Trends in poaching 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

April 2019

Abstract

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. These results are compared to two previous GLM methods reported by Brandão and Butterworth (2018a); a) a GLM applied to the "new" compliance database, and b) a GLM method applied to the "old" database in which estimates of relative effort efficiencies obtained from the "new" database are used to link effort to the number of confiscations ("old"-linked). By simultaneously modelling the "old" and the "new" databases, information on the trend in poaching, as well as on the relative efficiencies of different effort types from the "new" database, influence the estimates of poaching trends obtained from the combined model. However, due to the fact that the "old" database is a summation of operations which produced a number of instances of confiscations (including cases of zero confiscations), while the "new" database reports individual incidents with non-zero confiscations, undue weight is being given to the information available in the "new" database if data from each are equally weighted. We therefore advocate that an analysis which upweights the data in the "old" database should be used in the model which combines use of both the "old" and "new" database values, and that this should serve as the primary basis for inferring poaching trends.

Introduction

At the 2018 International Fisheries Stock Assessment Workshop, the Panel reviewed the analyses of Brandão and Butterworth (2018a) for estimating poaching trends that makes use of the analyses of the "new" database in which the policing effort is linked to the number of confiscations to obtain poaching trend estimates from the "old" database. This analysis led to quite imprecise estimates because of low sample sizes. The Panel recommended an alternative analysis that incorporates both datasets simultaneously (Cox *et al*. 2018). This paper gives the results for this alternative approach to obtain estimates for poaching trends. To aid comparison, results for the "old"-linked trends advocated by Brandão and Butterworth (2018a) 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 2017 form the "old" database. Data for the period 2012 to 2017 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 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 to which policing effort has been linked to the number of confiscations, 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 no zero confiscations are ever observed 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 distribution:

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

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

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 is 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 2017 for Super Area 8+ and 2009 to 2017 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 $\lambda^{new}, \alpha_m, \delta_y$ 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.

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 "new" database,
- to the "old" database using relative effort efficiencies from the "new" database model ("old"-linked),
- 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.

The precision of estimates for the month and year factors has been improved (with the exception of one or two values) when modelling the “new” and “old” databases simultaneously, compared to when modelling the “new” database alone. If the contribution of the “old” database is weighted as suggested, then the precision of the estimates is substantially improved. Standard errors for the parameter estimates for the combined model, which incorporates both the “old” and the “new” databases, are based on the Hessian.

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

The poaching trends obtained from the simultaneous analyses of the “old” and the “new” databases for the northern Super-areas (3-7) differ in the main in the last year, when modelling the “new” database alone. If the “old” database is not weighted in the combined model, the estimated poaching value for the last year is notably higher compared to both the estimates of the weighted combined approach and that of the “old”-linked model (Table 2 and Figure 1). This is indicative of the influence the trend information in the “new” database has on the estimated poaching values in the combined model. Weighting of data from the “old” database in the combined model shows similar trends to those of the “old”-linked model.

The effect of not weighting the “old” database is more marked for the southern Super-area 8+ (Figure 1), with the values for the last two years being reduced by the influence of the trend information in the “new” database. Again, weighting of data from the “old” database in the combined model shows similar trends to those of the “old”-linked model.

Conclusions

In the “old”-linked model, only information on the relative effort efficiencies obtained from the “new” database is used. By simultaneously modelling the “old” and the “new” databases, information on the trend in poaching in the “new” database is also used. However, due to the fact that data in the “old” database reflects a summation over operations which produced a number of confiscations (and some instances of no confiscations), while the “new” database reports individual positive confiscation incidents only, undue weight is being given to the information available in the “new” database when data from these two sources are given the same weight. We therefore advocate that the weighted combined ”old” and “new” database model trends shown in Figure 1 should serve as the primary basis for inferring poaching trends. These do suggest some downturn in the last two years in the northern areas, but seem to vary about a steady level from 2013 onwards for the southern Super-area 8+.

Figure 2 and Tables 4a-b compare these trends to the “old”-linked presented in Brandão and Butterworth (2018a), together with how they were smoothed for use in the base case assessment at that time. Importantly, because results here are shown relative to 2008 (for Super-area 8+) or 2009 (for the northern areas), the relative levels of poaching in the southern areas are notably different for the “old”-linked trends than those indicated by the combined model, even when a weight is applied. This “normalisation” relative to 2008/9 merits further discussion.

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.

