Updating and modifying the 2020 South Coast Rock Lobster assessment

Johnston, S.J. and Butterworth, D.S.

KEY WORDS: South Coast rock lobster, stock assessment

Summary

A proposed new 2020 assessment approach (which takes account of updated data) seems to be able to side-step the previous conflict between the CPUE and CAL data (with regard to current abundance estimates) and allows for sensible estimation of the λ parameters that allocate recruitment amongst sub-areas).

Key words: spatial assessment, catchability, Rock lobster, South Coast

Introduction

Two issues have arisen in past SCRL assessments:

(1) The estimation of the sub-area λ parameters (proportional split of recruitment amongst sub-areas) has been problematic; for some assessments these parameters were thus fixed (for example the 2017 assessment). This problem appeared to be evident again with the initial 2020 updated assessment, with the result being the λ_{A1E} was estimated to be close to 1.0 (i.e. all recruitment going into sub-area A1E).

(2) Sensitivity to down-weighting the CAL data.

Previous SCRL assessments have shown that down-weighting the CAL data produces different results from the RC (which gives equal weight to both the CPUE and CAL data). This feature remained evident in the 2018 assessments but the differences were ameliorated somewhat. Down-weighting the CAL data (by a multiplicative factor of 0.1 in the negative log-likelihood) data produced more optimistic results.

Table 1 reports results of the initial 2020 updated assessment – i.e. the same model structure as in 2018 except with more data now available. In the initial 2020 updated assessment it was evident that the fitting procedure had difficulty obtaining a sensible fit to the data, and would prefer a solution where virtually all recruitment was allocated to sub-area A1E (λ^{A1E}=0.99) – quite the reverse of a year previously where nearly all recruitment was allocated to sub-area A2+3. It was clear that (as in some cases in the past), the best fit resulted in the model estimating vary large λ_{A1E} - although the –lnL values were good! After much exploration into why and how this could occur, it was found that the model was able to produce good CPUE fits in A1E (which was estimated to have a ridiculously large biomass) by adjusting the catchability "*q*" value for that sub-area. The catchability values across sub-areas were different by several orders of magnitude.

Proposed new updated 2020 model to deal with both problems above

STEP 1

The first aspect examined was to fix the λ^{A1E} values at fixed levels over a range between 0.1 and 1.0, and then refit the model to examine the different contributions to –lnL more closely (note that to facilitate readier comparison, these are shown as values of -lnL DIFF, which subtracts from each contribution its minimal values across the λ^{A1E} range considered. The first set of results shown in Figure 1 compares the overall contributions from the CPUE and the CAL data. A data conflict is immediately apparent: the CPUE data do not favour values of λ^{A1E} below about 0.3, whereas the CAL data hardly support values of λ^{A1E} above about 0.3. The relative weighting given to these two contributions to -lnL is somewhat arbitrary, so that the assessment result is heavily dependent on that weighting choice, which then determines the "optimal" value of λ^{A1E} found by the fit.

Figure 2 repeats the plots Figure 1, but now showing the contributions from each sub-area separately. This makes clear that the contributions from sub-area A2+3 dominate for both CPUE and CAL, while at the other extreme the sub-area A1W contributions to -lnL are hardly informative about the value of λ^{A1E} . Figure 3 shows that the main consequence of changing the value of λ^{A1E} is to change the distribution of biomass across the overall area: a low value of λ^{A1E} sees most of the exploitable biomass in sub-area A2+3, with little in A1E, with the reverse applying for a high value of λ^{A1E}; the biomass in sub-area A1W is hardly affected.

Overall then, these results suggests that that likelihood values can be used to exclude parts of the λ^{A1E} range, but not to estimate λ^{A1E} itself satisfactorily. Probably only the range from about 0.15 to 0.35 for λ^{A1E} could be argued to be reasonably consistent with the CPUE and CAL data. However, other approaches need to be sought to attempt to narrow this range of values further.

This prompted a closer look at the estimates of the catchability (*q*) parameters which link CPUE to exploitable biomass. One might expect these to be similar in some way across sub-areas, but Figure 4 shows both that two differ depending on the value specified λ^{A1E}, and all three appreciably across the sub-areas. However, to take this idea further, it is necessary to remember that CPUE is basically an index of lobster density in a sub-area, not of biomass *per se*, so that an estimate of the size of the surface area within each sub-area where lobsters are to be found (called the "effective fishing area" below) is needed to relate indices of density to estimates of abundance.

The Appendix describes how these effective areas were calculated. In relative terms these effective $Area^{sub-area}$ values for each sub-area were estimated to be:

$$Area^{A1E} = 0.15$$

$$Area^{A1W} = 0.20$$

$$Area^{A2+3} = 0.65$$

Since CPUE relates more to density than to biomass, one can modify the basic relationship:

CPUE = q Bexp

to *CPUE = q*Bexp/Area*

One then might expect the modified q* values to be similar across sub-areas, where:

$$q^{*,sub-area} = q^{sub-area} Area^{sub-area}$$

The next step then is to modify the assessment to reflect that outcome.

Note that GLM standardised CPUE values have been used in the analyses reported here. This is potentially problematic when *q* values are interpreted as above, as the standardisation may impact absolute values compared to those for nominal CPUE differently for the different sub-areas. However, an investigation of the size of this effect showed that these differences were considered sufficiently small that no adjustment for this needed to be made at this stage.

STEP 2

This step was effected by adding a further penalty function to the overall –*ln*L function as follows:

$$q^{pen} = \omega \sum_{sub-area} (q^{*,sub-area} - q^{*,ave})^2$$

where $$q^{*,ave} = \frac{\sum_{sub-area} q^{*,sub-area}}{3}$$

Figure 5 plots the modified *q* values (*q**) across fixed values of λ^{A1E}. Results are compared for either no q^{pen} (ω*=0), and for two alternate larger values of ω* of 19 and 23, where to keep values in "reasonable" ranges we work with ω* where ω = exp(ω*).

The higher the ω*value, the more similar the *q** values for each sub-area.

STEP 3

The model was then extended so that all three λ values are estimated, and results computed for a range of ω* values from 17 to 23. Sensible model fits were obtained, and with positive-definite Hessians.

The question is how close should the different *q**'s be constrained to be? The assumption that the *q**'s should be similar for all three sub-areas is a major one, so sensitivity to different weights ω* for the q^{pen} were then explored. For further insight on this, Figure 6 plots as a function of ω* the difference (DIFF) of "-lnL total – q^{pen}" from its value for the highest value of ω*=23. This shows the extent to which the fit to the other information (primarily the CPUE and CAL data) is being compromised by imposing the penalty of the q^{pen} function.

