$	Natural mortality rate in zone $k$
$\tilde{N}_{kay}$	Number of fish in zone $k$ that are age $a$ at the beginning of year $y$ (all stocks combined)

Table 2. Overlap and diffusion equations for population dynamics.

Overlap model equations	Description
$N_{s,a+1,y+1} = N_{say} \sum_k T_{skay} e^{-Z_{kay}}$	Number of fish from stock $s$ that are age $a+1$ at the beginning of year $y$ ( $a+1 < A$ )
$N_{s,A,y+1} = \sum_{a=A-1}^A N_{say} \sum_k T_{skay} e^{-Z_{kay}}$	Number of fish from stock $s$ that are age $A$ or older at the beginning of year $y$
$\tilde{N}_{kay} = \sum_s T_{skay} N_{say}$	Number of fish in zone $k$ that are age $a$ at the beginning of year $y$ (all stocks combined)
$T_{skay}$	Fraction of stock $s$ residing in zone $k$ at the beginning of year $y$
Diffusion model equations	Description
$\tilde{N}_{k,a+1,y+1} = \sum_j \tilde{N}_{jay} \tilde{T}_{jkay} e^{-Z_{kay}}$	Number of fish in zone $k$ that are age $a+1$ at the beginning of year $y$ ( $a+1 < A$ )
$\tilde{N}_{k,A,y+1} = \sum_{a=A-1}^A \sum_j \tilde{N}_{jay} \tilde{T}_{jkay} e^{-Z_{kay}}$	Number of fish in zone $k$ that are age $A$ or older at the beginning of year $y$
$\tilde{T}_{jkay}$	Fraction of population in zone $j$ that moves to zone $k$ at the beginning of year $y$

**Table 3.** Models for indices of abundance (index series:  $i$ , zone:  $k$ , age:  $a$ , year:  $y$ ) that are available in VPA-2BOX.

Equations and variables		Description
$\sum_i \sum_k \sum_y 0.5 \left( \frac{\ln(I_{iky}/\hat{I}_{iky})}{\tilde{\mathbf{s}}_{iky}} \right)^2 + \ln \tilde{\mathbf{s}}_{iky}$	lognormal	negative log-likelihood terms $L(\tilde{I})$ for various error distributions.
$\sum_i \sum_k \sum_y 0.5 \left( \frac{I_{iky} - \hat{I}_{iky}}{\mathbf{s}_{iky}} \right)^2 + \ln \mathbf{s}_{iky}$	normal	Note: for the lognormal distribution $\mathbf{s}'$ is the variance on a logarithmic scale,
$\sum_i \sum_k \sum_y \hat{I}_{iky} - I_{iky} \ln(\hat{I}_{iky})$	Poisson	
$\sum_i \sum_k \sum_y 0.5 \frac{(I_{iky} - \hat{I}_{iky})^2}{\mathbf{s}_{iky}^2 (\hat{I}_{iky} + 1)}$	Chi - square	$\tilde{\mathbf{s}}_{iky} = \sqrt{\log e \left\{ \left( \frac{\mathbf{s}_{iky}}{\hat{I}_{iky}} \right)^2 + 1 \right\}}$ ,
$\sum_i \sum_k \sum_y \left( \frac{\sqrt{2}  I_{iky} - \hat{I}_{iky} }{\mathbf{s}_{iky}} \right)^2 + \ln \mathbf{s}_{iky}$	Laplace	and for the gamma distribution: $\mathbf{a} = (\hat{I}_{iky}/\mathbf{s}_{iky})^2$ and $\mathbf{b} = \hat{I}_{iky}/\mathbf{a}$
$\sum_i \sum_k \sum_y C \ln \mathbf{b} - (\mathbf{a} - 1) \ln I_{iky} - \frac{I_{iky}}{\mathbf{b}} - \ln \Gamma(\mathbf{a})$	gamma	
$I_{iky}$		observed value of index
$\hat{I}_{iky} = q_{iky} \Delta_{ik} \sum_a v_{ikay} w_{ikay} \tilde{N}_{kay}$		predicted value of index
$v_{ikay}$		relative selectivity/availability at age
$\mathbf{s}_{iky}$		standard error of index
$q_{iky}$		catchability coefficient
$w_{ikay}$		weight
$\Delta_{ik}$		adjustment for time of year

**Table 4.** Model for tag recoveries. The subscript  $\mathbf{t}$  denotes a unique group (cohort) of tag releases distinguished by the zone ( $\mathbf{k}$ ), year ( $\mathbf{y}$ ) and age ( $\mathbf{a}$ ) of release. The remaining subscripts are as defined previously (stock:  $s$ , zone:  $k$ , year:  $y$ , age:  $a = \mathbf{a} + y - \mathbf{y}$ ).