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. Monthly maximum number of positive confiscations recorded over the years covered by the "new" database.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Areas 3-7	19	8	8	17	4	9	5	1	4	7	23	21
Areas 8+	14	12	10	12	19	5	2	5	5	5	8	32

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

	Poaching ("new")	Poaching ("old"-linked)	Combined "old" and "new" – weight = 1	Combined "old" and "new" – weight = 15; 10
January	0	0	0	0
February	1.300 (0.684)	0.848 (0.365)	1.443(0.440)	1.068(0.151)
March	0.819 (0.772)	1.077 (0.329)	1.464(0.468)	1.175(0.155)
April	0.442 (0.760)	0.214 (0.359)	0.672(0.421)	0.499(0.146)
May	0.780 (1.060)	-0.038 (0.379)	0.358(0.478)	-0.288(0.149)
June	0.043 (1.312)	-2.928 (1.278)	-1.079(0.463)	-2.452(0.169)
July	-0.725 (2.247)	-3.141 (1.239)	-1.614(0.468)	-2.321(0.167)
August	-0.365 (3.248)	-2.477 (0.773)	-1.156(0.489)	-1.831(0.162)
September	0.998 (1.169)	-1.231 (0.628)	-0.418(0.502)	-1.496(0.176)
October	1.519 (0.726)	-2.163 (0.720)	0.107(0.515)	-1.687(0.167)
November	-1.490 (1.417)	-1.351 (0.478)	-0.552(0.426)	-0.849(0.159)
December	1.326 (0.665)	0.194 (0.361)	1.213(0.513)	0.164(0.160)
2008	—	—	—	—
2009	—	0.409 (0.383)	0.521(0.512)	0.375(0.158)
2010	—	1.225 (0.320)	1.556(0.474)	1.154(0.143)
2011	—	0.142 (0.363)	0.710(0.498)	0.342(0.153)
2012	-1.620 (1.044)	-0.835 (0.420)	-0.194(0.455)	-0.206(0.152)
2013	-1.044 (0.593)	-0.594 (0.407)	-0.341(0.415)	-0.881(0.144)
2014	0	0	0	0
2015	-0.094 (0.408)	0.084 (0.373)	0.540(0.412)	-0.036(0.138)
2016	-1.203 (1.531)	-1.350 (0.950)	-0.900(0.480)	-0.891(0.166)
2017	0.401 (0.490)	-1.063 (0.558)	-0.176(0.444)	-0.793(0.140)
coastal	0	—	0	0
slipway	0.611 (0.568)	—	1.074(0.890)	-0.756(0.464)
vehicles	1.013 (0.571)	—	1.374(0.758)	0.515(0.211)

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

	Poaching ("new")	**Poaching ("old"-linked)**	**Combined "old" and "new" – weight = 1**	**Combined "old" and "new" – weight = 20; 10**
January	0	0	0	0
February	1.547 (0.423)	1.152 (0.433)	1.811(0.527)	1.168(0.142)
March	0.272 (0.682)	-1.099 (0.791)	0.446(0.589)	-0.588(0.148)
April	-1.102 (0.715)	0.571 (0.463)	0.360(0.475)	0.513(0.140)
May	-0.250 (0.667)	0.140 (0.514)	0.014(0.479)	0.077(0.144)
June	0.339 (0.629)	-0.119 (0.535)	0.193(0.523)	0.117(0.166)
July	-0.619 (1.493)	-1.278 (0.739)	-0.444(0.584)	-0.766(0.167)
August	-1.333 (1.936)	-2.983 (1.391)	-2.152(0.534)	-2.464(0.163)
September	2.400 (0.520)	-0.086 (0.531)	1.322(0.563)	0.389(0.169)
October	-1.380 (2.410)	-0.729 (0.613)	-0.867(0.566)	-0.356(0.170)
November	-0.993 (0.858)	-1.223 (0.746)	-1.110(0.470)	-0.976(0.163)
December	-2.310 (0.971)	-0.582 (0.589)	-1.774(0.409)	-0.849(0.140)
2008	—	-1.223 (0.812)	-1.139(0.602)	-1.087(0.160)
2009	—	-1.137 (0.696)	-1.240(0.542)	-0.877(0.147)
2010	—	-0.595 (0.549)	-0.783(0.543)	-0.614(0.145)
2011	—	-0.113 (0.487)	-0.287(0.535)	-0.246(0.136)
2012	-1.012 (1.417)	-1.033 (0.626)	-0.714(0.492)	-0.981(0.137)
2013	1.206 (0.474)	0.750 (0.424)	0.905(0.417)	0.646(0.140)
2014	0	0	0	0
2015	0.988 (0.431)	0.090 (0.456)	0.104(0.396)	-0.242(0.137)
2016	-0.073 (0.686)	0.673 (0.478)	-0.373(0.457)	0.280(0.157)
2017	-0.953 (0.782)	0.070 (0.452)	-1.136(0.378)	-0.041(0.139)
coastal	0	—	0	0
slipway	-0.357 (0.314)	—	0.408(0.420)	0.761(0.282)
vehicles	3.032 (0.596)	—	3.567(1.070)	2.574(0.481)

Table 4a. Poaching series obtained from a) the “old” database using relative effort efficiencies from the “new” database model (“old”-linked), b) the unweighted (i.e. weight = 1) combined “old” and “new” databases, and c) 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.

