Compliance poaching trends updated up to 2020 for West Coast rock lobster from modelling the "old" and the "new" databases simultaneously

A. Brandão and D.S. Butterworth

Marine Resource Assessment & Management Group (MARAM)
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch 7701, Cape Town

June 2021

Abstract

Updated results are given for an approach which simultaneously models the "old" and the "new" databases as recommended by the Panel for the 2018 International Fisheries Stock Assessment Workshop, with poaching data for 2020 now included. These results are compared to those obtained previously. For case where the "old" database is upweighted (as has been the accepted recent practice), the 2020 poaching index estimated for the southern Super-area 8+ shows a statistically significant increase compared to 2019 and is also the highest estimated value over the whole time series. In the case of the northern Super-areas (3-7) a statistically significant decrease is evident. A three-point smoothing approach adopted previously is used for converting the poaching series obtained into the outputs to be used for the input series to the assessments.

Keywords: west coast rock lobster; poaching trends; confiscations; policing effort

Introduction

The analyses for estimating poaching trends recommended by the Panel for the 2018 International Fisheries Stock Assessment Workshop (Cox *et al*. 2018), and implemented by Brandão and Butterworth (2019a), are updated to include data for 2020. To aid comparison, the results for this approach reported in Brandão and Butterworth (2021), which included data to 2019, are shown as well.

Data

Monthly data on confiscations and policing effort obtained from one of the DFFE Directorates within the CD (Directorate: Compliance) for the period of April 2008 to December 2020[1] form the "old" database. Data for the period 2012 to 2020 on rock lobster confiscations that are linked to a policing

[1] Note that a "year" refers to a calendar year throughout this document.

effort type form the "new" database. The first three months of the 2016 compliance data have been omitted from the analyses to remove the effect of the greatly enhanced policing levels during those months when Operation Phakisa was launched.

Methods

The recommendation by the International Panel to obtain poaching trends is to apply the models of Brandão and Butterworth (2018) to analyse the "new" database, together with the "old" database for which confiscations were not linked to the associated policing effort type, in a way that combines these two databases. This keeps some parameter values common between the two models to improve the precision of parameter estimates for the "new" database model of Brandão and Butterworth (2018). The Panel also recommended assuming that the number of confiscations follow a Negative Binomial distribution, instead of an overdispersed Poisson as assumed by Brandão and Butterworth (2018). In the case of the "new" database, only positive confiscations are reported. Thus, because instances of zero confiscations are never recorded in the "new" database, a Zero-Truncated Negative Binomial distribution is assumed.

The numbers of confiscations from the "new" and the "old" databases are modelled together assuming the following distributions:

$$C_{y,m,t}^{new} \sim \text{Zero-Truncated Negative Binomial}\left(exp\left(\lambda^{new} + \alpha_m + \beta_t + \delta_y\right), \xi^{new}\right)$$

$$C_{y,m}^{old} \sim \text{Negative Binomial}\left(\sum_t \left(Q_t e^{\beta_t} E_{y,m,t}^{old}\right) exp\left(\lambda^{old} + \alpha_m + \delta_y\right), \xi^{old}\right)$$

where

$C_{y,m,t}^{new}$ is the number of confiscations made in a single compliance event in year *y*, month *m* and by policing type *t* as reported in the "new" database,

$C_{y,m}^{old}$ is the total number of confiscations made in year *y* and month *m* as reported in the "old" database,

$E_{y,m,t}^{old}$ is the total policing effort reported in the "old" database for year *y*, month *m* and by policing type *t*,

Q_t is a factor to account for the absences of inspections with zero rock lobster confiscations in the "new" database; the adjustments made are the averages over years of proportions of successful (illegally caught rock lobster confiscated) inspections as given in Table 1,

λ^{new} is the intercept for the "new" database,

λ^{old} is the intercept for the "old" database,

α_m is a common month effect for both databases,

β_t reflects the type of policing effort which is linked to the confiscations, where the "type" factor is associated with the different types of policing such as coastal patrols, slipway inspections and vehicles inspections; they provide relative policing effort efficiencies which can be used in the "old" database to link policing effort to the number of confiscations,

δ_y is the common year effect for both databases (2008 to 2020 for Super Area 8+ and 2009 to 2020 for the northern Super-areas 3 to 7) whose estimates provide the poaching trend, and

$\xi^{new/old}$ is the dispersion parameter of the Negative Binomial distribution for the "new"/"old" databases.

