Trends in poaching for West Coast rock lobster from modelling the "old" and the "new" databases simultaneously updated to 2018

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

July 2019

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, in which poaching data for 2018 is now included. These results are compared to those obtained previously. For cases where the "old" database is upweighted the 2018 poaching indices estimated for both the northern Super-areas (3-7) and for the southern Super-area 8+ show statistically significant decreases compared to 2017. A new smoothing method is put forward for converting the output poaching series obtained to the input series used in the assessments.

Introduction

The analyses for estimating poaching trends recommended by the Panel of the 2018 International Fisheries Stock Assessment Workshop (Cox *et al*. 2018), and implemented by Brandão and Butterworth (2019), are updated to include data for 2018. To aid comparison, the results for this approach reported in Brandão and Butterworth (2019), which included data to 2017, are duplicated as well.

Data

Monthly data on confiscations and policing effort obtained from one of the Directorates within the CD (Directorate: Compliance) for the period of April 2008 to December 2018 form the "old" database. Data for the period 2012 to 2018 on rock lobster confiscations that are linked to a policing 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 (2018a) 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. 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 (2018a). 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 (2018a). 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 number of confiscations from the “new” and the “old” databases are modelled simultaneously assuming the following distributions:

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

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

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* reported in the “new” database,

$C_{y,m}^{old}$ is the total number of confiscations made in year *y* and month *m* 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 2 of Brandão and Butterworth (2018b),

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 2018 for Super Area 8+ and 2009 to 2018 for the northern Super-areas 3 to 7) whose estimates provide the poaching trend, and

$\varphi^{new/old}$ 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:

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

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 φ^{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 $X_{1,i} = 1$ to represent 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:

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

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 φ^{new},

μ_i^{new} is determined by a set of s indicator variables to represent the categorical variables λ^{new}, α_m, δ_y and β_t and is given by:

$\mu_i^{new} = \exp\left(\varsigma_1 Y_{1,i} + \varsigma_2 Y_{2,i} + \cdots + \varsigma_s Y_{s,i}\right)$, where $Y_{1,i} = 1$ to represent the intercept λ^{new}. Note that the regression coefficients θ and ς (of μ_i^{old} and μ_i^{new} respectively) that correspond to the categorical variables for month and year will be the same in this case 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:

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

The w_i weighting factor in equation (1) may be set at a value more than 1 to upweight the contribution 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 a 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 information that has been lost by the summing of the confiscations in a month; these 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 clearly showed 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 area and 15 for the northern areas, and 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 (2019).

Results

Tables 1 and 2 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 one (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 1 and 2, for the two Super-area combinations.

For the preferred approach of upweighting the "old" database contributions to the likelihood, the poaching indices for 2018 show appreciable declines in relative terms compared to those for 2017, for both the northern Super-areas (3-7) and for the southern Super-area 8+.

Figure 2 and Tables 3a-b compare these trends to how they were smoothed for use in the 2016 base case assessment. Results here are shown relative to 2008 (for Super-area 8+) or 2009 (for the northern areas).

Table 4 shows the ratio of 2018 estimated poaching indices to the 2017 value together with the lower and upper bounds of the 95% confidence intervals. For both the northern Super-areas 3-7 and the southern Super-area 8+, a statistically significant drop in the poaching indices for 2018 is evident.

The smoothing method used in 2016 has become dated, given three further years with estimates. To deal with such smoothing, which is necessary given the high variance (Figure 1) of some of the estimates, a new method is proposed which is simple and objective. 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. Figure 3 compares the results from this smoothing process with the updated estimates and the earlier 2016 approach.

References

Brandão, A. and Butterworth, D.S. 2018a. 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-WCRL/18.

Brandão, A. and Butterworth, D.S. 2018b. Investigation of the comparability of the West Coast rock lobster compliance databases ("new" and "old"). Fisheries/2018/AUG/SWG-WCRL/17.

Brandão, A. and Butterworth, D.S. 2019. Trends in poaching for West Coast rock lobster from modelling the "old" and the "new" database simultaneously. Fisheries/2019/APR/SWG-WCRL/02.