	“Old”-linked	**Combined “old” and “new” – weight = 1**	**Combined “old” and “new” – weight = 15; 10**
2008	—	—	—
2009	1.000	1.000	1.000
2010	2.262	2.814	2.179
2011	0.766	1.207	0.968
2012	0.288	0.489	0.559
2013	0.367	0.422	0.285
2014	0.664	0.594	0.687
2015	0.722	1.019	0.663
2016	0.172	0.241	0.282
2017	0.229	0.498	0.311

Table 4b. Poaching series obtained from a) the “old” database using relative effort efficiencies from the “new” database model (“old”-linked), b) the unweighted (i.e. weight = 1) combined “old” and “new” databases, and c) the weighted (see text for details) combined “old” and “new” databases for the southern **Super-area 8+**. Results shown are normalised to 2008=1.

	“Old”-linked	**Combined “old” and “new” – weight = 1**	**Combined “old” and “new” – weight = 20; 10**
2008	1.000	1.000	1.000
2009	1.090	0.903	1.233
2010	1.874	1.428	1.604
2011	3.033	2.343	2.317
2012	1.209	1.529	1.111
2013	7.189	7.716	5.657
2014	3.396	3.123	2.964
2015	3.715	3.467	2.326
2016	6.657	2.150	3.921
2017	3.642	1.003	2.845

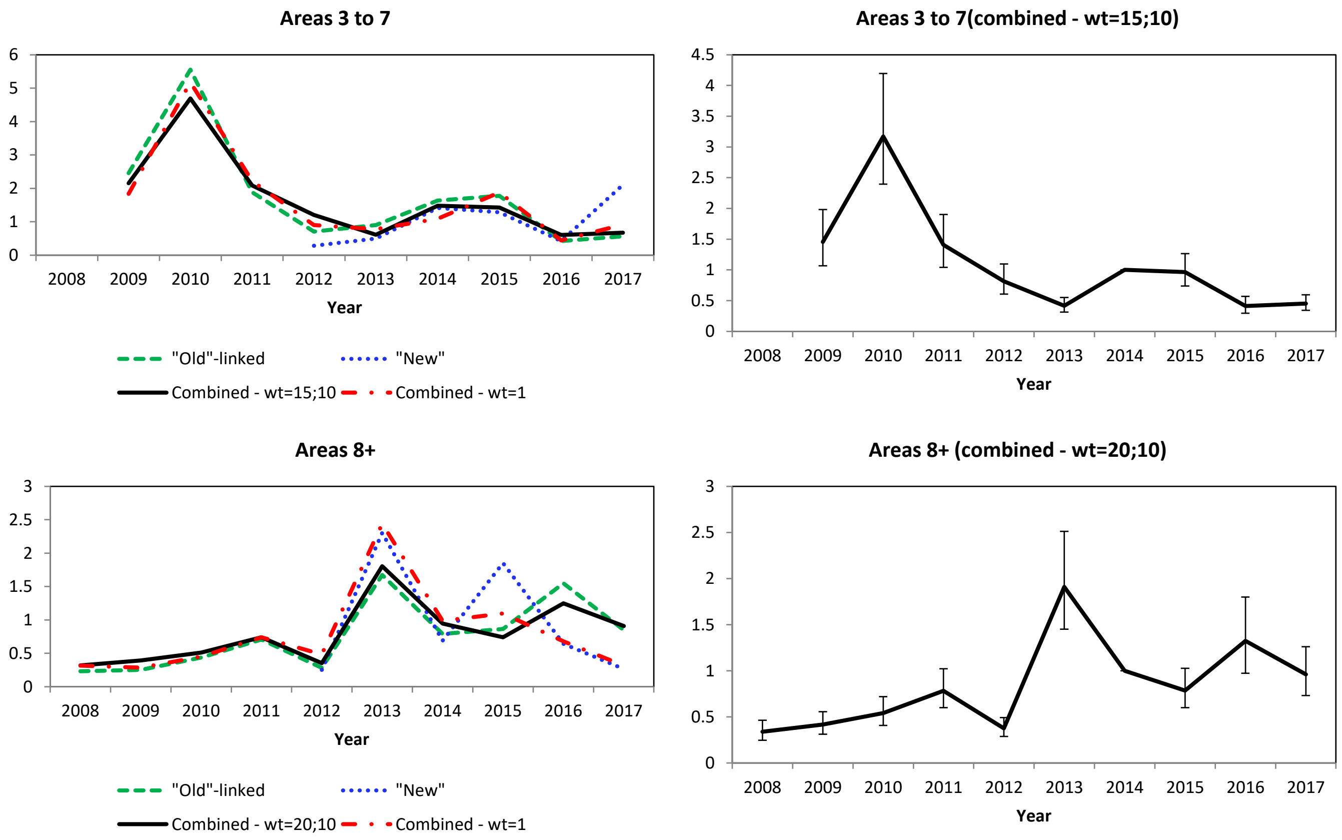
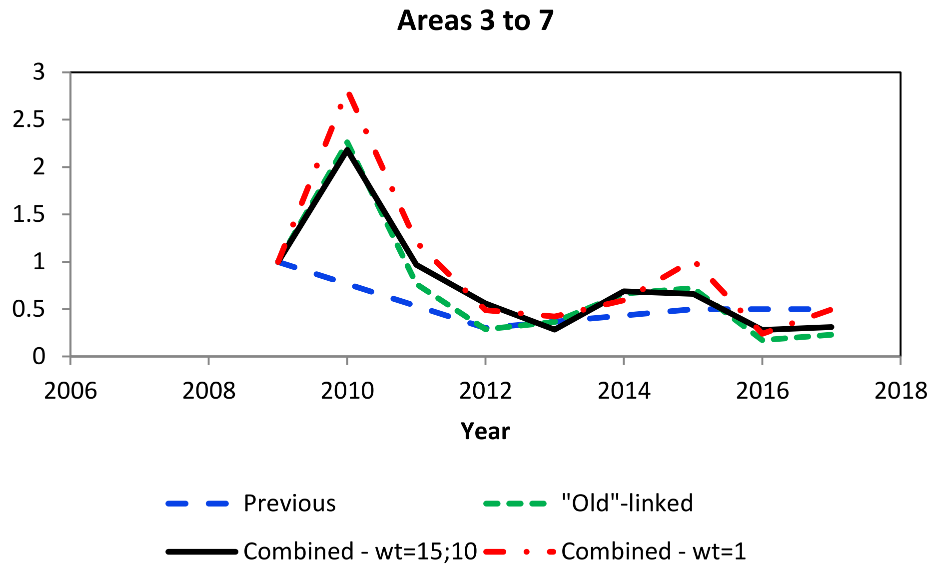