Note that "year" refers to a calendar year throughout this document.

The contribution of the "old" database to the negative log-likelihood function in terms of individual observations is given by:

$$-\ln L_{old} = \sum_i \left\{ w_i \left[\left(C_i^{old} + (\phi^{old})^{-1}\right) ln\left(1 + \phi^{old}\mu_i^{old}\right) - C_i^{old}\, ln\left(\phi^{old}\mu_i^{old}\right) - \sum_{j=0}^{C_i^{old}-1} ln(j + (\phi^{old})^{-1}) \right] \right\} \quad (1)$$

where

C_i^{old} represents a single record of $C_{y,m}^{old}$ for a particular year (*y*) and month (*m*),

ϕ^{old} is the reciprocal of the dispersion parameter ξ^{old},

w_i is a weighting factor applied to upweight the contribution of the "old" database to the overall negative log-likelihood (see later discussion), and

μ_i^{old} is determined by a set of *k* indicator variables to represent the categorical variables λ^{old}, α_m and δ_y, and is given by:

$$\mu_i^{old} = \sum_t \left(Q_t e^{\beta_t} E_{i,t}^{old}\right) exp\left(\theta_1 X_{1,i} + \theta_2 X_{2,i} + \cdots + \theta_k X_{k,i}\right),$$

where $X_{1,i} = 1$ represents the intercept λ^{old}, and $E_{i,t}^{old}$ represents the single records of $E_{y,m,t}^{old}$ for a particular year (*y*) and month (*m*) for policing type *t*.

Similarly, the contribution of the "new" database to the negative log-likelihood function in terms of individual observations and assuming a Zero-Truncated Negative Binomial distribution is given by:

$$-\ln L^{new} = \sum_i \left\{ \begin{array}{l} \left(C_i^{new} + (\phi^{new})^{-1}\right) ln(1 + \phi^{new}\mu_i^{new}) - C_i^{new}\, ln(\phi^{new}\mu_i^{new}) - \\ \sum_{j=0}^{C_i^{new}-1} ln(j + (\phi^{new})^{-1}) - ln\left(1 - (1 + \phi^{new}\mu_i^{new})^{-(\phi^{new})^{-1}}\right) \end{array} \right\} \quad (2)$$

where

C_i^{new} represents the single records of $C_{y,m,t}^{new}$ for a particular year (*y*), month (*m*) and policing type (*t*),

ϕ^{new} is the reciprocal of the dispersion parameter ξ^{new},

μ_i^{new} is determined by a set of *s* indicator variables to represent the categorical variables $\lambda^{new}, \alpha_m, \delta_y$ and β_t, and is given by:

$\mu_i^{new} = exp\left(\zeta_1 Y_{1,i} + \zeta_2 Y_{2,i} + \cdots + \zeta_s Y_{s,i}\right)$, where $Y_{1,i} = 1$ represents the intercept λ^{new}. Note that the regression coefficients θ and ζ (of μ_i^{old} and μ_i^{new} respectively) that correspond to

the categorical variables for month and year will be the same as in equation (1) as they are common between both model components.

In the equations above of the contributions to the negative log-likelihood function, the following relationship for the gamma function is used:

$$ln\left(\frac{\Gamma(C_i + \phi^{-1})}{\Gamma(\phi^{-1})}\right) = \sum_{j=0}^{C_i-1} ln(j + \phi^{-1}).$$

The w_i weighting factor in equation (1) may be set at a value of more than 1 to upweight the contributions of data in the "old" database to the overall negative log-likelihood, compared to those in the "new" database in equation (2). The reason that this factor is introduced is that the "old" and the "new" databases are not comparable in the sense that the confiscation entries in the "old" database have been summed for each month in a particular year, while those in the "new" database represent individual incidents of non-zero confiscations that occurred in a particular month and year. Thus, by upweighting the contribution of data in the "old" database, one is compensating for the input-related information that has been lost by the summations of the confiscations in a month; those entries pertain to multiple rather than single incidents.

