Robustness test results associated with the new West Coast rock lobster OMP

S.J. Johnston and D.S. Butterworth

Note: A wide variety of variants of the OMP have been evaluated. The aim of the variations was to try and produce narrower probability intervals for the resource recovery statistics (*B*(16/06) values). Variations that were attempted include:

- Alternate values of f_1 and f_2 (relative weighting factors for trap, hoop and FIMS CPUE in the OMP)
- Alternate values of *p* (see Eqn 1 in ASWS/JUL07/WCRL/MP/1)
- Alternate levels of "capping" of input data (see pg section iv of ASWS/JUL07/WCRL/MP/1)
- Alternate forms of the somatic growth "response" term (see pg 16 of ASWS/JUL07/WCRL/MP/1
- Limits on the extent of large inter-annual changes in the input indices, which would seem implausible
- Maintaining future somatic growth rate constant (and removing the somatic growth term from OMP) in an attempt to improve the OMP performance based on future recruitment variability only.

Results (in the form of *B*(16/06) summary statistics) for a range of empirical OMPs were also explored. These OMPs were tuned assuming future somatic growth rate remains low. The purpose was to determine if a more simple, i.e. empirical type OMP could be made to perform adequately or even better that the proposed candidate OMP. The results examined showed that the simple empirical OMPs examined resulted in even wider PIs on the B(16/06) statistic, and produced some very low 5^{th}%iles for A8. The more complicated model-based OMPs appeared to perform the best.

Ranking of robustness tests for OMP testing

Table 1 lists the full range of robustness tests being considered for OMP testing. Responses from OLRAC, MCM and MARAM for each robustness test with respect to whether the test should be included in the final set (Yes/Maybe/No) and the associated "*a priori*" plausibility weighting (H/M/L) are reported. The authors have then attempted to provide an "overall" set of "scores" for each test.

It is provisionally planned that rather than simply treat robustness tests as tick tests (as for previous OMP testing), the Working Group considers moving to a more formal incorporation of these results by an approach along the following lines:

a) agree a resource-wide target *B*(16/06) level (in either median or lower 5%ile terms);
b) rank the robustness tests as of high/medium/low (H/M/L) plausibility; and
c) require that all H tests meet the agreed criterion in a), and that all M tests meet a similar criterion with the *B*(16/06) level set somewhat lower (by an extent to

be agreed); L tests (if any) would purely be inspected to check that performance was not "outrageously" poor.

An initial suggestion is for a resource-wide (male) median target $B(16/06)$ value of 1.20 to be met by the RC and all H robustness tests, with M (including M+ and M-) robustness tests (but excluding test E3) required to meet a target of 1.1.

It is further suggested that an additional robustness test be included which accords 100% weighting to the no somatic growth increase scenario (scenario FSGL re ASWS/JUL07/WCRL/MP/1), and assigns this a plausibility weighting of M – this test is called "SG low".

"Category A" robustness test assessment model output compared with the reference case assessment model.

A number of "Category A" robustness tests have been identified for use in examining the robustness of the new area-disaggregated OMP being developed for the west coast rock lobster. These "A" robustness tests require re-fitting of the size-structured operating model for all five super-areas to input data. Table 2 reports the $-ln$L and the $B75(2005)$ estimates for each of these robustness test re-fits to data. These are compared with those for the reference case (RC, as estimated by MARAM). The column headed "T-RC" gives the [(robustness test $-ln$L total) – (RC $-ln$L total value)]. Negative values thus indicate an improved fit compared to the RC, and positive values a worse fit.

Priority for robustness tests

The final column in Table 1 lists a priority for each test – "I" being highest priority, with "III" being the lowest. Initial priorities were assigned on the following lines:

Priority "I": overall plausibility H or M+
Priority "II": overall plausibility M
Priority "III": overall plausibility M-

Robustness tests with an overall plausibility ranking of "L" are to be excluded from the final set.

Final priorities take account of the overall "inclusion" score (excluding the test if this is N) and the likelihood compared to the RC. Only in the case SG4 is there broad indication of a poorer fit of the model to the data for most of the super-areas, so that this is the only test excluded on the basis of the likelihood value.

Note that category C tests, though high plausibility, are accorded priority III because previous experience indicates that they are unlikely to give rise to difficulties in achieving targets so that they are considered less urgent to complete.

The reason for this prioritisation is that the time constraints may preclude carrying out computations for all tests.