Figure 7 plots the exploitable biomass in each sub-area (across ω^* values), and Figure 8 plots the spawning biomass. Finally Figure 9 plots the *q** values.

Table 1 provides details of the –lnL contributions, the estimated λ values as well as recent and current Bsp/K estimates. Table 2 repeats this in summary form for different choices for the value of ω*. This information shows again the opposite trends in –lnL contributions between the CPUE and CAL data.

From Figures 7 and 8 it is clear that the model for ω*=17 provides an unrealistic fit (this model puts effectively all the recruitment into sub-area A1E). Models for which ω*=18-23 provide more sensible fits to the data.

The way forward

Across the "sensible" range for ω*of 18-23, results in Table 2 indicates a range from 0.27 to 0.24 for λ^{A1E}. This lies well within the range of 0.15 to 0.35 suggested in STEP 1 not to result in serious misfits to either the CPUE or the CAL data.

For OMs to be used in an OMP update for the resource, perhaps the choice of ω*=23, corresponding to λ^{A1E}= 0.24 would provide a suitable reference case. However, λ^{A1E} is certainly not well determined, and robustness to alternatives would need to be tested – possibly λ^{A1E}values of 0.2 and 0.3 fitted for ω* values of both 18 and 23.

Note that the estimation of the q's becomes important with respect to the calculation of a combined CPUE across the three sub-areas for input into the OMP formula. The current relative split of the weights[1] used in this formula (dependent on the most recent assessment) are:

v^{A1E}=0.006

v^{A1W}=0.006

v^{A2+3}=0.988

i.e. sub-area A1E and A1W CPUEs get virtually no weight (as a result of the large difference in biomass across sub-areas).

The updated 2020 assessment (given the inclusion of the q penalty described above) with the q pen ω^{*}=23 would result in a quite substantial change to these weights as follows:

v^{A1E}=0.163

v^{A1W}=0.156

v^{A2+3}=0.681

Thus sub-areas A1E and A1W CPUE would consequently get a larger weight in the overall "recent CPUE" calculation when simulation testing new OMP candidates.

[1] Note that in earlier documents this weight was accorded the symbol λ , but a different symbol v has been used here as λ is used in this document for the proportional split of recruitment amongst the sub-areas.

Conflict between CPUE and CAL data

The updated proposed new 2020 assessment model appears to avoid this conflict between the CPUE and CAL data (with regards to current abundance estimates) and allows for sensible estimation of the λ parameters.

Table 1 shows that the recent spawning biomass estimates relative to Ksp are very similar for both CAL weights (Bsp(2019)/Ksp = **0.37** for CAL wt =1.0 and =**0.36** for CAL wt = 0.1). The 2018 assessment showed a much bigger difference in recent spawning biomass estimates between the two CAL wt options (Bsp(2018)/Ksp = **0.29** for CAL wt =1.0 and =**0.36** for CAL wt = 0.1).

Note that the 2020 assessment with CAL wt = 1.0 (Bsp(2019)/Ksp = **0.37**) is more in line with the more optimistic 2018 model where CAL data were down-weighted by a factor of 0.1 (Bsp(2018)/Ksp = **0.36**).

Table 1: Negative log likelihood values and other parameters obtained for four 2020 assessment models: CAL equal weight to CPUE (CAL wt 1.0) and CAL down-weighted (CAL wt 0.1) for both model variants that either exclude a q^{pen} weighting or include a q^{pen} weighting of $\omega^* = 23$. Results for the lower CAL weight used in the 2018 assessment are shown in the right hand columns of each pair of columns below. Circled vales reflect implausible results.

	2018 assessment model		2020 updated assessment (as for 2018 model)		2020 updated assessment (with q^{pen})	
	No q^{pen} wt	No q^{pen} wt	No q^{pen} wt	No q^{pen} wt	ω^* =23	ω^* =23
	CAL wt 1.0	CAL wt 0.1	CAL wt 1.0	CAL wt 0.1	CAL wt 1.0	CAL wt 0.1
-lnL T	**-450.63**	**-180.40**	**-573.66**	**-201.07**	**-562.16**	**-197.38**
q^{pen}	-	-	-	-	0.10	0.0001
-lnl CPUE	**-122.84**	**-185.45**	**-115.53**	**-190.23**	**-119.36**	**-188.09**
-lnl CPUE A1E	-24.09 (0.33)	-24.56 (0.32)	-22.43	-23.94	-23.61	-25.16
-lnl CPUE A1W	-49.44 (0.17)	-65.36 (0.11)	-51.89	-71.79	-50.60	-68.53
-lnl CPUE A2+3	-49.30 (0.17)	-95.53 (0.05)	-41.21	-94.51	-43.14	-94.40
-ln SCI CAL	**-408.40**	**-84.79**	**-515.43**	**-129.80**	**-504.55**	**-117.32**
-ln CAL A1E	12.39 (0.15)	20.55 (0.16)	26.27	20.82	27.15	20.21
-ln CAL A1W	-143.21 (0.08)	-68.19 (0.10)	-190.29	-95.06	-188.97	-88.11
-ln CAL 2+3W	-277.58 (0.06)	-37.15 (0.10)	-351.41	-55.57	-342.73	-49.42