The weighting factors were determined by examining the maximum number of positive confiscations that took place in each month over the years of the "new" database (unfortunately raw data with this information are not available). This examination showed clearly that there were months of typically higher and of typically lower numbers of positive confiscations. The values for w_i based on this exercise, and applied in this paper for the months December to May were 20 for the southern and 15 for the northern areas, while for the months of June to November, a weight of 10 was applied to both the southern and the northern areas. These choices were made based on the values listed in Table 1 of Brandão and Butterworth (2019a).

Results

Tables 2 and 3 show parameter estimates for Super-areas 3+4+5+6+7 and 8+ respectively for GLMs fitted as follows:

- to the combined "old" and "new" databases with data from the "old" database weighted by some factor (see text immediately above for details); and
- to the combined "old" and "new" databases with a weight of 1 (i.e. unweighted) applied to data from the "old" database.

Figure 1 shows the poaching trends obtained from the two different analysis approaches, as detailed above for the results shown in Tables 2 and 3, for the two Super-area combinations.

For the preferred approach of upweighting the "old" database contributions to the likelihood, the poaching indices for 2020 show an appreciable increase in relative terms compared to those for 2019 for the southern Super-area 8+, and a decrease for the northern Super-areas (3-7). Figure 2 shows the nominal annual number of confiscations and the annual number of policing incidences in the southern Super-area 8+ in both databases. This demonstrates the much larger number of confiscations made in

2020 compared to those made in earlier years while there was no substantial change in policing effort (Figure 2).

Figure 3 and Tables 4a-b report these poaching trends relative to 2008 (for Super-area 8+) or 2009 (for the northern areas). Figure 3 also compares these trends to the smoothing method proposed by Brandão and Butterworth (2019b). This is simply three-point smoothing, where each value is replaced by the average of itself with the values for the years immediately before and after. For the points at the end of the series, this simplifies further to averaging their values with those for the single adjacent year.

The bottom plot of Figure 3 shows the poaching trends above for Super-area 8+ when the largest confiscation records from 2020 are omitted from the analysis (one record from the "old" database close to 18 000 and one from the "new" database of over 13 000 lobsters). This is intended as a check of sensitivity to a possible "outlier" event. It should be noted, however, that while the confiscation of over 13 000 lobsters in the "new" database is more than twice the value of the previous highest record, in the "old" database large confiscations of lobsters have been recorded previously (close to 15 000 in 2013 and close to 12 000 in 2017).

Table 5 shows the ratio of 2020 estimated poaching indices to the 2019 value together with the lower and upper bounds of the 95% confidence intervals. For the southern Super-area 8+, a statistically significant increase in the poaching index for 2020 compared to 2019 is evident; this is also the highest estimated index over the whole time series. For the northern Super-areas 3-7 a statistically significant decrease in the poaching index for 2020 compared to 2019 is evident.

References

Brandão, A. and Butterworth, D.S. 2018. Refined trends in poaching for West Coast rock lobster using information from the "new" database for the period 2012 to 2017. Fisheries/2018/AUG/SWG-WCRL18.

Brandão, A. and Butterworth, D.S. 2019a. Trends in poaching for West Coast rock lobster from modelling the "old" and the "new" database simultaneously. Fisheries/2019/APR/SWG-WCRL02.

Brandão, A. and Butterworth, D.S. 2019b. Trends in poaching for West Coast rock lobster from modelling the "old" and the "new" database simultaneously updated to 2019. Fisheries/2019/JUL/SWG-WCRL14.

Brandão, A. and Butterworth, D.S. 2021. Updated trends in poaching for West Coast rock lobster from modelling the "old" and the "new" database simultaneously. Fisheries/2021/JUNE/SWG-WCRL02.

Cox, S., Gaichas, S., Haddon, M. and Punt, A.E. 2018. International Review Panel Report for the 2018 International Fisheries Stock Assessment Workshop, 26-30 November 2018, University of Cape Town.

Table 1. Difference between the number of observations in the "old" and the "new" databases that might be assumed to be the number of "zero" confiscations in the "new" database. Results are shown by common policing effort types and Super-areas. Percentages of observations in the "new" database that would constitute non-zero confiscations are also shown.