Robustness test results

Due to current time constraints, although most of priority I robustness tests have been run to date, only some of the priority II tests have been completed. These robustness tests have been run assuming the OMP which has the "2245 MT" tuning (for the average commercial TAC). The results are reported in Table 3 (for results for the resource as a whole) and Table 4 (for results of *B*(16/06) on a super-area basis). Results are also presented for the RC model where a constant future commercial catch is assumed. These CC results are either assuming:

i) **CC fixed** (at 2200 MT)**:** the commercial TACs in each super-area remain constant for all years at the following values:
A12 = 30 MT
A34 = 230 MT
A56 = 40 MT
A7 = 590 MT
A8 = 1310 MT

ii) **CC flexible:** the combined future commercial TAC is fixed at 2245 MT, but the OMP is used to split this between the super-areas.

[Note: Because these last results were added in haste, the **CCfixed** ones correspond to a slightly different constant catch than the 2245 MT of the tuning, as that run had been carried out earlier. For the **CC flexible** run, the 10-year average TAC is 2276 MT; this is because the 10 year period for which this statistic is reported includes the 2006 TAC which is already set (higher than 2245 MT). Despite these deficiencies in exact comparability in terms of catch taken, the qualitative features of results for these two cases compared to the OMP remain clear.]

Implications and future work

The relatively large drops evident in Table 3 in the lower 5%-iles of the B(16/06) distributions for the resource as a whole for the priority I robustness tests compared to the Reference Case result are of some concern. This concern increases given that Table 4 indicates that this 5%-ile drops to zero for most of these tests for the A7 super-area in isolation.

The fact that this particular statistic is notably worse for A7 for the **CC flexible** compared to the **CC fixed** option (see Table 4) suggests that the OMP may have been "over-tuned" for "optimality" for the Reference Case. In particular the parameters for the "transfer" of TAC from A8 to A7 (see item vi) on pg 17 of ASWS/JUL07/WCRL/MP/1) will need to be revisited, hoping that a lesser size for this transfer may be sufficient to provide adequate robustness test performance for super-area A7 without the need to consider a more conservative approach overall.

A further option which it is planned to investigate is that of allowing for some variation in the limited rights holder allocations in each super-area (currently taken to be fixed - see section 6 on pg. 10 of ASWS/JUL07/WCRL/MP/1), in line with resource trends. The current situation of fixed amounts is problematic particularly for super-areas A1-2 and A5-6 in which only these limited holders are permitted to operate given the small TACs available there. The option would be on similar lines to the manner in which the recreational allocation is adjusted (see pg 4 of

ASWS/JUL07/WCRL/MP/1): *viz*. allocations do not vary continuously with resource index values, but rather are held fixed unless certain thresholds are breached, which leads to adjustments to quotas of fixed amounts. Trials will investigate whether such a modification leads to improved performance in terms of lessening resource risk to a meaningful extent.

Tables 3 and 4 show that in broad terms the feedback performance of the form of OMP proposed is relatively poor. When compared to constant catch "equivalents", the OMP is unable to secure much improvement in the risk statistic, the lower 5%-ile of B(16/06). The investigations reported in the opening Note of this document suggest that this is not a reflection of poor OMP construction – many approaches were considered, but all demonstrated worse performance in terms of risk than the form of OMP tested here. The underlying reason seems likely to be the poor precision of the resource indices available (see their associated large CVs as listed on pg 6 of ASWS/JUL07/WCRL/MP/1), which means that trends in these indices have relatively weak ability to detect trends in the underlying resource abundance, and hence provide a good basis to adjust TACs in appropriate response to those trends. A future priority is to improve both FIMS (perhaps by less intensive but more frequent repeat surveys each year) and CPUE indices (perhaps through collecting spatial data on catch and effort at a finer scale) to improve the precision and hence information content of the inputs to any OMP.

Table 1: Possible list of robustness and other tests for evaluation in 2007. MARAM, MCM and OLRAC responses as well as a suggested "integrated" response for each test are given, together with a "priority" rating (see text) for the order in which these are computed. "EX" indicates "exclude".

