Testing time-varying selectivity models of the South Coast rock lobster with data generated by a model with spatial structure

L.B. Furman and D.S. Butterworth

MARAM
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch, 7701

January 17, 2011

Abstract

Both gear selectivity and availability combine to form selectivity. When selectivity is allowed to change with time, it is difficult to determine the source of the variation. If availability changes with time, perhaps as a consequence of changing spatial distributions of fishing effort, selectivity-at-age will appear to vary over time. It then becomes unclear how to renormlise selectivity so that the catchability q which relates CPUE to underlying abundance does not change. This study uses a model with spatial structure in the age distribution of the population together with changes in the spatial distribution of fishing over time to simulate variations in availability. The model is used to generate data for simulation testing of two different approaches to modeling time-varying selectivity.

1 Introduction

In resource assessments selectivity is used to model the vulnerability of fish of different ages to gear, as well as their availability. Age-specific availability reflects a combination of different age distributions of fish in different area and the spatial distribution of fishing effort. Availability often changes with time, and time-varying selectivity models can be used to try to take this into account. However, since catchability q and selectivity-at-age S_a are confounded, conventionally with time-independent S_a the largest of the S_a's is set to 1 to remove ambiguity. Thus interpretation problems arise if S_a changes over time, as it is not immediately clear how S_a should then be renormalised so that the catchability q in the

CPUE-abundance relationship can be assumed to remain unchanged over time.

The objective of this study is to assess the performance of two approaches to modeling time-varying selectivity, with test data that reflects underlying changes in availability, but constant gear selectivity. To this end, the data are generated by a two-area model with constant selectivity. As a result of migration from the one area to the other, the age-distribution differs between the areas. If the proportions of the catch taken in these areas varies from year to year, this appears as time-varying availability from the perspective of a single area assessment approach which assumes spatial homogeneity. This two-area model then produces the "reality" (operating model) against which to test time-varying selectivity models.

The two-area model is first fitted to observed catch-per-unit-effort (CPUE) and catch-at-age data for the South Coast rock lobster to tune its parameters to roughly resemble an actual situation. It would not be enough to simply test the time-varying selectivity methods against the single best CPUE and catch-at-age estimates of the two-area model. They must take account of the random variations in observed data that one realistically expects. Therefore, simulation tests are performed where the models are fitted to many sets of generated data which incorporate such random noise.

2 Data

The total annual catch by mass (MCM document WG/06/04/SCRL1) and GLM standardised CPUE time series (Glazer, 2007) data for the South Coast rock lobster used are given in Table 1. The commercial catch-at-age data (Bergh pers. commn) as derived from length data are shown in Table 2. All data are given as reported by Johnston and Butterworth (2007).

3 The two-area model

The following subsections describe features that are unique to the two-area model. Elements that are common to both the two-area model and the single-area time-varying selectivity models are detailed in the appendix.

3.1 The population dynamics

It is assumed that the resource is split between two areas, a *recruitment area* and an *immigration area*. At the start of each year new recruits are produced in the recruitment area, and throughout the year lobsters migrate from the recruitment area to the immigration area. This is intended to produce two areas with very different age structures, to simulate the effects of variations in availability-at-age as the spatial distribution of fishing changes from year to year.

The equations governing the population dynamics are:

$$N^r_{y+1,0} = R_{y+1} \tag{1}$$

$$N^r_{y+1,a+1} = N^r_{y,a} e^{-(M+\mu)} \left(1 - S_a F^r_y\right) \qquad a = 0, \ldots, m-2 \tag{2}$$

$$N^r_{y+1,m} = N^r_{y,m-1} e^{-(M+\mu)} \left(1 - S_{m-1} F^r_y\right) + N^r_{y,m} e^{-(M+\mu)} \left(1 - S_m F^r_y\right) \tag{3}$$

$$N^i_{y+1,0} = 0 \tag{4}$$

$$N^i_{y+1,a+1} = \left[N^i_{y,a} + N^r_{y,a}\left(1 - e^{-\mu}\right)\right] e^{-M} \left(1 - S_a F^i_y\right) \qquad a = 0, \ldots, m-2 \tag{5}$$

$$\begin{aligned} N^i_{y+1,m} = &\left[N^i_{y,m-1} + N^r_{y,m-1}\left(1 - e^{-\mu}\right)\right] e^{-M} \left(1 - S_{m-1} F^i_y\right) \\ &+ \left[N^i_{y,m} + N^r_{y,m}\left(1 - e^{-\mu}\right)\right] \left(1 - S_m F^i_y\right) \end{aligned} \tag{6}$$

$$N_{y,a} = N^r_{y,a} + N^i_{y,a} \tag{7}$$

where

$\boldsymbol{N^r_{y,a}}$ is the number of lobsters of age a in the recruitment area at the start of year y,
$\boldsymbol{N^i_{y,a}}$ is the number of lobsters of age a in the immigration area at the start of year y,
$\boldsymbol{N_{y,a}}$ is the total number of lobsters of age a at the start of year y,
$\boldsymbol{R_y}$ is the number of recruits at the start of year y,
$\boldsymbol{M}$ is the natural mortality rate of lobsters,
$\boldsymbol{\mu}$ is an estimable parameter describing the rate of migration from the recruitment area to the immigration area,
$\boldsymbol{S_a}$ is the age-specific time-invariant selectivity as described by equation (20),
$\boldsymbol{F^r_y}$ is the fully selected fishing mortality in the recruitment area in year y,
$\boldsymbol{F^i_y}$ is the fully selected fishing mortality in the immigration area in year y, and
$\boldsymbol{m}$ is the minimum age of the plus-group.

The proportion of the total catch caught in each area varies annually. Equations (8) and (9) describe the annual catch by mass in the recruitment area and the immigration area respectively:

$$p_y C_y = \sum_{a=0}^{m} w_{a+\frac{1}{2}} N^r_{y,a} e^{-(M+\mu)/2} S_a F^r_y \tag{8}$$

$$(1 - p_y) C_y = \sum_{a=0}^{m} w_{a+\frac{1}{2}} e^{-M/2} \left[N^i_{y,a} + N^r_{y,a}\left(1 - e^{-\mu/2}\right)\right] S_a F^i_y \tag{9}$$

where

$\boldsymbol{C^y}$ is the area combined total catch by mass for year y,
$\boldsymbol{p_y}$ is an additional estimable parameter giving the proportion of the total catch that is caught in the recruitment area in year y, and
$\boldsymbol{w_{a+\frac{1}{2}}}$ is the mid-year mass of lobsters at age a.

The model estimate of mid-year exploitable biomass for year y is given by:

$$\hat{B}^r_y = \sum_{a=0}^{m} w_{a+\frac{1}{2}} S_a N^r_{y,a} e^{-(M+\mu)/2} \left(1 - F^r_y/2\right) \tag{10}$$

$$\hat{B}^i_y = \sum_{a=0}^{m} w_{a+\frac{1}{2}} S_a \left[N^i_{y,a} + N^r_{y,a}\left(1 - e^{-\mu/2}\right)\right] e^{-M/2} \left(1 - F^i_y/2\right) \tag{11}$$

$$\hat{B}_y = B^r_y + B^i_y \tag{12}$$

where

$\boldsymbol{\hat{B}^r_y}$ is the mid-year exploitable biomass for the recruitment area in year y,

$\boldsymbol{\hat{B}^r_y}$ is the mid-year exploitable biomass for the immigration area in year y, and

$\boldsymbol{\hat{B}_y}$ is the total mid-year exploitable biomass in year y.