	Difference in number of observations between "old" and "new"			**Percentage of observations that would constitute non-zero confiscations in the "new" database**		
Northern Super-areas 3+4+5+6+7						
	Coastal Patrols	Slipway inspections	Vehicle inspections	Coastal Patrols	Slipway inspections	Vehicle inspections
2008	—	—	—	—	—	—
2009	1822	1939	3724	—	—	—
2010	2510	3287	3587	—	—	—
2011	3223	4197	3953	—	—	—
2012	2935	3537	7068	0.17%	0.06%	0.24%
2013	2513	2902	6464	0.28%	0.68%	0.68%
2014	2646	2722	4174	0.94%	1.45%	0.69%
2015	2742	2994	3551	0.47%	1.22%	0.84%
2016	2623	3134	2442	0.57%	0.22%	0.12%
2017	2545	2781	1841	0.90%	0.36%	0.22%
2018	2122	2315	1108	0.47%	0.30%	0.09%
2019	1876	1982	1373	0.90%	0.65%	0.65%
2020	1770	2221	5796	0.28%	0.31%	0.00%
Average	***2444***	***2834***	***3757***	***0.55%***	***0.58%***	***0.39%***
Southern Super-area 8+						
2008	3923	2585	3179	—	—	—
2009	4857	3766	3951	—	—	—
2010	5540	3507	3847	—	—	—
2011	7722	4876	2604	—	—	—
2012	5991	4547	3366	0.02%	0.57%	0.03%
2013	4931	4291	2966	0.78%	0.44%	0.00%
2014	4208	3848	1806	0.82%	0.67%	0.06%
2015	5656	4621	2593	0.88%	0.75%	0.04%
2016	5579	3661	3539	0.46%	0.97%	0.06%
2017	6161	4534	4567	0.63%	0.74%	0.00%
2018	4858	3258	3838	0.39%	0.82%	0.00%
2019	4503	3323	3501	0.27%	0.39%	0.00%
2020	3300	2517	3229	0.54%	0.79%	0.03%
Average	***5171***	***3795***	***3307***	***0.51%***	***0.65%***	***0.02%***

Table 2. GLM parameter/coefficient (and standard error) estimates[2] for **Super-areas 3+4+5+6+7.**

	Previous weight = 1	Previous weight = 15; 10	Updated weight = 1	Updated weight = 15; 10
January	0	0	0	0
February	1.380 (0.449)	0.668 (0.143)	1.541 (0.398)	0.755 (0.136)
March	1.281 (0.431)	0.424 (0.153)	1.293 (0.407)	0.409 (0.145)
April	0.484 (0.401)	0.108 (0.148)	0.649 (0.396)	0.244 (0.140)
May	0.039 (0.502)	-0.848 (0.148)	0.098 (0.440)	-0.796 (0.142)
June	-1.533 (0.449)	-2.986 (0.165)	-1.439 (0.446)	-2.834 (0.154)
July	-2.092 (0.478)	-3.048 (0.165)	-2.051 (0.441)	-3.057 (0.157)
August	-1.449 (0.473)	-2.444 (0.159)	-1.413 (0.445)	-2.435 (0.152)
September	-0.687 (0.537)	-2.070 (0.174)	-0.583 (0.493)	-1.976 (0.165)
October	-0.209 (0.471)	-1.954 (0.155)	-0.085 (0.451)	-1.757 (0.149)
November	-0.898 (0.534)	-1.492 (0.158)	-0.477 (0.402)	-1.026 (0.153)
December	0.630 (0.667)	-0.127 (0.157)	0.879 (0.588)	0.136 (0.149)
2008	—	—	—	—
2009	0.613 (0.603)	0.479 (0.161)	0.759 (0.535)	0.609 (0.161)
2010	1.504 (0.588)	1.216 (0.150)	1.691 (0.494)	1.399 (0.149)
2011	0.682 (0.589)	0.355 (0.159)	0.881 (0.515)	0.541 (0.157)
2012	-0.204 (0.512)	-0.085 (0.154)	-0.072 (0.452)	0.056 (0.154)
2013	-0.417 (0.469)	-0.755 (0.151)	-0.270 (0.481)	-0.579 (0.147)
2014	0	0	0	0
2015	0.525 (0.575)	0.104 (0.142)	0.663 (0.414)	0.214 (0.143)
2016	-0.815 (0.638)	-0.799 (0.164)	-0.676 (0.482)	-0.688 (0.164)
2017	0.091 (0.648)	-0.884 (0.146)	0.314 (0.447)	-0.776 (0.144)
2018	-0.681 (0.544)	-1.277 (0.157)	-0.445 (0.466)	-1.120 (0.152)
2019	-0.048 (0.541)	-0.560 (0.142)	0.114 (0.425)	-0.474 (0.141)
2020	—	—	-0.524 (0.428)	-0.794 (0.140)
coastal	0	0	0	0
slipway	0.409 (0.433)	-0.189 (0.319)	0.498 (0.394)	-0.155 (0.332)
vehicles	0.747 (0.376)	0.435 (0.217)	0.874 (0.428)	0.439 (0.228)