	2018 assessment model		**2020 updated assessment (as for 2018 model)**		**2020 updated assessment (with q^{pen})**	
	No q^{pen} wt	**No q^{pen} wt**	**No q^{pen} wt**	**No q^{pen} wt**	ω^* **=23**	ω^* **=23**
	CAL wt 1.0	**CAL wt 0.1**	**CAL wt 1.0**	**CAL wt 0.1**	**CAL wt 1.0**	**CAL wt 0.1**
Bsp(2019) (Bsp(2010)/K)	945 (**0.29**)	1429 (**0.36**)	4797050 (0.85)	3401790 (0.84)	1225 (**0.37**)	1615 (**0.36**)
Bsp(2020) (Bsp(2020)/K)	-	-	4749500 (0.84)	3439620 (0.85)	1182 (0.35)	1663 (0.37)
$B_{exp}(2018)$ ($B_{exp}(2018)/K_{exp}$) A1E	199 (0.44)	179 (0.56)	1090580 (0.85)	4259330 (0.86)	500 (0.58)	278 (0.62)
$B_{exp}(2018)$ ($B_{exp}(2018)/K_{exp}$) A1W	168 (0.12)	179 (0.28)	374 (0.24)	263 (0.39)	481 (0.27)	312 (0.45)
$B_{exp}(2018)$ ($B_{exp}(2018)/K_{exp}$) A2+3	1976 (0.36)	1785 (0.41)	2787 (0.45)	1167 (0.39)	2080 (0.39)	1143 (0.42)
$B_{exp}(2019)$ ($B_{exp}(2019)/K_{exp}$) A1E	-	-	10767500 (0.84)	4325280 (0.87)	479 (0.56)	296 (0.66)
$B_{exp}(2019)$ ($B_{exp}(2019)/K_{exp}$) A1W	-	-	350 (0.22)	264 (0.39)	456 (0.26)	316 (0.45)
$B_{exp}(2019)$ ($B_{exp}(2019)/K_{exp}$) A2+3	-	-	2742 (0.44)	1199 (0.40)	2052 (0.38)	1143 (0.43)
λ^{A1E}	0.15	0.15	0.99	0.99	0.24	0.24
λ^{A1W}	0.26	0.26	0.00005	0.00001	0.21	0.26
λ^{A2+3}	0.59	0.59	0.00002	0.00001	0.54	0.49
q_{A1E}	0.00558	0.0000003	0.0000001	0.0000003	0.00252	0.00527
q_{A1W}	0.00299	0.00277	0.00235	0.00461	0.00191	0.00395
q_{A2+3}	0.000549	0.00074	0.00041	0.0012	0.00058	0.00121
	Om18	Om18s2	Om20old	Om20old01	Om20a	Om20a1

Table 2: Negative log likelihood values and other parameters obtained for different ω^* values (when estimating all three lambdas). The -lnL values in red indicate minima across the row concerned.

	ω^* =**17**	ω^* =**18**	ω^* =**19**	ω^* =**21**	ω^* =**23**
-lnL T	-570.90	-566.26	-564.80	-562.73	-562.16
q^{pen}	2.36	1.51	1.35	0.58	0.10
-lnL T-q^{pen}	-573.25	-567.77	-566.15	-563.32	-562.26
-lnl CPUE	-110.98	-113.11	-115.42	-118.64	-119.36
-lnl CPUE A1E	-22.40	-23.38	-23.58	-23.62	-23.61
-lnl CPUE A1W	-50.80	-50.94	-50.89	-50.70	-50.60
-lnl CPUE A2+3	-37.77	-38.79	-40.95	-44.28	-43.14
-ln SCI CAL	-517.90	-513.37	-509.94	-505.67	-504.55
-ln CAL A1E	26.71	27.22	27.44	27.31	27.15
-ln CAL A1W	-190.30	-190.47	-190.27	-189.38	-188.97
-ln CAL 2+3W	-354.31	-350.12	-347.11	-343.60	-342.73
Bsp(2019)/K	0.43	0.43	0.40	0.37	0.37
λ^{A1E}	**1**	0.27	0.25	0.24	0.24
λ^{A1W}	**0**	0.15	0.18	0.21	0.21
λ^{A2+3}	**0**	0.58	0.58	0.55	0.54

Figure 1: Total -lnL contributions, and the –lnL contributions for the CPUE and CAL data are provided below. The plots shows the "DIFF" scores where DIFF = -lnL - -lnL$_{lowest}$.

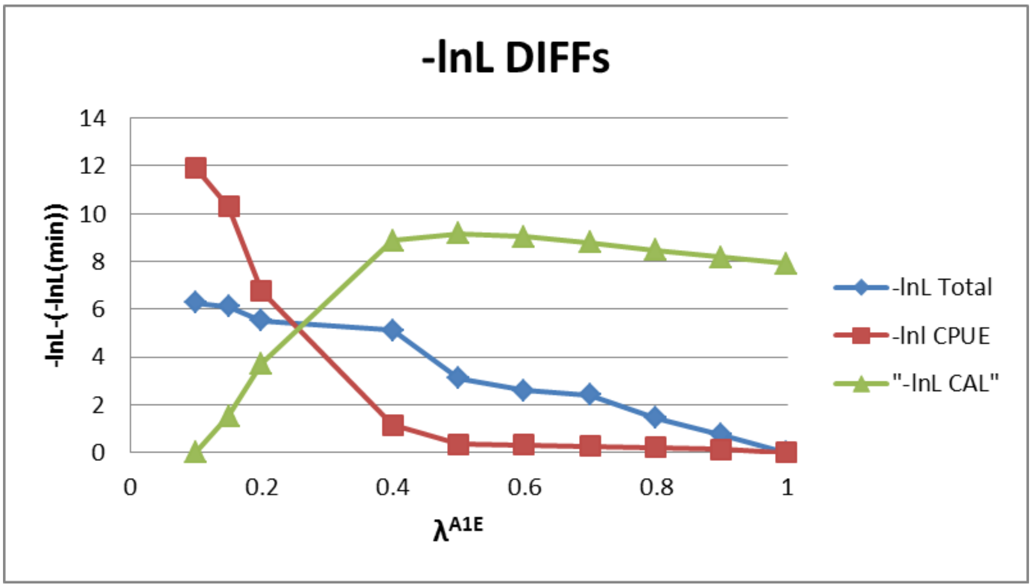


Figure 2: Total -lnL contributions, and the –lnL contributions for each sub-area. The top plot shows values for the CPUE data, and the bottom plot shows values for the CAL data. The plots shows the "DIFF" scores where

DIFF = -lnL - -lnL$_{lowest}$.