3.2 Generating test data

The purpose of the two-area model is to generate realistic data exhibiting the effects of variations availability, against which the time-varying selectivity models can be tested. The model's parameters are tuned by fitting it to the observed CPUE and catch-at-age data in Table 1 and Table 2 respectively. The fit also provides estimates of the standard deviations of both the CPUE and catch-at-age residuals, which are used to add random noise to generated data according to:

$$CPUE_y = \hat{CPUE}_y e^{\varepsilon^{cpue}_y} \tag{13}$$

$$C'_{y,a} = \hat{C}'_{y,a} e^{\varepsilon^{caa}_{y,a}} \tag{14}$$

where

$\boldsymbol{CPUE_y}$ is the observed CPUE for year y,

$\boldsymbol{\hat{CPUE}_y}$ is the expected CPUE for year y,

$\boldsymbol{\varepsilon^{cpue}_y} \sim N\left(0, \sigma^2_{cpue}\right)$ reflects fluctuations about the expected CPUE for year y,

$\boldsymbol{\sigma_{cpue}}$ is the standard deviation associated with CPUE residuals,

$\boldsymbol{C'_{y,a}}$ is the observed proportion of lobster caught in year y that are of age a (in terms of number),

$\boldsymbol{\hat{C}'_{y,a}}$ is the expected proportion of lobster caught in year y that are of age a (in terms of number),

$\boldsymbol{\varepsilon^{caa}_{y,a}} \sim N\left(0, \left[\sigma_{caa}/\sqrt{C'_{y,a}}\right]^2\right)$ reflects fluctuations about the expected catch-at-age proportion at age a for year y, and

$\boldsymbol{\sigma_{caa}}$ is the standard deviation associated with catch-at-age residuals.

Adding random noise to catch-at-age proportions presents a difficulty as the proportions in a year must always sum to unity. This is solved by renormalising, $C'_{y,a} \rightarrow C'^*_{y,a} = C'_{y,a}/\sum_a C'_{y,a}$.

4 The single-area time-varying selectivity models

4.1 The population dynamics

The single-area models are equivalent to the two-area model with $\mu = 0$, $p_y = 1$ and $S_a \rightarrow S_{y,a}$.

Equations (4), (5) and (6) fall away as there can be no pre-exploitation population in the immigration area and equations (1), (2) and (3) become:

$$N_{y+1,0} = R_{y+1} \tag{15}$$

$$N_{y+1,a+1} = N_{y,a}e^{-M}\left(1 - S_{y,a}F_y\right) \qquad a = 0, \ldots, m-2 \tag{16}$$

$$N_{y+1,m} = N_{y,m-1}e^{-M}\left(1 - S_{y,m-1}F_y\right) + N_{y,m}e^{-M}\left(1 - S_{y,m}F_y\right) \tag{17}$$

where
$\boldsymbol{N_{y,a}}$ is the number of lobsters at age a at the start of year y, and
$\boldsymbol{S_{y,a}}$ is the selectivity of lobsters at age a for year y,
$\boldsymbol{F_y}$ is the fully selected fishing mortality for year y.

The total catch by mass is given by:

$$C_y = \sum_{a=0}^{m} w_{a+\frac{1}{2}} N_{y,a} e^{-M/2} S_{y,a} F_y \tag{18}$$

Expression (12) for mid-year exploitable biomass simply becomes:

$$\hat{B}_y = \sum_{a=0}^{m} w_{a+\frac{1}{2}} S_{y,a} N_{y,a} e^{-M/2}\left(1 - F_y/2\right) \tag{19}$$

4.2 The selectivity functions

4.2.1 Time invariant selectivity

The commercial time-invariant selectivity function is a logistic curve of the form:

$$S_a = \frac{1}{1 + e^{-\kappa\ (a - a_{50})}} \tag{20}$$

where
$\boldsymbol{\kappa} = \frac{\ln 19}{a_95 - a_{50}}$,
$\boldsymbol{a_{50}}$ is the age such that $S_{a_{50}} = 50\%$, and
$\boldsymbol{a_{95}}$ is the age such that $S_{a_{95}} = 95\%$.

4.2.2 Time-varying selectivity - MARAM method

The selectivity function is allowed to vary over the time period for which catch-at-age data are available. The form of the selectivity function is:

$$S_{y,a} = \frac{1}{1 + e^{-\kappa\left[\left(a - \delta_y\right) - a_{50}\right]}} \tag{21}$$

where $\kappa = \frac{\ln 19}{a_{95} - a_{50}}$.

Thus, the model introduces an additional estimable parameters δ_y which shift the selectivity curve along the age-axis. δ_y is only estimated for $y = 1994, \ldots, 2005$, excluding 1999, as these are the years for which catch-at-age data are available. It is assumed that $\delta_{1999} = (\delta_{1998} + \delta_{2000})/2$, while for all other years without catch-at-age data $\delta_y = 0$.

The CPUE estimate in any given year will increase monotonically as δ_y decreases, as a decreasing δ_y increases the selectivity on younger lobsters. Consequently, a CPUE estimate can be chosen to fall anywhere between zero and its maximal value with an appropriate selection of δ_y. To counter this, the selectivity function is rescaled according to:

$$S_{y,a} \to S^*_{y,a} = S_{y,a} / \left[\sum_{a_1}^{a_2} \frac{S_{y,a}}{a_2 - a_1 + 1} \right] \tag{22}$$

In other words, the selectivity function is scaled by the inverse of its average value over a certain age range, so that if the selectivity increase on younger lobsters, $S^*_{y,a}$ will decrease for large a to compensate.

A term is added to the likelihood function to moderate the degree of change in the selectivity function:

$$-\ln L \to -\ln L + \sum_{\substack{y=1994 \\ y \neq 1999}}^{2005} \left(\frac{\delta_y}{\sigma_{sel}} \right)^2 \tag{23}$$

Johnston and Butterworth (2007) found through experimentation that $\sigma_{sel} = 0.75$, $a_1 = 8$ and $a_2 = 12$ gave reasonable performance.

4.2.3 OLRAC method

The selectivity function is of the form:

$$S_{y,a} = \alpha_{y,a} S_a \tag{24}$$

where

$\boldsymbol{S_a}$ is the time-invariant selectivity function described by equation (20),

$\boldsymbol{\alpha_{y,a}}$ is given by:

$$\alpha_{y,a} = \begin{cases} \frac{x_y}{X_y} & \text{if } a \leq 6 \\ \frac{x_y + (a-6)(1-x_y)}{(a_{kink}-6)X_y} & \text{if } 6 \leq a \leq a_{kink} \\ \frac{1}{X_y} & \text{if } a > a_{kink} \end{cases} \tag{25}$$

$\boldsymbol{X_y}$ is given by:

$$X_y = \left\{ \sum_{a=a_1}^{6} x_y + \sum_{a=7}^{a_{kink}} \left[x_y + \frac{(a-6)(1-x_y)}{a_{kink} - 6} \right] + \sum_{a=a_{kink}}^{a_2} 1 \right\} / (a_2 - a_1 + 1) \tag{26}$$

Figure 1 shows the shape of $\alpha_{y,a}$.

The model introduces extra estimable parameters x_y for $y = 1973, \ldots, 2005$. For $y > 2005$ it is assumed that the mean of x_y over its estimable period applies.