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

	Combined “old” and “new” – weight = 1	**Combined “old” and “new” – weight = 15; 10**	**Updated combined “old” and “new” – weight = 1**	**Updated combined “old” and “new” – weight = 15; 10**
January	0	0	0	0
February	1.443(0.440)	1.068(0.151)	1.181(0.430)	0.636(0.149)
March	1.464(0.468)	1.175(0.155)	1.388(0.447)	0.494(0.158)
April	0.672(0.421)	0.499(0.146)	0.320(0.567)	-0.186(0.150)
May	0.358(0.478)	-0.288(0.149)	-0.024(0.645)	-0.901(0.152)
June	-1.079(0.463)	-2.452(0.169)	-1.534(0.600)	-3.091(0.169)
July	-1.614(0.468)	-2.321(0.167)	-1.964(0.493)	-2.913(0.169)
August	-1.156(0.489)	-1.831(0.162)	-1.325(0.488)	-2.380(0.163)
September	-0.418(0.502)	-1.496(0.176)	-0.534(0.591)	-2.085(0.180)
October	0.107(0.515)	-1.687(0.167)	-0.238(0.537)	-2.001(0.161)
November	-0.552(0.426)	-0.849(0.159)	-0.807(0.606)	-1.522(0.164)
December	1.213(0.513)	0.164(0.160)	0.626(0.632)	-0.487(0.163)
2008	—	—	—	—
2009	0.521(0.512)	0.375(0.158)	0.607(0.614)	0.363(0.162)
2010	1.556(0.474)	1.154(0.143)	1.492(0.511)	1.097(0.149)
2011	0.710(0.498)	0.342(0.153)	0.632(0.508)	0.173(0.161)
2012	-0.194(0.455)	-0.206(0.152)	-0.218(0.657)	-0.211(0.155)
2013	-0.341(0.415)	-0.881(0.144)	-0.383(0.601)	-0.926(0.148)
2014	0	0	0	0
2015	0.540(0.412)	-0.036(0.138)	0.553(0.485)	-0.008(0.143)
2016	-0.900(0.480)	-0.891(0.166)	-0.817(0.680)	-0.938(0.166)
2017	-0.176(0.444)	-0.793(0.140)	0.184(0.483)	-0.957(0.145)
2018	—	—	-0.637(0.507)	-1.376(0.157)
coastal	0	0	0	0
slipway	1.074(0.890)	-0.756(0.464)	0.703(0.551)	0.139(0.365)
vehicles	1.374(0.758)	0.515(0.211)	1.051(0.749)	0.844(0.259)

Table 2. GLM parameter/coefficient (and standard error) estimates for **Super-area 8+.**

	Combined "old" and "new" – weight = 1	Combined "old" and "new" – weight = 20; 10	Updated combined "old" and "new" – weight = 1	Updated combined "old" and "new" – weight = 20; 10
January	0	0	0	0
February	1.811(0.527)	1.168(0.142)	1.778(0.476)	1.168(0.142)
March	0.446(0.589)	-0.588(0.148)	0.236(0.506)	-0.588(0.148)
April	0.360(0.475)	0.513(0.140)	0.939(0.458)	0.513(0.140)
May	0.014(0.479)	0.077(0.144)	0.079(0.431)	0.077(0.144)
June	0.193(0.523)	0.117(0.166)	0.213(0.481)	0.117(0.166)
July	-0.444(0.584)	-0.766(0.167)	-0.497(0.526)	-0.766(0.167)
August	-2.152(0.534)	-2.464(0.163)	-2.084(0.502)	-2.464(0.163)
September	1.322(0.563)	0.389(0.169)	1.255(0.532)	0.389(0.169)
October	-0.867(0.566)	-0.356(0.170)	-0.816(0.497)	-0.356(0.170)
November	-1.110(0.470)	-0.976(0.163)	-0.927(0.434)	-0.976(0.163)
December	-1.774(0.409)	-0.849(0.140)	-1.827(0.354)	-0.849(0.140)
2008	-1.139(0.602)	-1.087(0.160)	-1.175(0.601)	-1.211(0.163)
2009	-1.240(0.542)	-0.877(0.147)	-1.269(0.535)	-0.982(0.151)
2010	-0.783(0.543)	-0.614(0.145)	-0.898(0.535)	-0.845(0.146)
2011	-0.287(0.535)	-0.246(0.136)	-0.283(0.537)	-0.302(0.138)
2012	-0.714(0.492)	-0.981(0.137)	-0.608(0.448)	-0.929(0.137)
2013	0.905(0.417)	0.646(0.140)	0.858(0.412)	0.460(0.141)
2014	0	0	0	0
2015	0.104(0.396)	-0.242(0.137)	0.157(0.397)	-0.328(0.138)
2016	-0.373(0.457)	0.280(0.157)	-0.372(0.449)	0.158(0.158)
2017	-1.136(0.378)	-0.041(0.139)	-1.097(0.376)	-0.102(0.140)
2018	—	—	-1.139(0.398)	-0.862(0.153)
coastal	0	0	0	0
slipway	0.408(0.420)	0.761(0.282)	0.343(0.290)	0.761(0.282)
vehicles	3.567(1.070)	2.574(0.481)	3.879(1.050)	2.574(0.481)

Table 3a. 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**. Results shown are normalised to 2009=1. Previous results for these two approaches are given as well.