CATEGORY A TESTS	Description	**Inclusion Yes/Maybe/No**				**Plausibility weighting H/M/L**				**Priority**	
		MARAM	**MCM**	**OLRAC**	**OVERALL**	**MARAM**	**MCM**	**OLRAC**	**OVERALL**	**Initial**	**Final**
NS1	Male natural survivorship = 0.88	Y	Y	Y	**Y**	H	H	M	**H**	**I**	**I**
NS2	Male natural survivorship = 0.92	Y	Y	Y	**Y**	H	H	M	**H**	**I**	**I**
D2	Discard mortality = 0.2	M		N	**M**	H	M		**M+**	**I**	**I**
D3	Discard mortality increases 5 yrs prior to min size change	M	Y	N	**M**	L	H		**M**	**II**	**II**
SG1	Adult growth is 0.5mm more than thought	Y	Y	Y	**Y**	L	H	M	**M**	**II**	**II**
SG2	1910-1967 growth = 68-88 average	M	M	N	**M**	H	M		**M+**	**I**	**I**
SG3	Pre-1990 growth shifted down to 1990+ average level	M	M	M	**M**	M	M		**M**	**II**	**II**
SG4	1990+ growth shifted up to pre-1990 average level	Y	Y	M	**Y**	L	M-H		**M**	**II**	**EX**
W1	1990+ 225 MT walkout each yr* (but not in future)	M	Y	N	**M**	M	L		**M-**	**III**	**III**
W2	Once every decade 1910-1990 500 MT walkout	M	M	N	**M**	M	L		**M-**	**III**	**III**
B4	Hoop and trap CPUE 99-04 negatively biased by a factor of 1.3	Y	Y	Y	**Y**	L			**L**	**EX**	**EX**
CATEGORY B TESTS											
E1	R drops 50% for 3 years, once in 1998-2006	Y	Y	Y	**Y**	M	H	M/L	**M**	**II**	**II**
E3	25% all lobsters die once during 2006-	Y	Y	Y	**Y**	M	H	M/L	**M**	**II**	**II**

	2015										
P1	Poaching reduced next 5 years to 200 MT	M	M	M	**M**	M	L		**M-**	**III**	**II**
TH1	Future trap:hoop changes? (see bottom for details)	N	N	N	**N**	M	L		**M-**	**III**	**EX**
B1	CPUE 2007+ stays constant	Y	M	Y	**Y**	M	M		**M**	**II**	**II**
B2	Future adult somatic growth 0.5mm than reported	Y	Y		**Y**	L	M		**M-**	**III**	**III**
B3	Future adult somatic growth 0.5mm less than reported	Y	Y		**Y**	L	H		**M**	**II**	**II**
W1 future	Future walkouts continue at 1990s rate	M	Y	N	**M**	M	H		**M+**	**I**	**I**
W3	W1 above, but 400 MT walk-out annually 2006+*	M		N	**M**	L	L		**L**	**EX**	**EX**
COMP	Hard combination of tests	Y	Y	N	**Y**	M	H		**M+**	**I**	**I**
RECR1	Future recreational take is ?	Y	N	N	**N**	M	L		**M-**	**III**	**EX**
SG low	Future somatic growth remains low for all simulations								**M**	**II**	**II**
CATEGORY C TESTS	**(How to combine super-area data when some are not available)**										
M1	FIMS index missing	Y	Y	N	**Y**				**H**	**III**	**III**
M2	Somatic growth index missing	Y	Y	N	**Y**				**H**	**III**	**III**
M3	Trap CPUE index is missing	Y	Y	N	**Y**				**H**	**III**	**III**
M4	Hoop cpue index is missing	Y	Y	N	**Y**				**H**	**III**	**III**
CATEGORY D TESTS	**(How to split global TAC into super-area TACs)**										
DD1	Split global TAC at current (2006) TAC proportions throughout the period	Y	Y	Y	**Y**			L	**L**	**EX**	**EX**

* Super-area division to be specified after discussion

For M1-M4: the OMP would assume the average of the previous 3 years' data

Table 2: Comparison of "Category A" robustness test results with those of the reference case (as estimated by MARAM). The "T-RC" column refers to the total robustness test –*ln*Ls less the total RC –*ln*L values.

		-*ln*L						***B*75(2005)**				
		T-RC	**A12**	**A34**	**A56**	**A7**	**A8**	**A12**	**A34**	**A56**	**A7**	**A8**
RC		**0**	**-20.45**	**24.99**	**29.78**	**0.34**	**-55.23**	**526**	**3104**	**1326**	**4944**	**9386**
NS1	Male natural survivorship = 0.88	**-1.28**	-20.86	25.45	28.17	0.27	-54.88	641	3345	1723	5612	11238
NS2	Male natural survivorship = 0.92	**2.19**	-19.91	25.55	26.91	4.76	-55.69	532	2903	1516	4412	8134
D2	Discard mortality = 0.2	**3.68**	-20.15	24.43	32.91	1.31	-55.39	544	3132	1321	5055	9494
D3	Discard mortality increases 5 yrs prior to min size change	**-2.93**	-20.23	24.39	26.69	0.71	-55.06	526	3063	1358	5178	9552
SG1	Adult growth is 0.5mm more than thought	**-22.09**	-17.06	30.29	9.66	-12.13	-53.42	337	2872	1588	4190	8390
SG2	1910-1967 growth = 68-88 average	**-17.73**	-19.91	28.00	15.94	-6.09	-56.24	532	2951	1266	5163	9214
SG3	Pre-1990 growth shifted down to 1990+ average level	**17.6**	-20.45	48.80	19.80	-6.58	-44.64	526	7587	1741	5916	16980
SG4	1990+ growth shifted up to pre-1990 average level	*****	*	71.86	89.16	*	-30.03	*	3659	932	*	5184
W1	1990+ 225 MT walkout each yr* (but not in future)	**-2.07**	-20.45	23.34	29.36	0.34	-55.23	526	2815	975	4944	9386
W2	Once every decade 1910-1990 500 MT walkout	**1.06**	-20.45	25.12	30.71	0.34	-55.23	526	3180	1340	4944	9386
B4	Hoop and trap CPUE 99-04 negatively biased by a factor of 1.3	**3.57**	-23.01	22.29	33.23	2.10	-51.61	923	4182	1411	6485	13462