A term is added to the likelihood function to smooth the annual variations in selectivity:

$$-\ln L \rightarrow -\ln L + w_{pen} \sum_{y=1973}^{2004} (x_y - x_{y+1})^2 \tag{27}$$

Johnston and Butterworth (2007) found through experimentation that $w_{pen} = 5$, $a_1 = 5$, $a_{kink} = 9$ and $a_2 = 20$ provided reasonable performance.

5 Results

Table 3 presents results for the two-area model as best fitted to data and the median results for each of three single area models with constant or time varying selectivity fitted to 100 pseudo datasets generated from the two-area model. The best-fit time-series of the split of total catch between the recruitment and immigration areas for the two area model is illustrated by Figure 2. CPUE and catch-at-age residuals for the fit of the two-area model are shown in Figures 3 and 4 respectively. Figures 5 to 8 show how the time-varying selectivity functions differ when fitted to the observed data in contrast to the generated test data. Figures 9 to 15 show the performance of the time-varying selectivity models in estimating important fishery management statistics, by comparing estimates to their true values. In these figures solid markers indicate the median estimates across all 100 simulated datasets, while error bars indicate the 5^{th} and 95^{th} percentiles. Finally, Figures 16 and 17 compare the median estimated exploitable biomass and spawning biomass trajectories to their true trajectories as given by the underlying two-area model.

6 Discussion

Table 3 shows that the MARAM model of time-varying selectivity produces a much better fit to catch-at-age data than the time-invariant selectivity model; however the fit to CPUE data is only slightly enhanced. In contrast, the OLRAC model appreciably improves the fit to CPUE data, but performs worse than the MARAM model in fitting catch-at-age data. The MARAM model provides the best overall fit, which is surprising given that the OLRAC model uses twenty-three more estimable parameters then the MARAM model, which itself uses ten more than the time-invariant model.

The MARAM model is the only model which struggles to provide an reasonably accurate estimate of catchability, q.

For years without catch-at-age data, estimates of p_y tends to be at either the upper bound (0.8) or lower bound (0.2). When catch-at-age data is available, p_y changes erratically from year to year.

The catch-at-age residuals (Figure 4) are better behaved than the CPUE residuals (Figure 3) in that they show fewer trends with time. Although neither set of residuals are uncorrelated, the two-area model seems to provide a sufficiently reasonable fit to the observed data to serve as the basis for a test of different single-area model estimators.

Figures 5-8 demonstrate that both time-varying selectivity models can explain the generated data with far smaller variations in their selectivity functions than is required to explain the observed data, to which the two-area model is calibrated. This effect is very pronounced for the OLRAC model. It would be beneficial to further investigate the causes.

The MARAM model and the time-invariant model both outperform the OLRAC model by providing estimates with lower bias and variance; however there is little difference between the estimates of the MARAM model and the time-invariant model (Figures 9-15).

All three models provide accurate estimates of spawning biomass (Figure 17). The MARAM method overestimates exploitable biomass (Figure 16).

References

S.J. Johnston and D.S. Butterworth. The 2007 age-structured production model assessments and projections for the South Coast rock lobster resource - routine update using Pope's approximation model fitting to catch-at-age data including scenarios for time-varying selectivity. Technical Report WG/08/07/SCL05, MCM, 2007.

Table 1: Total annual catch and GLM standardised CPUE data for the South Coast rock lobster fishery as reported in Johnston and Butterworth (2007).

Year	**Total catch (MT of tails)**	**CPUE (kg tails/trap)**
1973	372	
1974	973	
1975	551	
1976	712	
1977	667	0.2187
1978	461	0.2059
1979	122	0.1607
1980	176	0.2041
1981	348	0.1930
1982	407	0.1657
1983	524	0.1958
1984	450	0.1625
1985	450	0.1589
1986	450	0.2074
1987	452	0.1864
1988	452	0.2211
1989	452	0.2050
1990	477	0.1737
1991	524.54	0.1428
1992	529.96	0.1393
1993	524.27	0.1271
1994	507.89	0.1161
1995	504.89	0.1077
1996	442.69	0.0900
1997	416.39	0.0823
1998	516.03	0.0786
1999	512.16	0.0800
2000	423.4	0.0896
2001	288	0.0998
2002	340	0.1107
2003	350	0.1154
2004	382	0.1298
2005	382	0.1136
2006	382	

Table 2: Commerical catch-at-age data for the South Coast rock lobster as reported in Johnston and Butterworth (2007). Note that poor sampling levels in 1999 lead to that year's omission from the assessment.

						Year					
Age	**1994**	**1995**	**1996**	**1997**	**1998**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0039	0.0000	0.0056	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
7	0.0003	0.0006	0.0140	0.0003	0.0201	0.0012	0.0009	0.0011	0.0009	0.0004	0.0004
8	0.0029	0.0093	0.0266	0.0066	0.0484	0.0069	0.0137	0.0190	0.0092	0.0075	0.0059
9	0.0215	0.0554	0.0478	0.0609	0.0834	0.0389	0.0465	0.0510	0.0218	0.0379	0.0223
10	0.0709	0.1265	0.0819	0.1467	0.1233	0.1166	0.0877	0.0767	0.0446	0.0690	0.0540
11	0.1441	0.1838	0.1202	0.2080	0.1429	0.2099	0.1290	0.0930	0.0816	0.0924	0.0989
12	0.1537	0.1369	0.1256	0.1373	0.0939	0.1648	0.1393	0.0986	0.1033	0.1106	0.1108
13	0.1493	0.1110	0.1184	0.1079	0.0844	0.1223	0.1342	0.1143	0.1278	0.1180	0.1186
14	0.1343	0.0829	0.1054	0.0775	0.0744	0.0782	0.1209	0.1242	0.1453	0.1196	0.1203
15	0.0677	0.0440	0.0603	0.0412	0.0462	0.0397	0.0636	0.0708	0.0868	0.0734	0.0733
16	0.0786	0.0548	0.0782	0.0498	0.0637	0.0461	0.0768	0.0927	0.1155	0.1003	0.1003
17	0.0386	0.0342	0.0419	0.0262	0.0361	0.0252	0.0388	0.0510	0.0564	0.0534	0.0557
18	0.0293	0.0319	0.0349	0.0215	0.0315	0.0213	0.0303	0.0434	0.0433	0.0443	0.0479
19	0.0238	0.0274	0.0296	0.0192	0.0271	0.0195	0.0252	0.0368	0.0372	0.0380	0.0419
20+	0.0849	0.1013	0.1113	0.0968	0.1191	0.1096	0.0929	0.1275	0.1266	0.1350	0.1498

Table 3: Results for the underlying two-area model (as best fitted to data) and median results for the fits of the single area model to data simulated from the two-area model. [Note that likelihoods for the two-area model are not comparable to the likelihoods for other models as the former was fitted to different data.]