[2] The reference level for the GLM is set as January for the month factor, 2014 for the year factor and coastal for the policing type factor, thus these values are given as zero in the Table as their effect is incorporated in the intercept.

Table 3. GLM parameter/coefficient (and standard error) estimates[3] for **Super-area 8+.**

	Previous weight = 1	Previous weight = 20; 10	Updated weight = 1	Updated weight = 20; 10
January	0	0	0	0
February	1.804 (0.450)	1.355 (0.134)	1.846 (0.447)	1.350 (0.136)
March	0.526 (0.498)	-0.561 (0.140)	0.878 (0.499)	-0.267 (0.144)
April	0.992 (0.444)	1.289 (0.143)	1.019 (0.447)	1.217 (0.145)
May	0.246 (0.427)	0.335 (0.139)	0.297 (0.415)	0.320 (0.141)
June	0.458 (0.465)	0.547 (0.160)	0.499 (0.451)	0.763 (0.163)
July	-0.148 (0.503)	-0.298 (0.160)	0.130 (0.505)	0.371 (0.169)
August	-1.959 (0.471)	-2.147 (0.159)	-1.855 (0.465)	-2.001 (0.159)
September	1.352 (0.522)	0.473 (0.164)	1.768 (0.518)	0.711 (0.167)
October	-0.758 (0.480)	-0.399 (0.164)	-0.254 (0.478)	0.253 (0.166)
November	-0.679 (0.418)	-0.556 (0.158)	-0.789 (0.422)	-0.483 (0.160)
December	-1.729 (0.343)	-0.813 (0.136)	-1.651 (0.385)	-0.719 (0.135)
2008	-1.285 (0.611)	-1.264 (0.170)	-1.188 (0.636)	-1.067 (0.174)
2009	-1.425 (0.541)	-1.081 (0.157)	-1.283 (0.561)	-0.871 (0.157)
2010	-0.979 (0.540)	-0.890 (0.153)	-0.841 (0.562)	-0.645 (0.157)
2011	-0.318 (0.558)	-0.337 (0.144)	-0.380 (0.569)	-0.309 (0.153)
2012	-0.628 (0.459)	-0.989 (0.141)	-0.594 (0.536)	-1.163 (0.150)
2013	0.870 (0.412)	0.507 (0.147)	1.039 (0.418)	0.715 (0.151)
2014	0	0	0	0
2015	0.139 (0.392)	-0.281 (0.145)	0.120 (0.389)	-0.218 (0.149)
2016	-0.499 (0.443)	0.054 (0.165)	-0.505 (0.453)	0.144 (0.173)
2017	-1.132 (0.375)	-0.132 (0.146)	-1.070 (0.387)	0.082 (0.147)
2018	-1.157 (0.399)	-0.864 (0.161)	-1.121 (0.474)	-0.581 (0.163)
2019	-0.504 (0.460)	-0.236 (0.153)	-0.443 (0.475)	-0.125 (0.154)
2020	—	—	1.739 (0.493)	1.395 (0.158)
coastal	0	0	0	0
slipway	0.349 (0.281)	0.379 (0.260)	0.316 (0.282)	0.301 (0.266)
vehicles	4.481 (1.076)	3.576 (0.347)	3.871 (1.197)	1.832 (0.616)

[3] The reference level for the GLM is set as January for the month factor, 2014 for the year factor and coastal for the policing type factor, thus these values are given as zero in the Table as their effect is incorporated in the intercept.