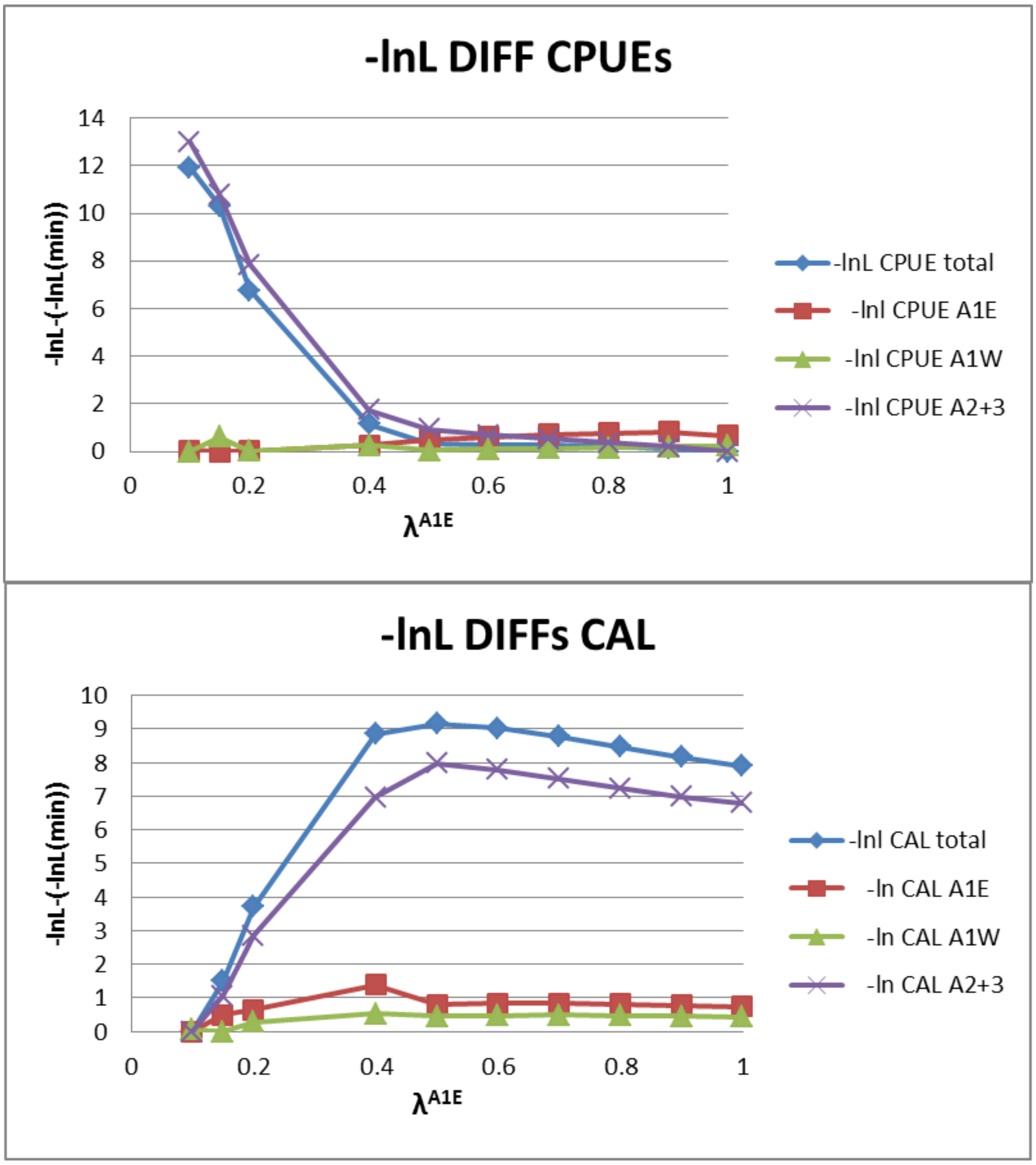


Figure 3: Exploitable biomass trends across a range of fixed λ^{A1E} values. For A1E two plots are shown – the RHS omits the λ_{1E} = 0.9 option for clarity.

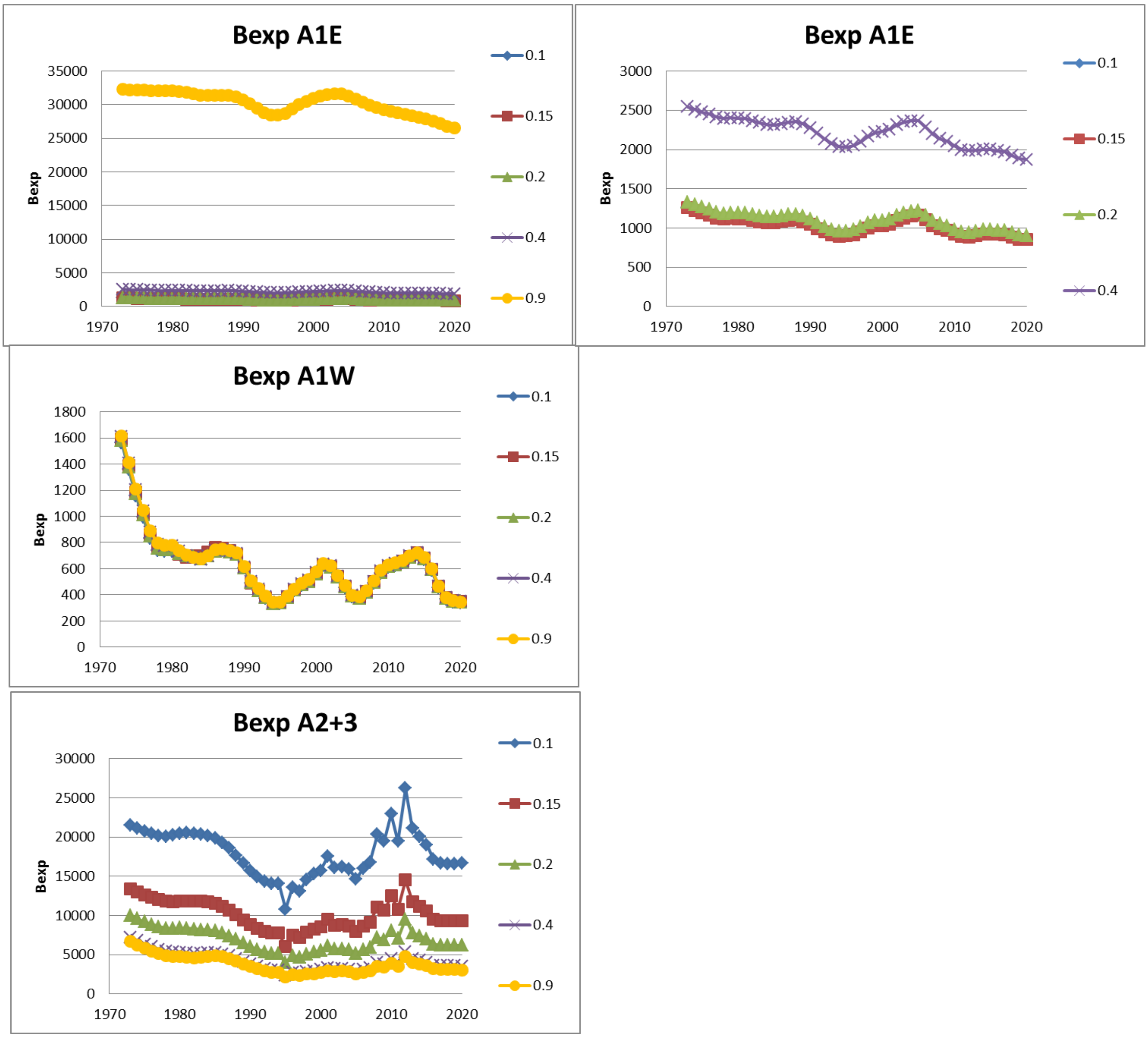


Figure 4: The catchability values *q* across a range of fixed λ^{A1E} values.