	Combined "old" and "new" – weight = 1	**Combined "old" and "new" – weight = 15; 10**	**Updated combined "old" and "new" – weight = 1**	**Updated combined "old" and "new" – weight = 15; 10**
2008	—	—	—	—
2009	1.000	1.000	1.000	1.000
2010	2.814	2.179	2.418	2.084
2011	1.207	0.968	1.025	0.827
2012	0.489	0.559	0.438	0.564
2013	0.422	0.285	0.372	0.276
2014	0.594	0.687	0.545	0.696
2015	1.019	0.663	0.947	0.690
2016	0.241	0.282	0.241	0.272
2017	0.498	0.311	0.655	0.267
2018	—	—	0.288	0.176

Table 3b. 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+**. Results shown are normalised to 2008=1. Previous results for these two approaches are given as well.

	Combined "old" and "new" – weight = 1	**Combined "old" and "new" – weight = 20; 10**	**Updated combined "old" and "new" – weight = 1**	**Updated combined "old" and "new" – weight = 20; 10**
2008	1.000	1.000	1.000	1.000
2009	0.903	1.233	0.910	1.258
2010	1.428	1.604	1.319	1.441
2011	2.343	2.317	2.440	2.481
2012	1.529	1.111	1.763	1.326
2013	7.716	5.657	7.639	5.316
2014	3.123	2.964	3.238	3.357
2015	3.467	2.326	3.788	2.417
2016	2.150	3.921	2.231	3.933
2017	1.003	2.845	1.081	3.030
2018	—	—	1.036	1.417

Table 4. The ratio of 2018 poaching indices to the 2017 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	0.658	0.502	0.861
Southern 8+	0.468	0.352	0.621