	Two-area	**Time invariant**	**MARAM method**	**OLRAC method**
K^{sp} (1000 t)	9047.58	8455.13	8304.49	8478.53
h	0.902	0.883	0.892	0.934
M	0.099	0.095	0.106	0.113
a_{50} (yr)	10.58	10.11	10.37	10.07
a_{95} (yr)	13.25	12.73	12.90	12.43
$\ln L_{cpue}$	-40.52	-39.74	-41.86	-46.78
σ_{cpue}	0.150	0.154	0.143	0.121
q	3.69×10^{-5}	3.95×10^{-5}	1.75×10^{-5}	3.31×10^{-5}
$\ln L_{caa}$	-128.46	-131.51	-143.97	-137.81
σ_{caa}	0.061	0.061	0.055	0.058
$\ln L_{SR}$	1.37	5.19	3.27	4.90
$\ln L_{sel}$	-	-	2.91	3.97
$\ln L_{h}$	-0.661	-0.635	-0.648	-0.689
$-\ln L_{total}$	-168.27	-167.01	-180.59	-175.05

Figure 1: Shape of the $\alpha_{y,a}$function. Reproduced from Johnston and Butterworth (2007).

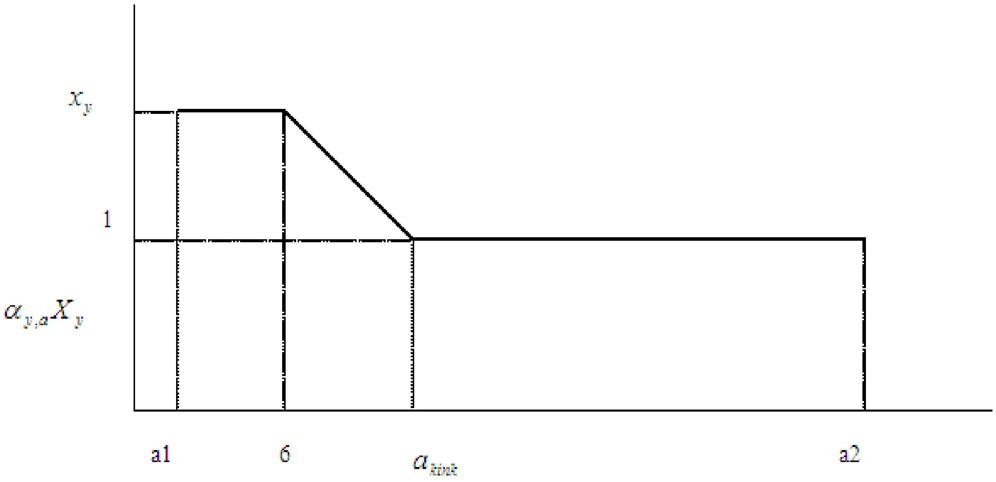


Figure 2: Proportion of the total catch taken in the recruitment area, p_y.

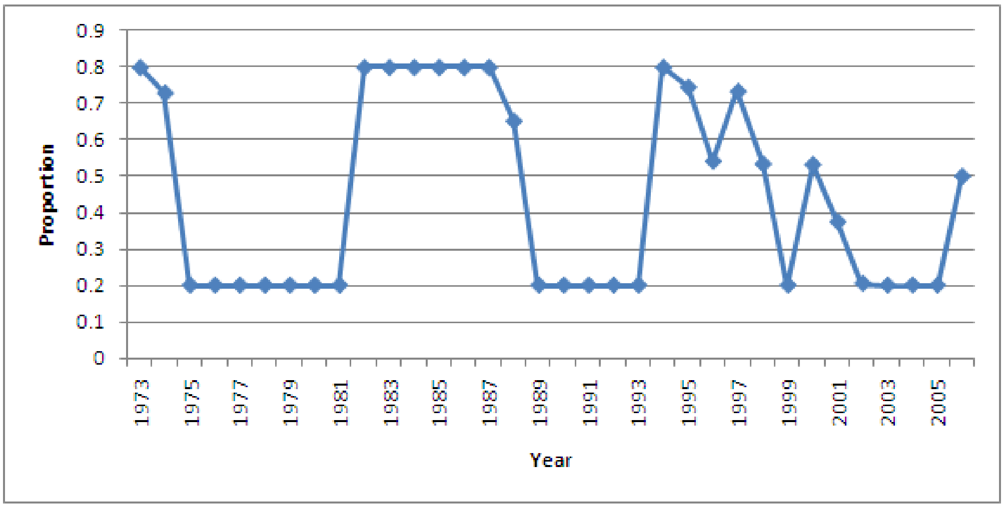


Figure 3: Observed and estimated CPUE for the two-area model.

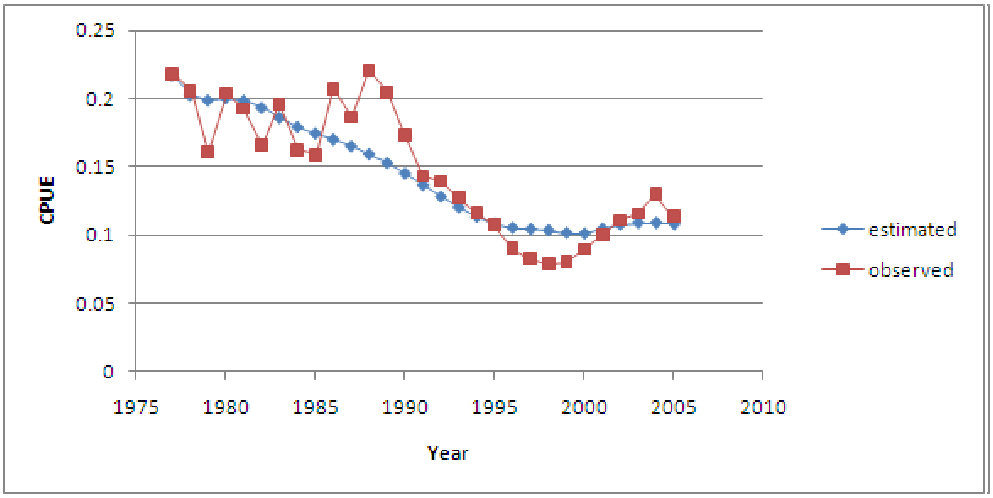


Figure 4: Catch-at-age residuals for the two-area model.

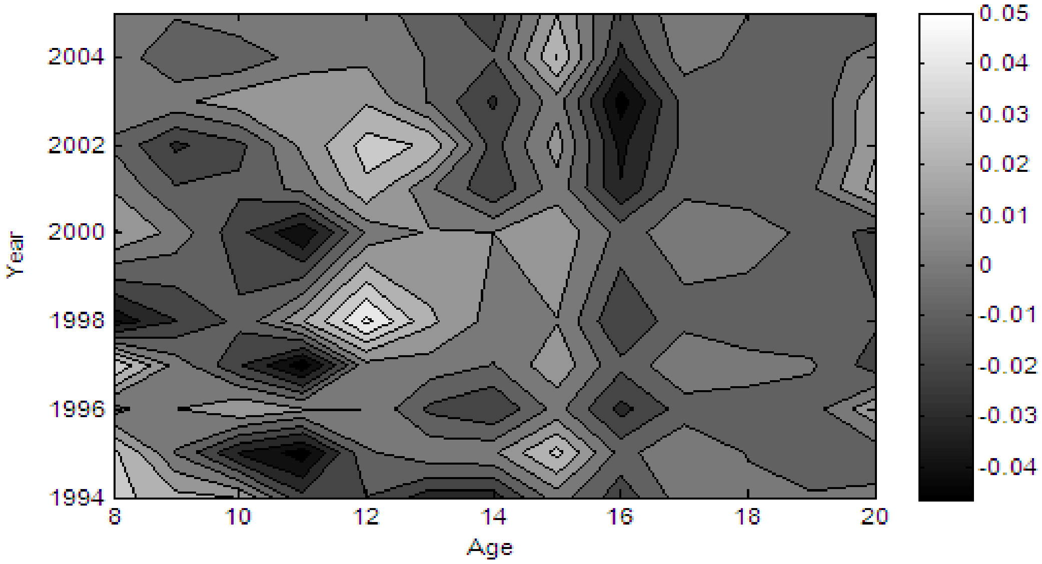


Figure 5: Selectivity function for the MARAM model fitted to the observed data in Tables 1 and 2.