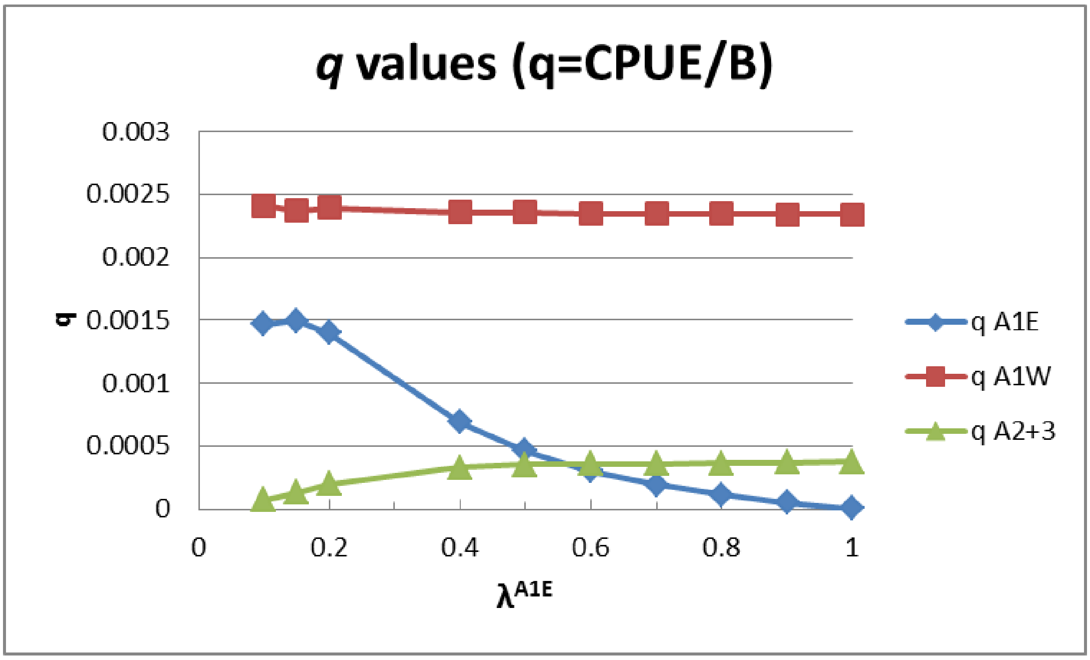


Figure 5: The modified catchability values *q across a range of fixed λ_{1E} values. Results are shown for "no q* penalty", "q* penalty w*= 23" and an intermediate scenario of "q* penalty w*= 19".**

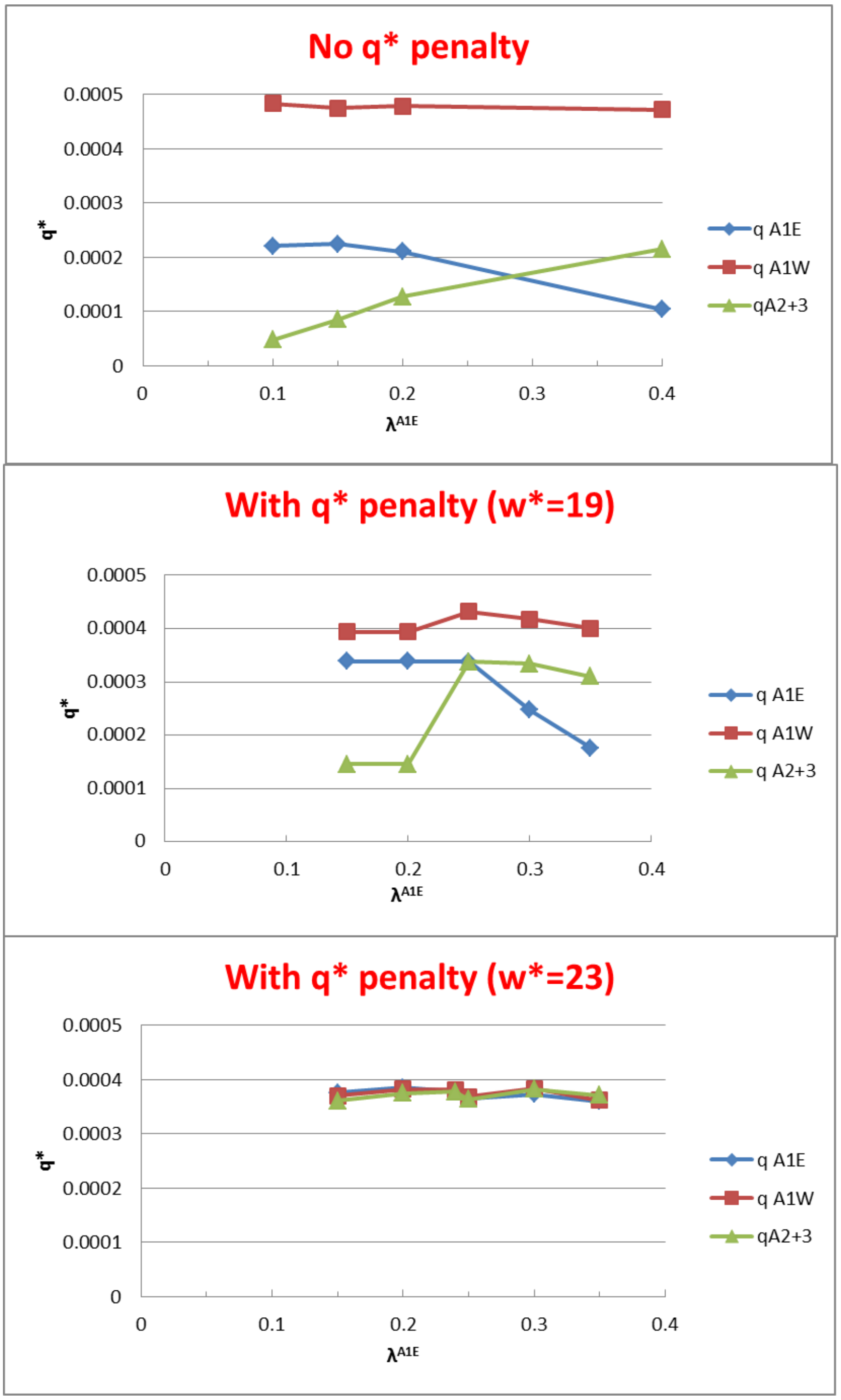


Figure 6: For a model that estimates all three λ values, the plot below shows –lnL values for a range of ω* weights. Here the "–lnL total – q pen" values are plotted where the DIFF score is reported, where DIFF is the difference between "-lnL total – q pen" and that for the largest ω*=23.