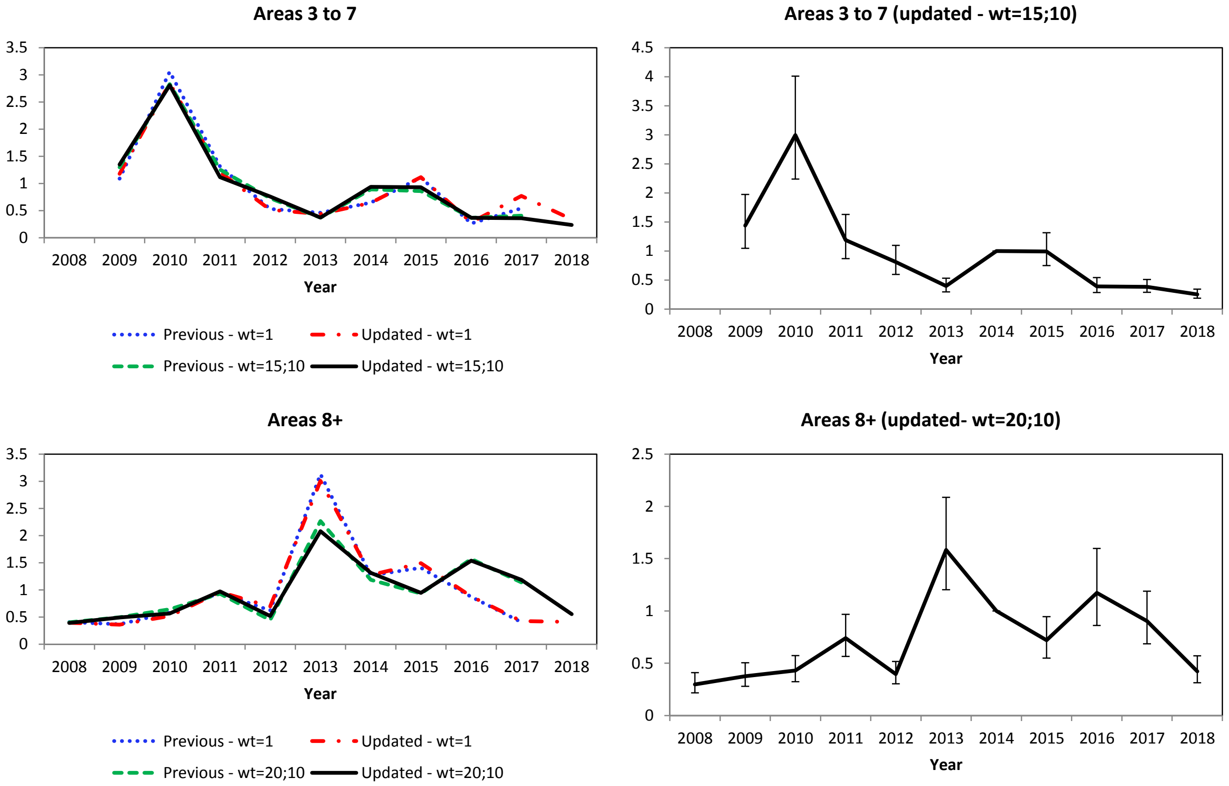

Figure 1. The plots on the left hand side show poaching trends (corresponding to year effects) for two different analysis approaches, with the previous results (to 2017 - Brandão and Butterworth, 2019) also plotted for comparison:

- 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 2017 for which they overlap.

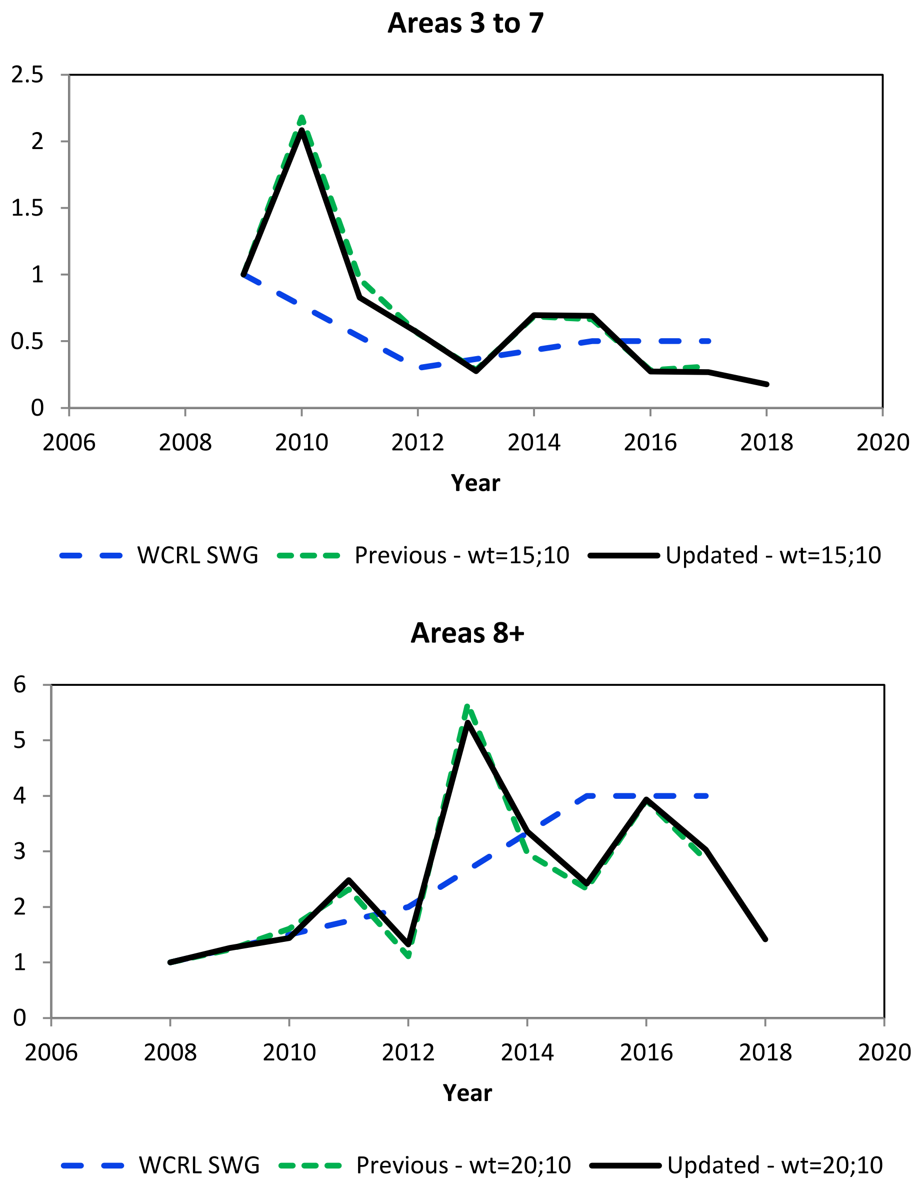

Figure 2. Poaching trends obtained using two different approaches:

- 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
- the WCRL SWG agreements on a simple characterisation of the poaching trends as assumed for the 2016 assessment ("WCRL SWG").