Table 4a. Poaching series obtained from a) the unweighted (i.e. weight = 1) combined "old" and "new" databases, and b) the weighted (see text for details) combined "old" and "new" databases for the northern **Super-areas 3+4+5+6+7**. The results shown are normalised to a 2009 value of 1. Previous results for these two approaches are given as well.

	Previous weight = 1	**Previous weight = 15; 10**	**Updated weight = 1**	**Updated weight = 15; 10**
2008	—	—	—	—
2009	1.000	1.000	1.000	1.000
2010	2.437	2.090	2.541	2.203
2011	1.072	0.883	1.131	0.934
2012	0.442	0.569	0.436	0.575
2013	0.357	0.291	0.357	0.305
2014	0.542	0.619	0.468	0.544
2015	0.916	0.687	0.909	0.674
2016	0.240	0.279	0.238	0.273
2017	0.593	0.256	0.641	0.250
2018	0.274	0.173	0.300	0.177
2019	0.516	0.354	0.525	0.339
2020	—	—	0.277	0.246

Table 4b. Poaching series obtained from a) the unweighted (i.e. weight = 1) combined "old" and "new" databases, and b) the weighted (see text for details) combined "old" and "new" databases for the southern **Super-area 8+**. The results shown are normalised to a 2008 value of 1. Previous results for these two approaches are given as well.

	Previous weight = 1	**Previous weight = 20; 10**	**Updated weight = 1**	**Updated weight = 20; 10**
2008	1.000	1.000	1.000	1.000
2009	0.870	1.200	0.910	1.216
2010	1.358	1.454	1.415	1.526
2011	2.631	2.527	2.244	2.135
2012	1.930	1.317	1.811	0.909
2013	8.632	5.874	9.276	5.946
2014	3.616	3.540	3.281	2.908
2015	4.154	2.673	3.699	2.338
2016	2.196	3.736	1.980	3.359
2017	1.166	3.100	1.126	3.155
2018	1.137	1.492	1.069	1.626
2019	2.184	2.796	2.106	2.566
2020	—	—	18.679	11.732

Table 5. The ratio of 2020 poaching indices to the 2019 values, together with the lower and upper bounds provided by 95% confidence intervals, for the northern Super-areas 3-7 and the southern Super-area 8+.

Super-area	Ratio	Lower bound	Upper bound
Northern 3-7	0.726	0.547	0.964
Southern 8+	4.573	3.346	6.250