Equations and variables	Description
$\mathcal{L}(\bar{r}) = \sum_{\mathbf{t}} \frac{1}{w_{\mathbf{t}}^2} \left\{ (R_{\mathbf{t}} - \sum_{k,y} r_{k\mathbf{t}y}) \ln \left[ 1 - \sum_{k,y} \frac{\hat{r}_{k\mathbf{t}y}}{R_{\mathbf{t}}} \right] - \sum_{k,y} r_{k\mathbf{t}y} \ln \left[ \frac{\hat{r}_{k\mathbf{t}y}}{R_{\mathbf{t}}} \right] \right\}$	negative log-likelihood for tag recoveries from all cohorts
$R_{\mathbf{t}}$	number of tag releases in cohort $\mathbf{t}$
$r_{k\mathbf{t}y}$	observed recoveries from cohort $\mathbf{t}$
$\hat{r}_{k\mathbf{t}y} = \mathbf{r}_{kay} \tilde{n}_{k\mathbf{t}y} \frac{f_{k\mathbf{t}ay}}{z_{k\mathbf{t}ay}} (1 - e^{-z_{k\mathbf{t}ay}})$	expected recoveries from cohort $\mathbf{t}$
$z_{k\mathbf{t}ay} = f_{k\mathbf{t}ay} + (M_{kay} + \mathbf{l}_{\mathbf{t}y})(1 - t_{\mathbf{t}y})$	total loss rate of cohort $\mathbf{t}$
$f_{k\mathbf{t}ay} = \mathbf{f}_{\mathbf{t}y} \mathbf{g}_{\mathbf{t}y} F_{kay}$	fishing mortality rate of cohort $\mathbf{t}$
$\mathbf{g}_{\mathbf{t}y}$	pre-mixing adjustment for cohort $\mathbf{t}$
$\mathbf{f}_{\mathbf{t}y}$	fishing fraction (= 1 for $y > \mathbf{y}$ )
$\mathbf{l}_{\mathbf{t}y}$	chronic tag loss rate of cohort $\mathbf{t}$
$d_{\mathbf{t}}$	immediate tag loss of cohort $\mathbf{t}$
$t_{\mathbf{t}y}$	release date of cohort $\mathbf{t}$ (= 0 for $y > \mathbf{y}$ )
$\mathbf{r}_{kay}$	reported fraction of recaptured tags
<i>Overlap tag attrition model</i>	
$\tilde{n}_{k\mathbf{t}y} = \begin{cases} R_{\mathbf{t}}(1-d_{\mathbf{t}}) & (y = \mathbf{y}, k = \mathbf{k}) \\ \sum_s T_{skay} n_{s\mathbf{t}y} & (y > \mathbf{y}) \end{cases}$	number of survivors with tags from cohort $\mathbf{t}$ in zone $k$ at start of year $y$
$n_{s, \mathbf{t}, y+1} = \begin{cases} R_{s\mathbf{t}}(1-d_{\mathbf{t}}) e^{-z_{k\mathbf{t}ay}(1-t_{\mathbf{t}y})} & (y = \mathbf{y}) \\ n_{s\mathbf{t}y} \sum_k T_{skay} e^{-z_{k\mathbf{t}ay}} & (y > \mathbf{y}) \end{cases}$	number of survivors with tags from cohort $\mathbf{t}$ and stock $s$ at start of year $y$ ( $R_{s\mathbf{t}}$ is the number of releases by stock)
<i>Diffusion tag attrition model</i>	
$\tilde{n}_{k, \mathbf{t}, y+1} = \begin{cases} R_{\mathbf{t}}(1-d_{\mathbf{t}}) e^{-z_{k\mathbf{t}ay}(1-t_{\mathbf{t}y})} & (y = \mathbf{y}) \\ e^{-z_{k\mathbf{t}ay}} \sum_j T_{jkay} \tilde{n}_{j\mathbf{t}y} & (y > \mathbf{y}) \end{cases}$	number of survivors with tags from cohort $\mathbf{t}$ in zone $k$ at start of year $y$