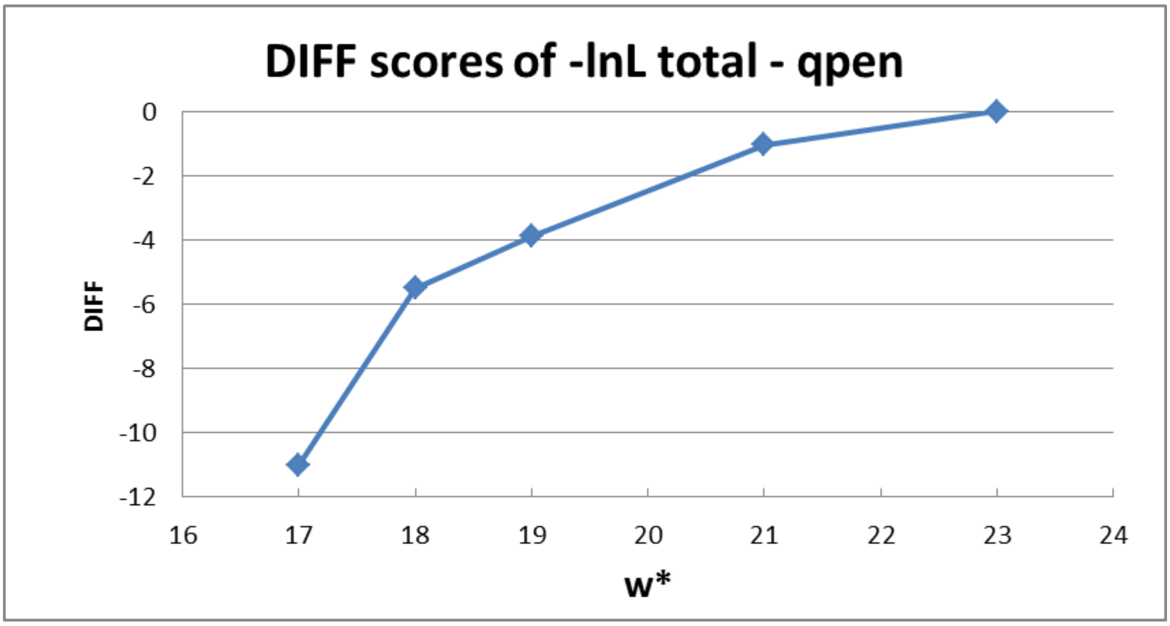


Figure 7: For a model that estimates all three λ values, the plot below shows the exploitable biomass trends for each sub-area for a range of ω* weights. The LHS is for ω* weights from 17 to 23, whereas the RHS omits the ω* =17 scenario for clearer comparison of the other results.