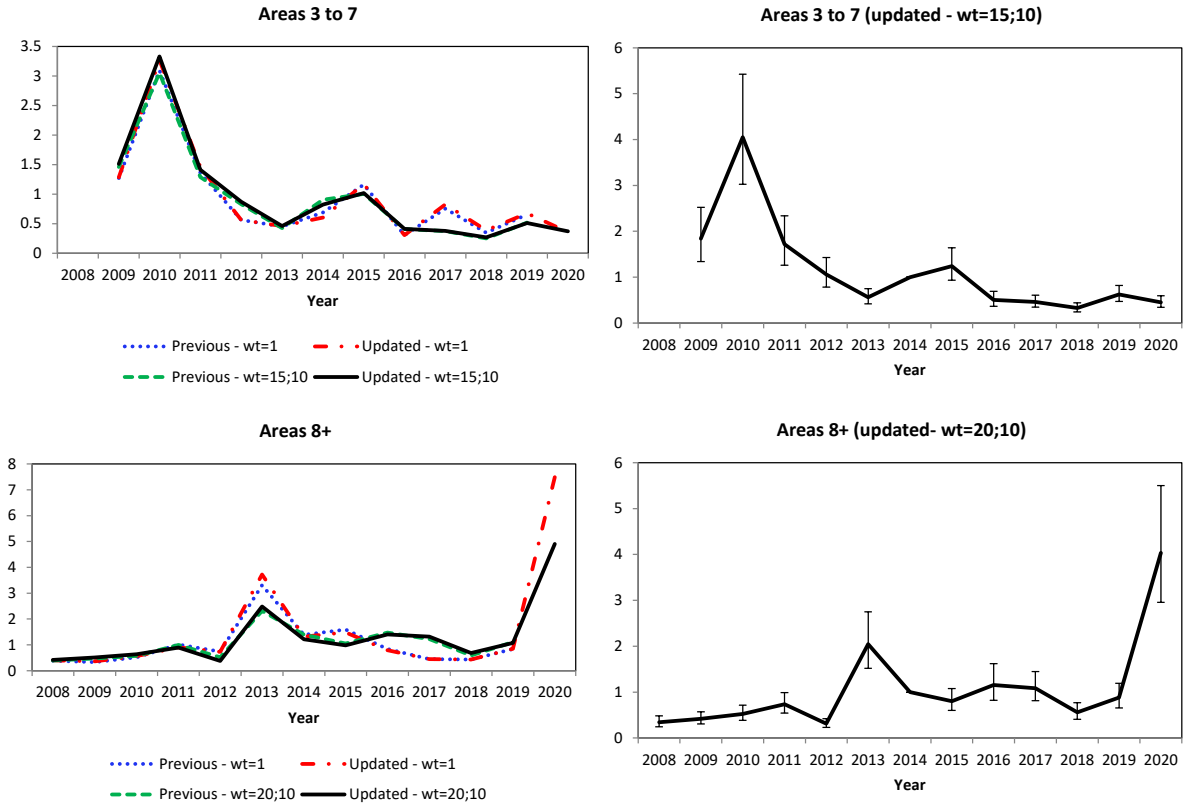

Figure 1. The plots on the left hand side show poaching trends (corresponding to year effects) for two different analysis approaches, with the previous results (to 2019 - Brandão and Butterworth, 2021) also plotted for comparison:

- modelling of the combined "old" and "new" databases with the "old" database data weighted by some factor (see text for details) – the approach now recommended; and
- modelling of the combined "old" and "new" databases with a weight of one (i.e. unweighted) applied to the "old" database data.

The plots on the right hand side show poaching trends for the weighted combined "old" and "new" databases (the recommended approach), together with 95% confidence limits. The plots described above are given for the northern **Super-areas 3+4+5+6+7 (top)** and the southern **Super-area 8+ (bottom)**. The series plotted on the left hand side have been normalised to the period from 2012 to 2019 for which they overlap.