* Convergence not achieved

Table 3: Robustness test results using the "2245 MT" tuned OMP. Median values are presented with values in parenthesis being the 5th and 95th %iles. These results refer to the resource as a whole.

TEST		***B*(16/06)**	TAC_{comm}^{ave}	**Effort(16/06)**
Reference Case		**1.26 [0.62; 3.00]**	**2245 [1831; 2587]**	**0.72**
CC fixed		**1.24 [0.53; 2.98]**	**2200 [2200; 2200]**	**-**
CC flexible				
Priority I tests				
NS1	Male natural survivorship = 0.88	1.40 [0.38; 3.57]	2152 [1358; 2753]	0.73
NS2	Male natural survivorship = 0.92	1.45 [0.46; 4.36]	1988 [1364; 2643]	0.63
D2	Discard mortality = 0.20	1.36 [0.48; 6.38]	2045 [1400; 2633]	0.68
SG2	1910-1967 growth = 68-88 average	1.21 [0.45; 3.95]	1990 [1375; 2539]	0.79
W1 future	Future walkouts continue at 1990s rate	1.45 [0.52; 4.14]	2210 [1806; 2625]	0.67
Priority II tests				
SG low	Future somatic growth remains low for all simulations	1.07 [0.54; 2.21]	2118 [1788; 2385]	0.73
E1	R drops 50% for 3 years, once in 1998-2006	1.03 [0.49; 2.54]	2203 [1805; 2568]	0.85
E3	25% all lobsters die once during 2006-2015	0.81 [0.35; 2.31]	2125 [1699; 2540]	1.02
P1	Poaching reduced next 5 years to 200 MT	1.37	2253	0.64

Table 4: Robustness test results using the "2245 MT" tuned OMP. Median values are presented with values in parenthesis being the 5th and 95th %iles. These results refer to the individual super-areas *B*(16/06) values.

	A12	**A34**	**A56**	**A7**	**A8**
RC	0.79 [0.50; 1.32]	1.06 [0.62; 2.58]	1.78 [0.61; 11.29]	1.26 [0.36; 3.26]	1.06 [0.39; 2.83]
CC fixed	0.77 [0.48; 1.31]	1.00 [0.55; 2.56]	1.75 [0.58; 11.26]	1.29 [0.42; 3.44]	0.96 [0.21; 2.86]
CC flexible					
NS1	0.93 [0.11; 39.73]	1.00 [0.33; 4.03]	2.11 [0.24; 16.89]	1.26 [0; 3.25]	1.07 [0.03; 3.17]
NS2	0.86 [0.28; 5.83]	1.18 [0.48; 6.20]	1.78 [0.46; 9.88]	1.56 [0; 3.44]	0.97 [0.18; 2.73]
D2	0.81 [0.10; 2.28]	1.12 [0.40; 6.82]	1.47 [0.31; 19.79]	1.42 [0; 3.99]	1.06 [0.39; 2.77]
SG2	0.63 [0.36; 1.58]	0.87 [0.35; 5.44]	1.33 [0.23; 19.66]	1.07 [0; 3.49]	1.03 [0.41; 2.54]
W1 future	0.79 [0.50; 1.32]	1.06 [0.29; 5.53]	2.22 [0; 20.06]	1.35 [0.55; 3.30]	1.00 [0.42; 2.79]
SG low	0.79 [0.51; 1.33]	0.95 [0.56; 2.01]	1.55 [0.55; 8.48]	1.25 [0.41; 3.10]	0.77 [0.33; 1.53]
E1	0.66 [0.42; 1.12]	0.94 [0.57; 2.21]	1.55 [0.56; 9.88]	1.09 [0.27; 3.01]	0.77 [0.30; 2.19]
E3	0.52 [0.29; 0.96]	0.78 [0.43; 2.01]	1.33 [0.43; 0.78]	0.89 [0.17; 2.69]	0.58 [0.16; 1.94]