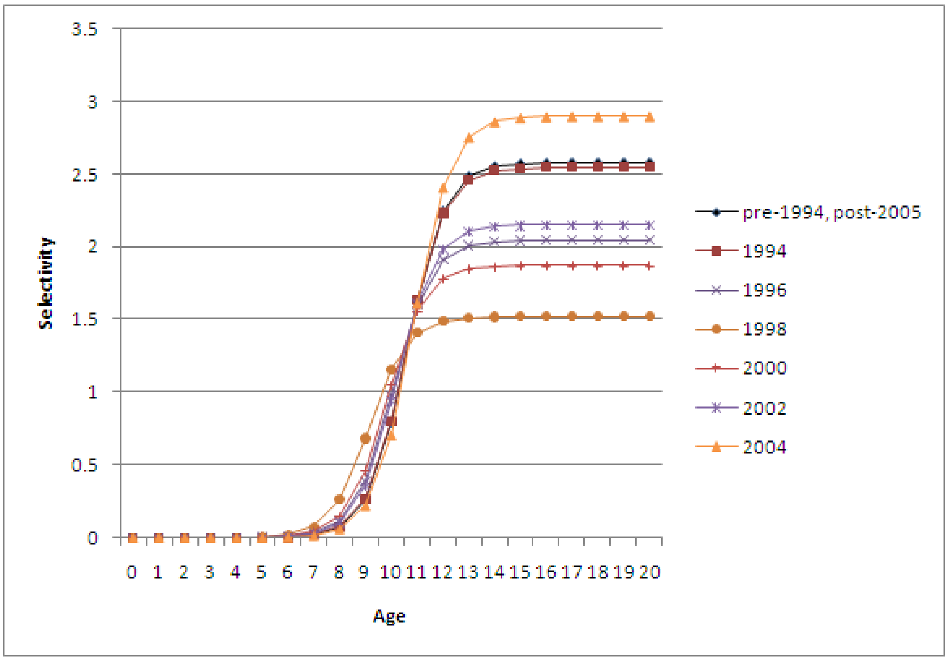


Figure 6: Median selectivity function for the MARAM model fitted to 100 generated test datasets.

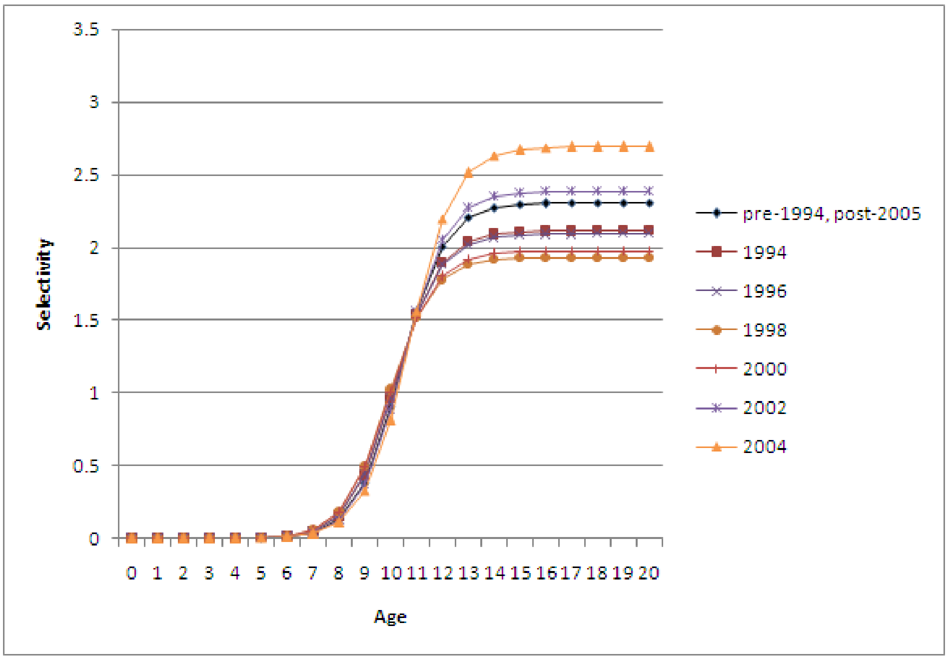


Figure 7: Selectivity function for the OLRAC model fitted to the observed data in Tables 1 and 2.

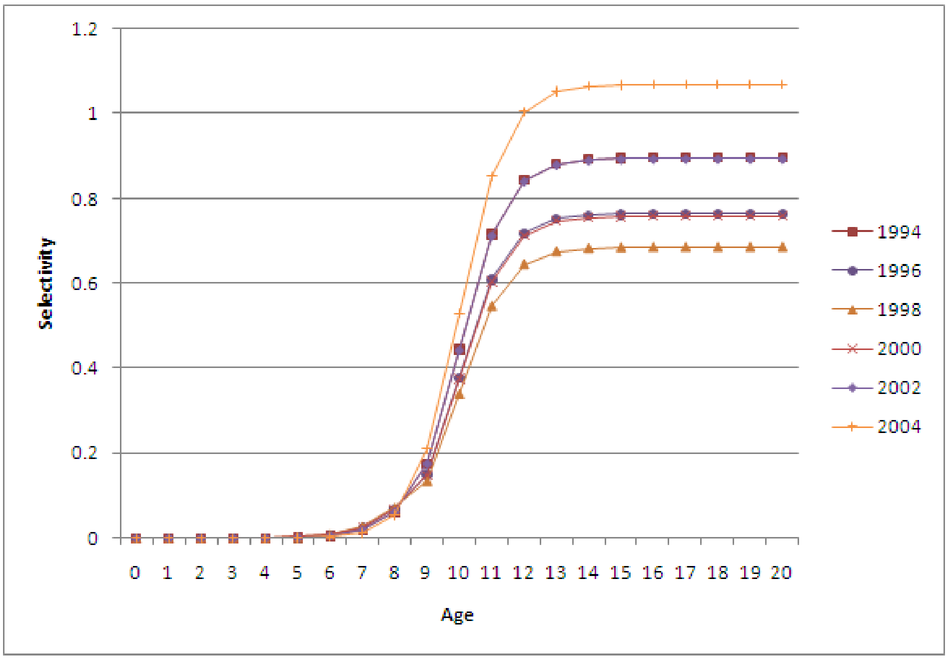


Figure 8: Median selectivity function for the OLRAC model fitted to 100 generated test datasets.