For comparison, the previous results for the first approach is also plotted. The plots described above are given for **Super-areas 3+4+5+6+7 (top)** and **Super-area 8+ (bottom)**. Results shown are normalised to 2008=1 for Super-area 8+ or to 2009=1 for Super-areas 3+4+5+6+7 as assumed for that previous assessment and projections.

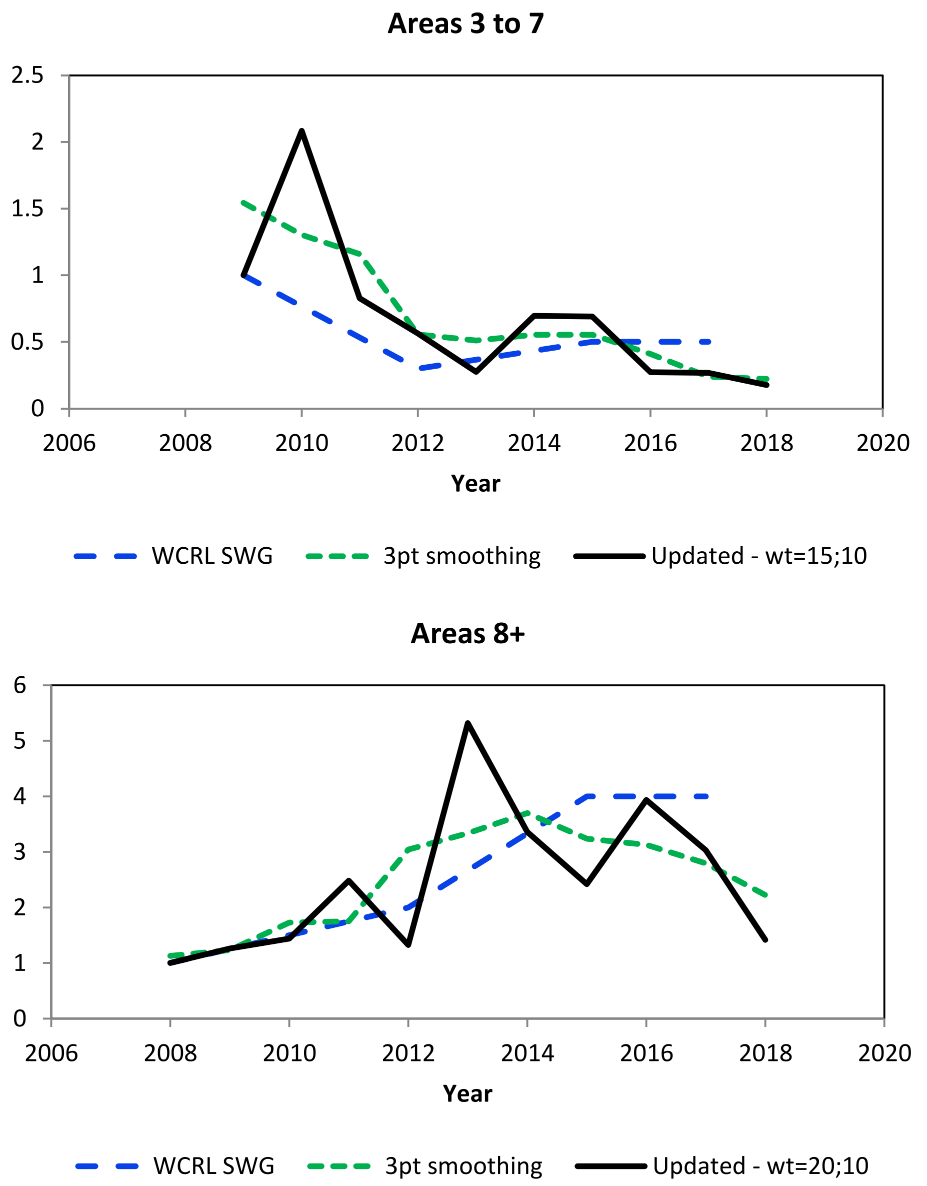

Figure 3. Poaching trends obtained using three different approaches:

- modelling of the combined "old" and "new" databases with the "old" database weighted by some factor – the approach now recommended (see text for details),
- the WCRL SWG agreements on a simple characterisation of the poaching trends as assumed for the 2016 assessment ("WCRL SWG"), and
- applying three-point smoothing to the poaching indices from the first approach.

The plots described above are given for **Super-areas 3+4+5+6+7 (top)** and **Super-area 8+ (bottom)**. Results shown are normalised to 2008=1 for Super-area 8+ or to 2009=1 for Super-areas 3+4+5+6+7 as assumed for that previous assessment and projections, for the first two approaches, but not the third.