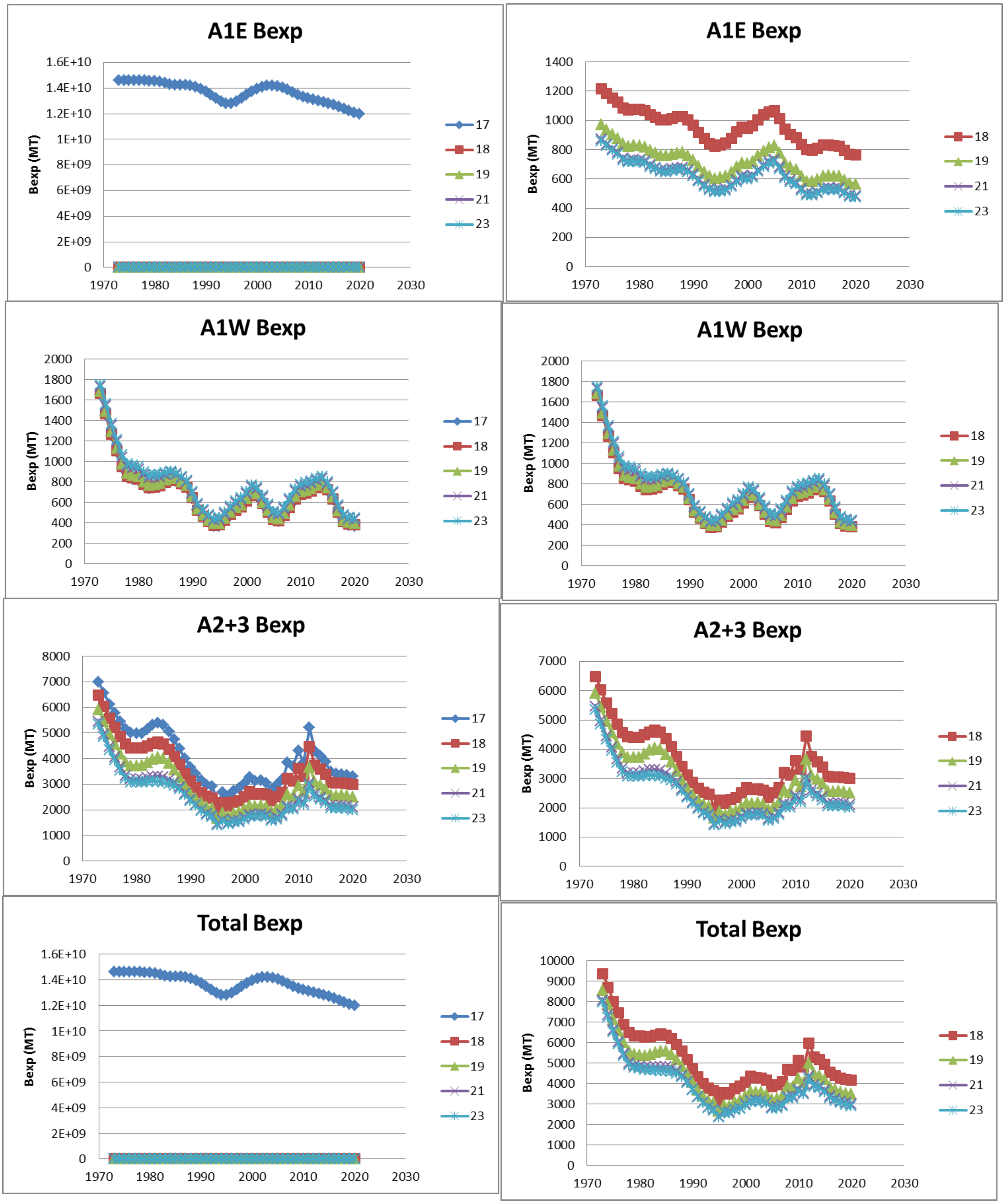


Figure 8: For a model that estimates all three λ values, the plot below shows the spawning biomass trends (for the whole resource) for a range of ω* weights. The top panels show the absolute biomass values, with the bottom panels showing Bsp/K. The LHS is for ω* weights from 17 to 21, whereas the RHS omits the ω* =17 scenario for clearer comparison of other results.

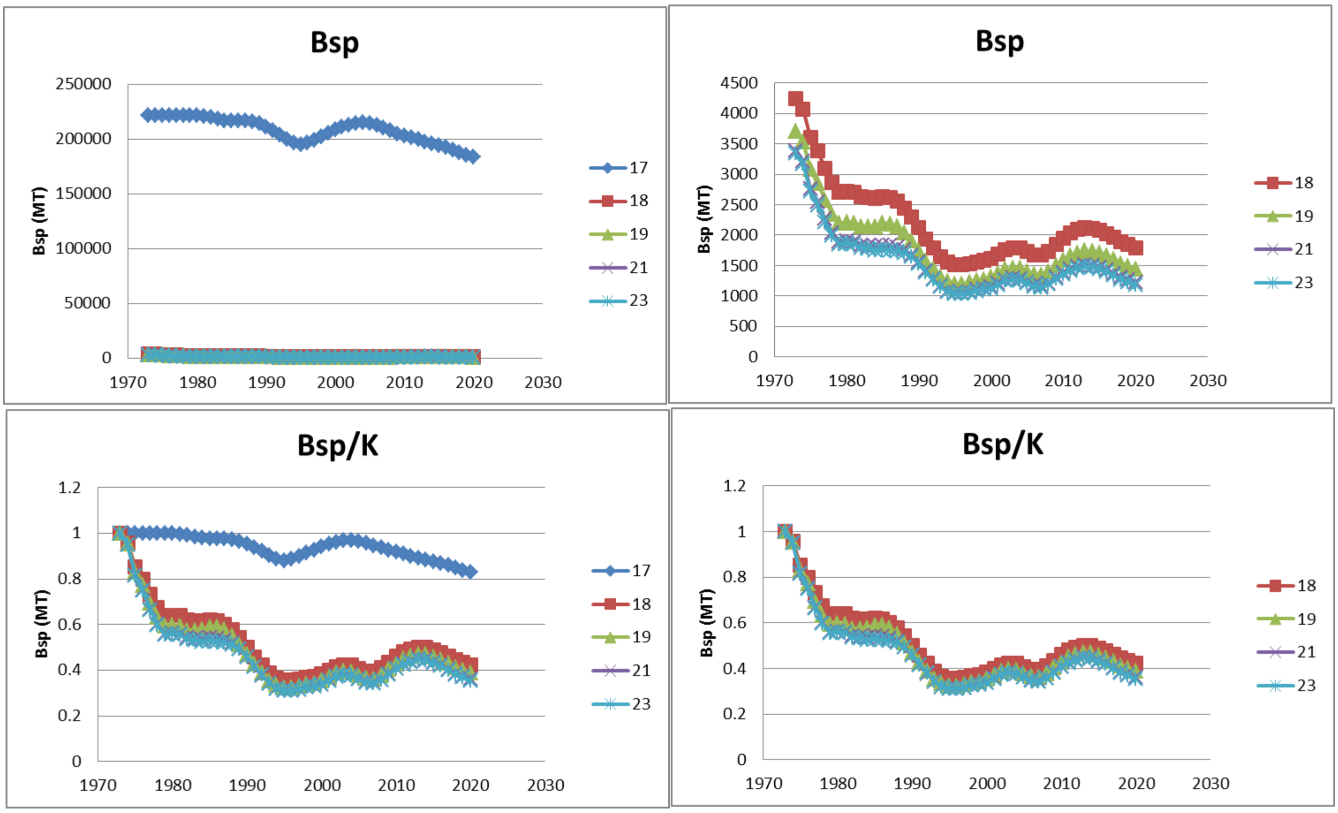
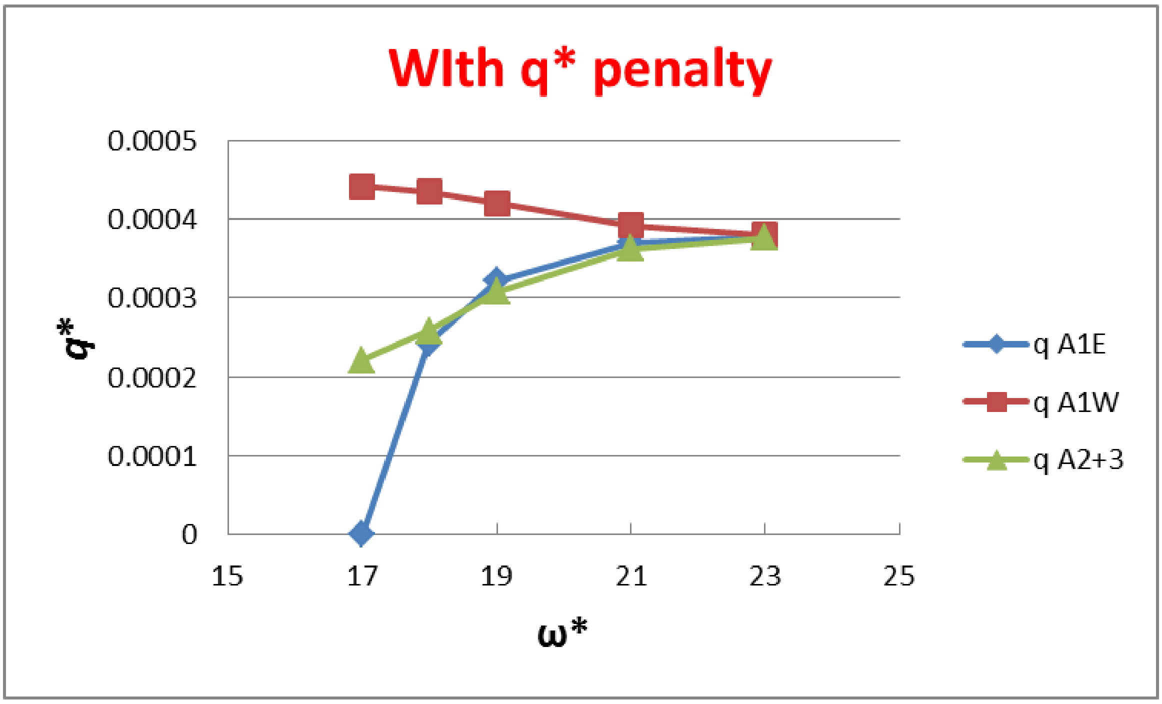


Figure 9: For a model that estimates all three λ values, the plot below shows the modified q* values over a range of ω* weights.