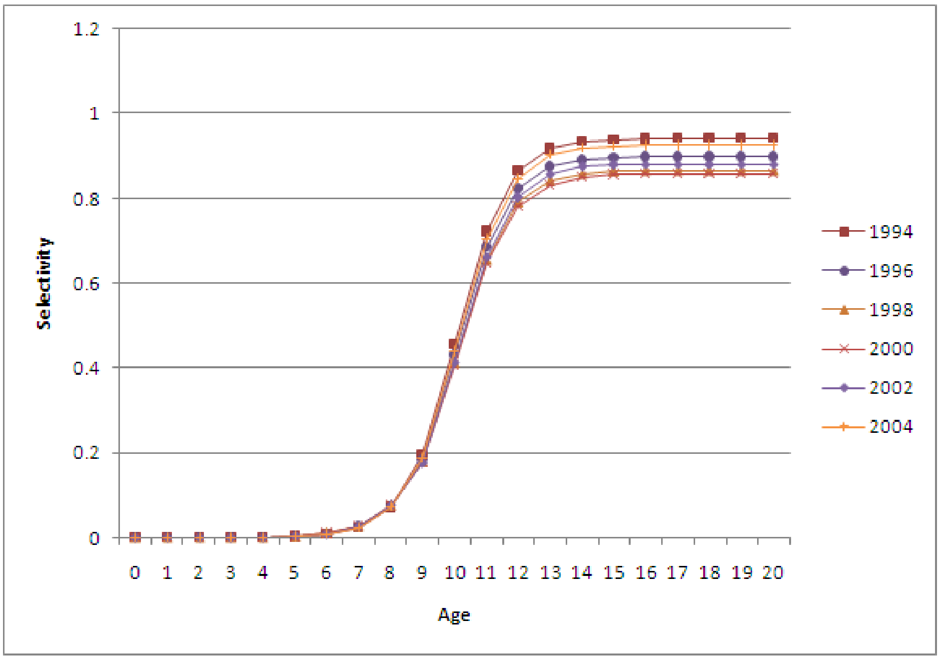


Figure 9: Comparison of the time-varying selectivity models' estimates of MSY. The bars here and in the plots following show 90% probability intervals.

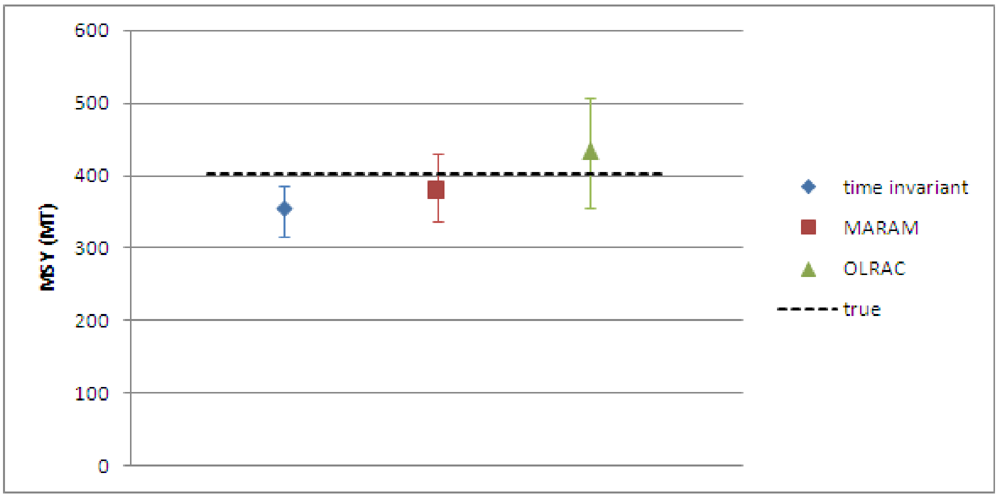


Figure 10: Comparison of the time-varying selectivity models' estimates of $MYSL^{exp}/K^{exp}$.

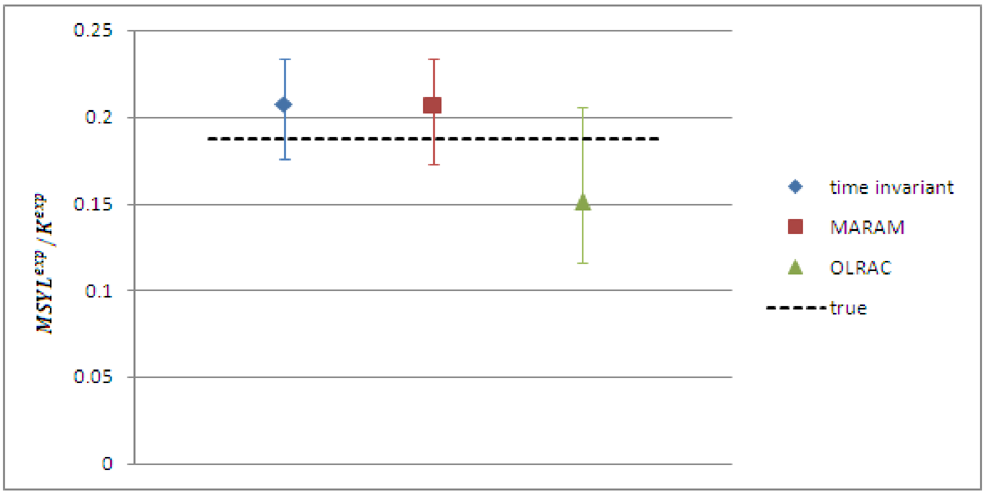


Figure 11: Comparison of the time-varying selectivity models' estimates of current exploitable biomass in relation to MSYL.

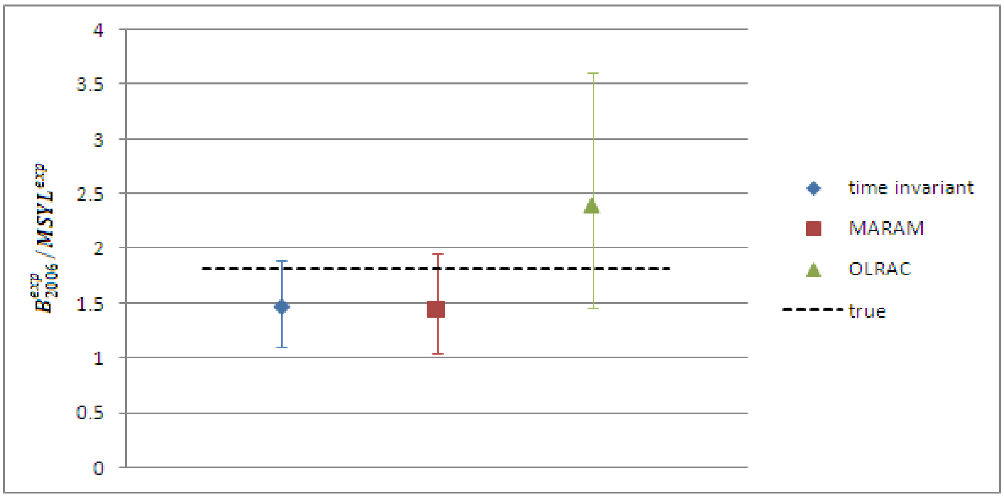


Figure 12: Comparison of the time-varying selectivity models' estimates of current spawning biomass depletion.