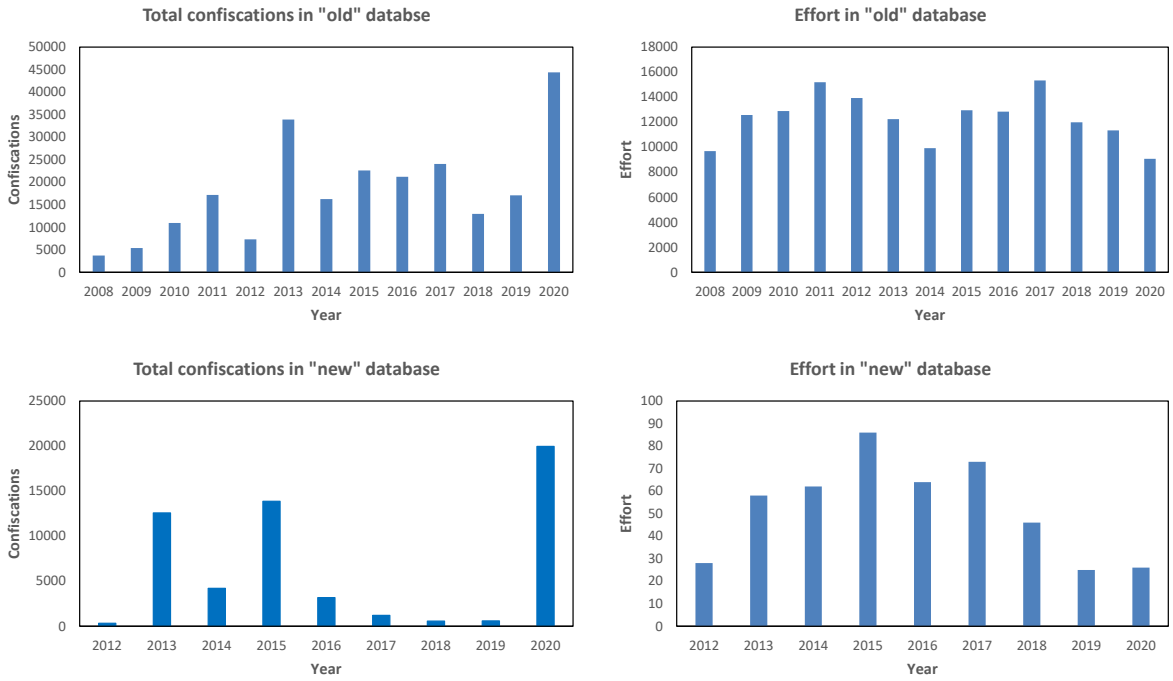

Figure 2. Nominal annual number of confiscations of rock lobster (left) and total number of policing incidences (right) in the “old” (top) and the “new” (bottom) databases for the southern **Super-area 8+**.

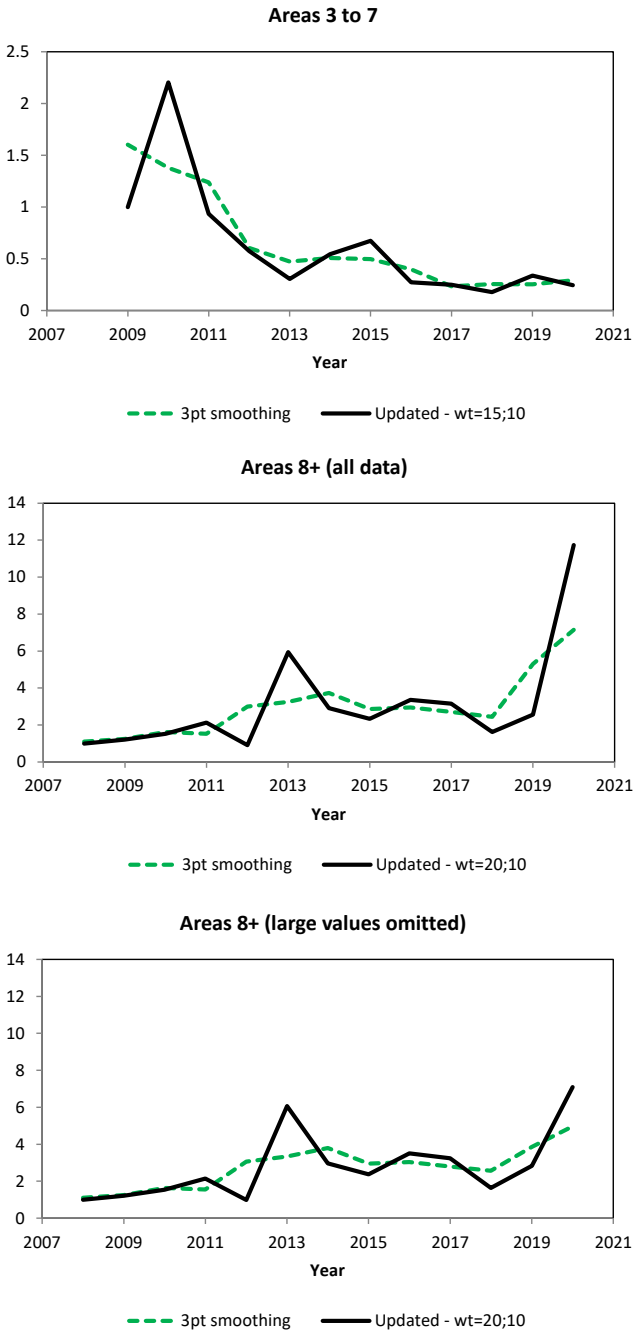

Figure 3. Poaching trends obtained using two different approaches:

- modelling of the combined "old" and "new" databases with the "old" database weighted by some factor – the approach now recommended (see text for details); and
- in addition, applying three-point smoothing to the poaching indices from the approach above.

The plots described above are given for **Super-areas 3+4+5+6+7 (top)** and **Super-area 8+ (middle and bottom)**. The bottom plot shows results when, as a sensitivity, the largest confiscation records from 2020 are omitted from the analysis (one record from the "old" dataset close to 18 000 and one from the "new" database of over 13 000 lobsters). Results shown are normalised to a 2008 value of 1 for Super-area 8+ and to a 2009 value of 1 for Super-areas 3+4+5+6+7, as assumed for the 2016 assessment and projections for the first approach, but not the second.