Appendix: Analysis of relative areal coverage of SCRL fishing

For each block in each of the three sub-areas (A1E, A1W and A2+3), one can calculate the % which that block has contributed to the overall catch in that sub-area over all years from 1977 to 2018. Figure A1 plots these %'s from smallest to highest.

Table A1 reports the number of blocks in each sub-area from which either 70%, 80% or 90% of the total catch for that sub-area has been taken. These can then be converted into relative proportions of the total "area". These proportions are quite similar across the 70% range-90%.

Figure A2 is a map of the SCRL fishing grounds showing the three sub-areas A1E, A1W and A2+3. The blue rectangles indicate the grid blocks where catches have been recorded over the 1977-2018 period. The red ellipses indicate the grid blocks from which 70% of the catches have been taken in each sub-area.

Conclusion: The total effective areal coverages of the three sub-areas A1E, A1W and A2+3 are reasonably taken to be 0.15: 0.20: 0.65 respectively when considered in relative terms.

Table A1:

	# of blocks			Proportions by sub-area		
	A1E	A1W	A2+3	A1E	A1W	A2+3
70% of catch	8	11	36	0.145	0.200	0.655
80% of catch	11	15	50	0.145	0.197	0.658
90% of catch	15	21	73	0.138	0.193	0.670

Figure A1: % of the sub-area total catch made in each grid block over 1977-2018. Grid block numbers do not correspond to the numbers shown in Figure A2.

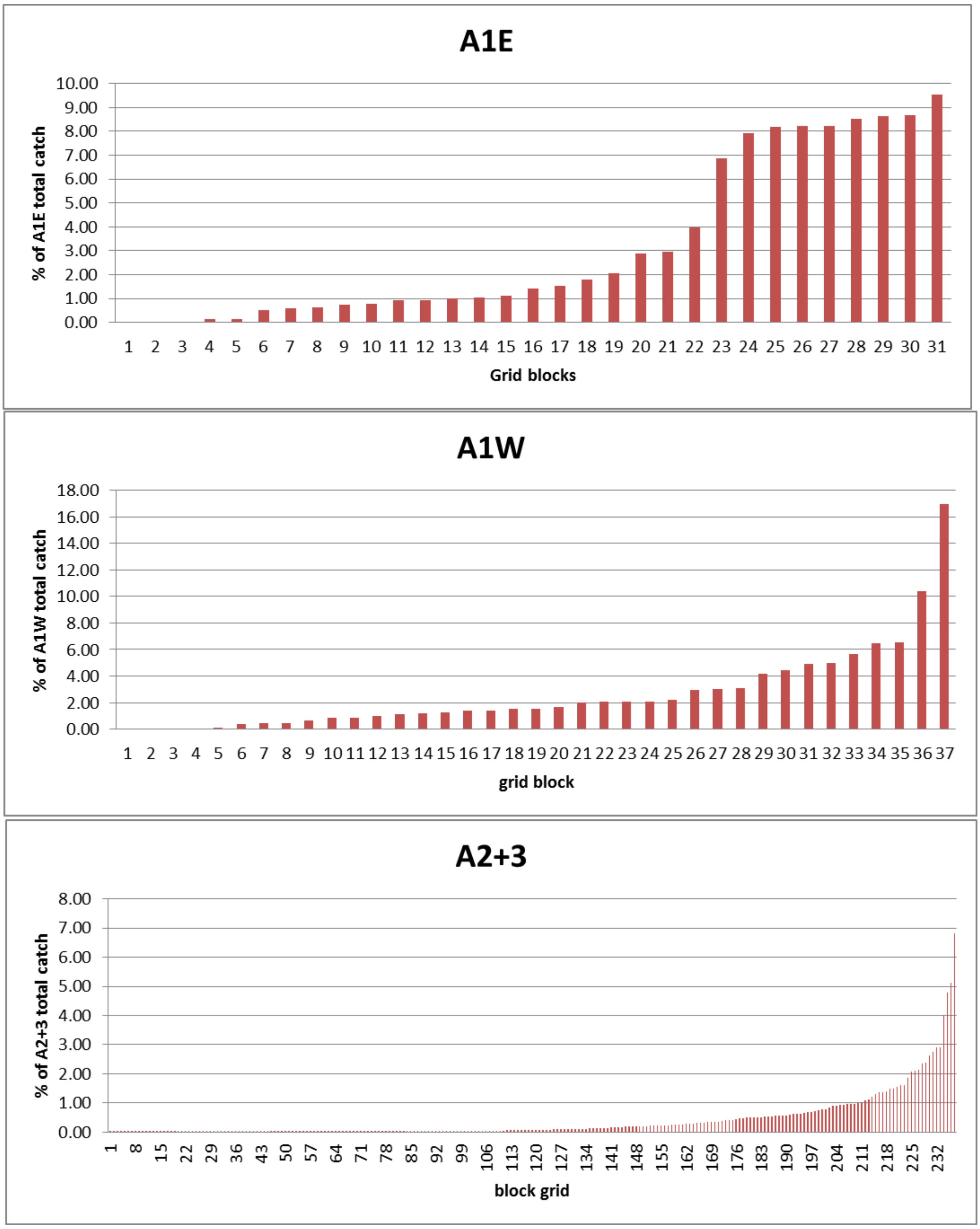


Figure A2: A map of the SCRL fishing grounds showing the three sub-areas A1E, A1W and A2+3. The blue rectangles indicate the grid blocks where catches have been recorded over the 1977-2018 period. The red ellipses indicate the grid blocks within which 70% of the catches in each sub-area has been taken.