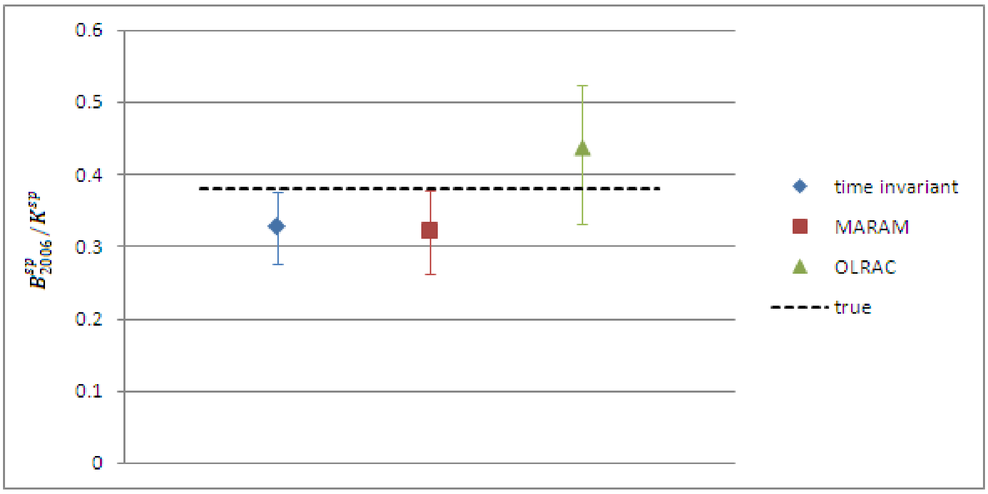


Figure 13: Comparison of the time-varying selectivity models' estimates of current exploitable biomass depletion.

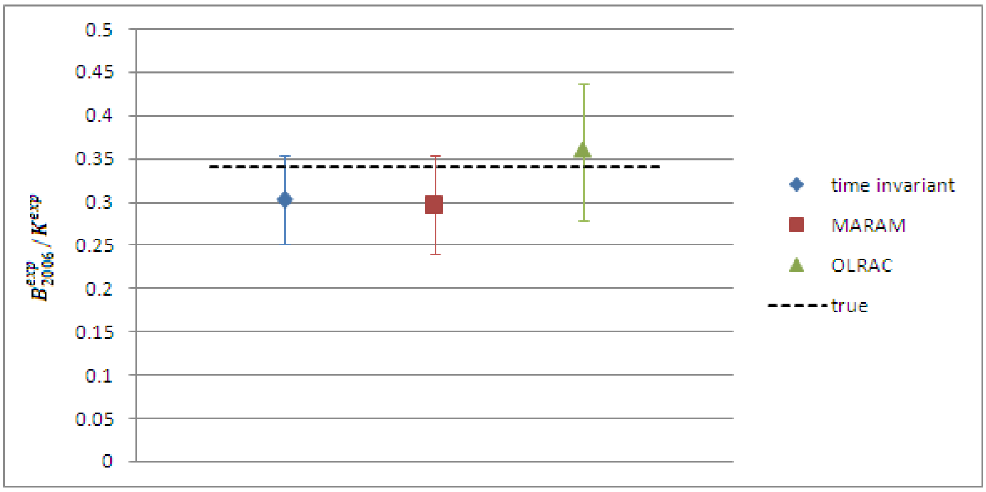


Figure 14: Comparison of the time-varying selectivity models' estimates of pristine spawning biomass.

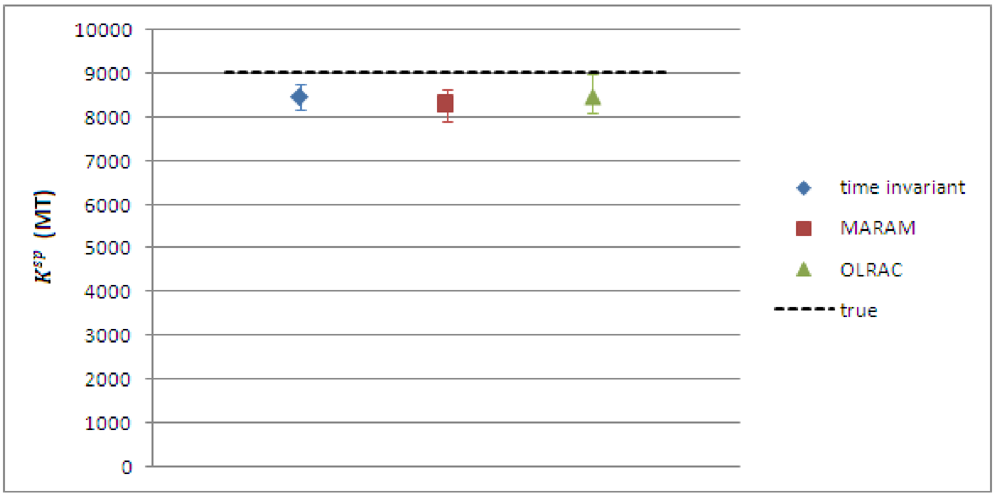


Figure 15: Comparison of the time-varying selectivity models' recommendation for the 2007 catch. The recommended catch is taken as the predicted catch for 2007 if $F_{2007} = F_{MSY}$.

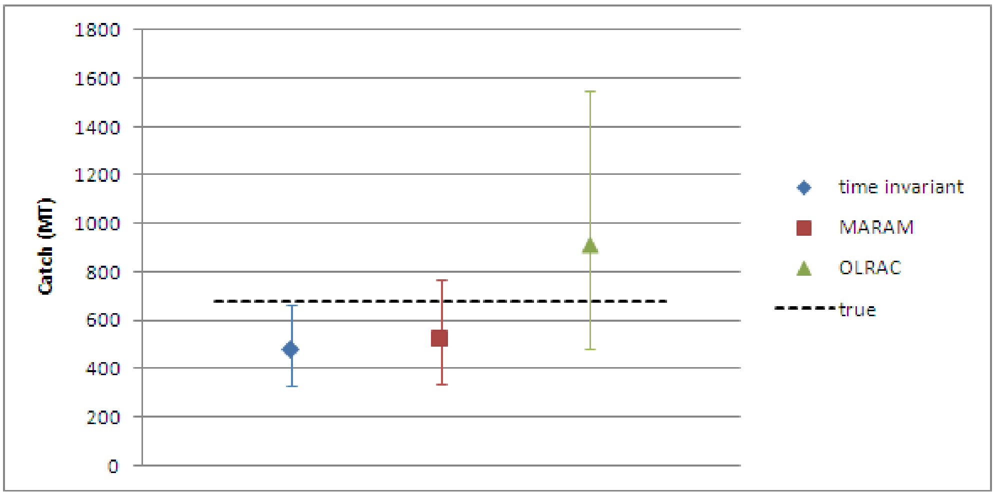


Figure 16: Comparison of the time-varying selectivity models' median estimates of the mid-year exploitable biomass trajectory.