Figure 1. The plots on the left hand side show poaching trends (corresponding to year effects) for four different analysis approaches:

- modelling of the “old” database using relative effort efficiencies from the “new” database model (“old”-linked),
- modelling of the “new” database,
- 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.



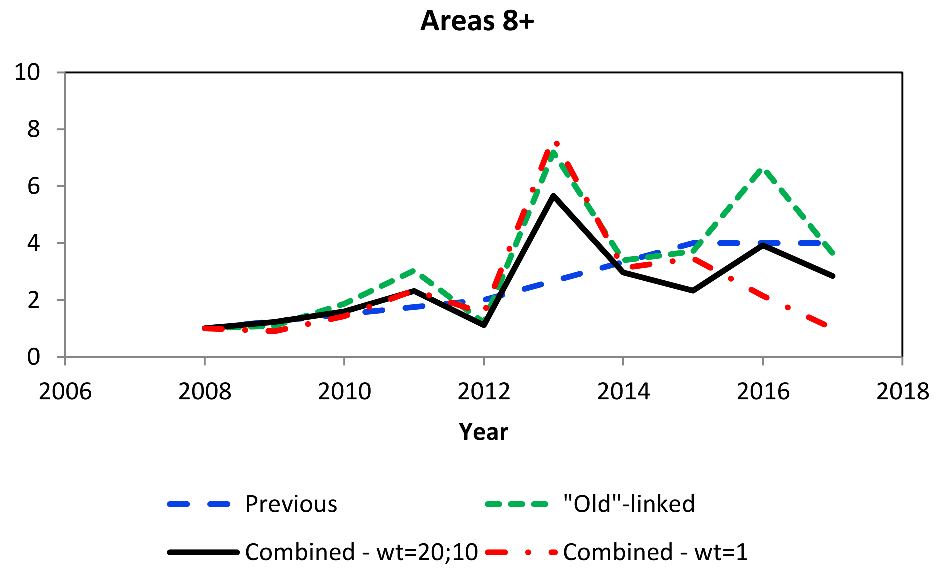


Figure 2. Poaching trends obtained using four different approaches:

- modelling of the “old” database using relative effort efficiencies from the “new” database model (“old”-linked),
- modelling of the combined “old” and “new” databases with the “old” database weighted by some factor – the approach now recommended (see text for details),
- modelling of the combined “old” and “new” databases with a weight of one (i.e. unweighted) applied to the “old” database, and
- the WCRL SWG agreements on a simple characterisation of the poaching trends as assumed for the 2016 assessment (“Previous”).

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.