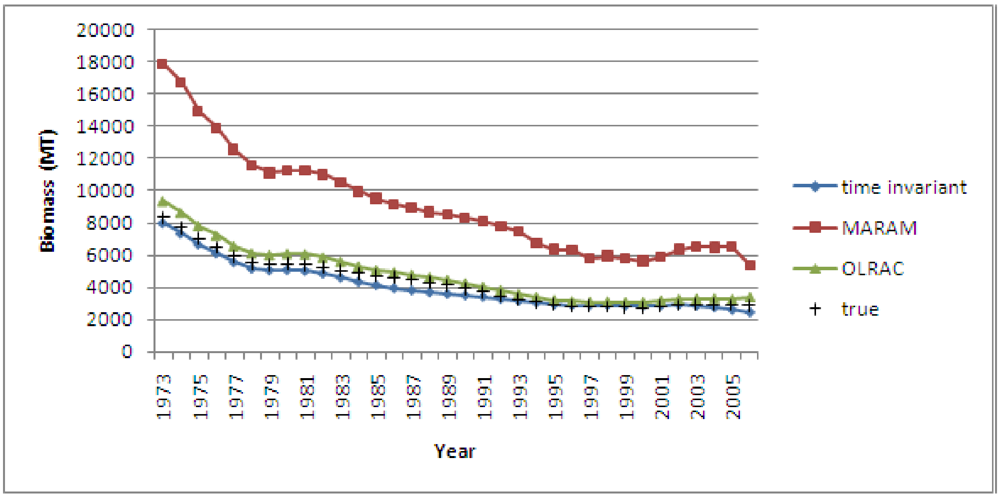
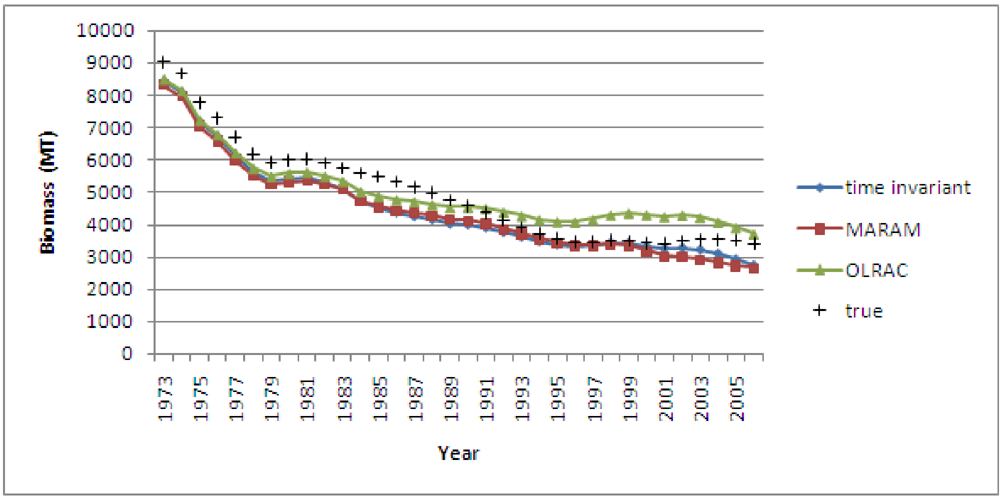


Figure 17: Comparison of the time-varying selectivity models' median estimates of the spawning biomass trajectory.



A Details common to the time-varying selectivity models and the two-area model

A.1 Population dynamics

Spawning biomass is given by:

$$B_y^{sp} = \sum_{a=0}^{m} f_a w_a N_{y,a} \tag{28}$$

where

$\boldsymbol{B_y^{sp}}$ is the spawing biomass of the resource at the start of year y,

$\boldsymbol{f_a}$ is the proportion of lobsters at age a that are sexually mature, and

$\boldsymbol{w_a}$ is the begin-year mass of lobsters at age a.

The number of recruits at the start of year y is related to spawning biomass by a Beverton-Holt relationship:

$$R_y = \frac{4hR_1 B_y^{\mathrm{sp}}}{(1-h)\,K^{\mathrm{sp}} + (5h-1)\,B_y^{\mathrm{sp}}} e^{-\varsigma_y} \tag{29}$$

where

$\boldsymbol{K^{sp}}$ is the pristine spawning biomass,

$\boldsymbol{R_1}$ is pristine recruitment,

$\boldsymbol{h}$ reflects the steepness of the stock-recruitment curve (specifically, it is ratio of recruitment at $B^{sp} = 0.2K^{sp}$ to R_1), and

$\boldsymbol{\varsigma_y}$ describes the fluctuation about the expected recruitment for year y.

The mass of lobsters at age a is given by the growth curve;

$$w_a = \alpha \left[l_\infty \left(1 - e^{-\kappa(a-t_0)} \right) \right]^\beta \tag{30}$$

where the values for the growth parameters are set to the values given by Glazer and Groenveld (1999), as reported by Johnston and Butterworth (2007).

It is assumed all lobsters become sexually mature at age 10:

$$f_a = \begin{cases} 0 & \text{if } a \leq 9 \\ 1 & \text{if } a > 9 \end{cases} \tag{31}$$

A.2 Estimable parameters

The estimable parameters common to all models are:

- pristine spawning biomass, K^{sp};
- natural mortality, M;
- the steepness of the stock-recruitment curve, h;
- the age at 50% selectivity, a_{50};
- the age at 95% selectivity, a_{95}; and
- stock-recruitment residuals, ς_y for $y = 1974, \ldots, 1997$.

A.3 The likelihood function

The models are fitted to CPUE and catch-at-age data. Stock-recruitment residuals and a prior for h also contribute to the negative log-likelihood function.

A.3.1 CPUE

It is assumed that:

$$CPUE_y = q\hat{B}_y e^{\varepsilon_y^{cpue}} \tag{32}$$

where q is the catchability coefficient. The contribution of CPUE data to the negative log-likelihood function is given by:

$$-\ln L = \sum_y \left[\left(\varepsilon_y^{cpue}\right)^2 / 2\sigma_{cpue}^2 + \ln \sigma_{cpue} \right] \tag{33}$$

where

$$\sigma_{cpue} = \sqrt{1/n \sum \left(\ln CPUE_y - \ln \hat{q}\hat{B}_y \right)^2} \tag{34}$$

and where

$\boldsymbol{n}$ is the number of years with CPUE data, and

$\hat{\boldsymbol{q}}$ is the maximum likelihood estimate of q:

$$\ln \hat{q} = 1/n \sum_y \left(\ln CPUE_y - \ln \hat{B}_y \right) \tag{35}$$

A.3.2 Catches-at-age

The contribution of catch-at-age data to the negative log-likelihood function is given by:

$$\ln L = \sum_y \sum_a \left[\ln \left(\sigma_{caa} / \sqrt{C'_{y,a}} \right) + C'_{y,a} \left(\ln C'_{y,a} - \ln \hat{C}'_{y,a} \right)^2 / 2\sigma_{caa}^2 \right] \tag{36}$$

where

$$\sigma_{caa} = \sqrt{\sum_y \sum_a C'_{y,a} \left(\ln C'_{y,a} - \ln \hat{C}'_{y,a} \right)^2 / \sum_y \sum_a 1} \tag{37}$$

Note that a "minus" group of lobsters aged 8 and younger is used when calculating the catch-at-age contribution to ensure that $C'_{y,a} \neq 0$.

A.3.3 Stock-recruitment residuals

It is assumed that these residuals are log-normally distributed and uncorrelated. Therefore, the contribution to the negative log-likelihood function is given by:

$$-\ln L = \sum_{y=1974}^{y=1997} \frac{\varsigma_y^2}{2\sigma_R^2} \tag{38}$$

where σ_R is the standard deviation of the log residuals, which is assumed be equal to 0.4.

A.3.4 Prior distribution for h

The prior distribution for h is assumed to be $N\left(0.95, \sigma_h^2\right)$, where $\sigma_h = 0.2$. Therefore, the contribution to the (penalised)negative log-likelihood is given by:

$$-\ln L = -\ln\left[\frac{1}{\sqrt{2\pi\sigma_h^2}}e^{-(h-0.95)^2/\left(2\sigma_h^2\right)}\right] \tag{39}$$