NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

Northwest Atlantic Fisheries Organization

Serial No. N6907 **NAFO SCR Doc. 19/005**

SCIENTIFIC COUNCIL MEETING – JANUARY 2019

3M cod MSE: survey indices in the projection years

by

Carmen Fernández[1], Diana González-Troncoso[2], Fernando González-Costas[2], Thomas Brunel[3], Ricardo Alpoim[4], Antonio Ávila[4], Jose de Oliveria[5], Agurtzane Urtizberea[6] and Panayiota Apostolaki[7]

[1]Instituto Español de Oceanografía, Gijón, Spain
[2]Instituto Español de Oceanografía, Vigo, Spain
[3]Wageningen Marine Research, The Netherlands
[4]Instituto Português do Mar e da Atmosfera, Lisbon, Portugal
[5]Centre for Environment, Fisheries and Aquaculture Science, Lowestoft, UK
[6]Azti-Tecnalia, Pasaia, Spain
[7]MRAG Ltd, London, UK

Abstract

The survey indices in weight in the future to carry out the 3M Cod MSE are calculated from the numbers-at-age of each OM in the future multiplied by the weights in the projected scenario and by the catchability by age of the survey and raised by an error. Three different methods are presented to estimate that error, the first and the third based on a regression over the total biomass of the survey and the second based on a regression over the numbers-at-age of the survey. The first regression was fitted over three different periods of time, being the more rationale the one in which the years in which $SSB \geq B_{lim}$. As the results for all the methods seem to be quite the same, we chose the errors estimated by the first method for estimating the future survey indices in the MSE process.

Introduction

An MSE process for the 3M cod is in development, based on the Bayesian SCAA model approved in June 2018 by the Scientific Council (González-Troncoso *et al.*, 2018). Several OMs are projected in order to test different model-free HCRs, based on the survey indices. The aim of this document is to explain how the indices of the survey are estimated for the future. A 20 years projection (2018-2037) is performed for each of the iterations generated by the OMs.

Material and Methods

Data used

The data used in this work are those used as inputs in the 3M cod assessments and most of them come from the Flemish Cap survey (1988-2017) (González-Troncoso *et al.*, 2018). Table 1 presents the survey abundance indices at age, the total abundance and the total biomass by year of the Flemish Cap survey. Table 2 shows the mean weight-at-age observed in the survey.

Methodology

The survey indices in weight in the future ($Iwproj_y$) to carry out the 3M Cod MSE are calculated from the numbers-at-age of the OM (N_y^a) multiplied by the weights in the projected scenario ($wstockproj_y^a$) and by the catchability of the survey, q^a(constant over the years), and raised by an error:

$$Iwproj_y = \sum_{a=1}^{8+} q^a * N_y^a * wstockproj_y^a * e^{\varepsilon_y}$$

$wstockproj_y^a$ is the weight-at-age in the stock for the projected years. Four different scenarios are considered, as explained in González-Troncoso *et al.* (2019).

Different methods are analyzed in this document to produce the error in the equation, e^{ε_y}. Before describing the methods, let´s introduce some variables:

- lnI_a^y is the Naperian logarithm of the observed number-at-age by year in the survey (this is an input of the assessment, Table 1), *y*=1988,…,2017; *a*=1,…,8

-Iw_y is the "real" biomass of the survey by year estimated by the swept area method, *y*=1988,…,2017 (Table 1)

- $wstock_a^y$ is the mean weight-at-age in the survey (Table 2), *y*=1988,…,2017; *a*=1,…,8

- $lnIEst_a^y$ is the Naperian logarithm of the number-at-age by year in the survey generated from the results of the OM (*y*=1988,…,2017; *a*=1,…,8). We have 1000 iterations of this variable:

$$lnIEst_a^y = logphi_a + gama_a * \left[ln(N)_a^y + ln\left(e^{-alfa*Z_a^y} - e^{-beta*Z_a^y}\right) - ln(beta - alfa) - ln\left(Z_a^y\right)\right]$$, y=1988,…,2017; a=1,…8 (1)

- $IwEst_y$ is the total biomass of the survey by year from the numbers-at-age obtained in (1) and the mean weight-at-age in the survey, *y*=1988,…,2017:

$$IwEst_y = \sum_{a=1}^{8} e^{lnIest_a^y} * wstock_a^y \qquad (2)$$

The $IwEst_y$ median values for all the OMs are presented in Table 3.

The methods that we are going to discuss are:

1. First method: In this case, we do not use the numbers-at-age but the total biomass in weight, using $IwEst_y$ as a proxy of the "real" biomass of the index. For that, we fit a regression of $IwEst_y$ versus Iw_y (in logarithms) in the historic period (y=1988-2017) to get the error between them. After that, we can estimate Iw_y in the future adding to $IwEst_y$ the error got in this regression:

$$ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon, \quad \text{being} \quad \varepsilon \sim N(0, \sigma)$$

The intercept is set to be 0 and the slope to be 1, as it is supposed that the values of the estimated biomass should be very close to the "real" biomass. This fact was corroborated by a regression fitting the intercept and the slope ($ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$).

From here, we get the σ of the regression (from the R code). So, to generate the errors of the indices, we simulate, for each year of the projected period, 1000 values of a $N(0\,,1)$ and then multiply these values by the σ of the regression:

$$errorI\_1_y = N(0\,,1)_y * \sigma \text{, y=1,...,20, by iteration}$$

The same 1000 x 20 values from a $N(0\,,1)$ have been being used for all the OMs (we just generate once the 1000 values of the standard Normal distribution).

For fitting the regression, three trials were made changing the period used in the fit: with all the period of the survey (1988-2017), with the years in which the biomass was above B_{lim}=20000 tons (1988-1994, 2007-2017) and with the years in which the biomass was below B_{lim} (1995-2006).

In order to see the level of the correlation of the indices by year, the autocorrelation between the residuals of the indices, ϒ, is extracted from the regressions for all the OMs:

$$Residual_y = ln(IwEst_y) - ln(Iw_y)$$

$$\Upsilon = corr\left(ln(Residual_y), ln(Residual_{y-1})\right)$$

2. Second method: In this case we take into account the numbers-at-age and consider the CV used in the assessment fit (Bayesian SCAA). In the OMs 1, 2, 3, 4 and 6, the index of the assessment by age (numbers-at-age) is fitted as:

$$lnI_a^y \sim N(lnIEst_a^y, CV = 0.3)$$

So, in the future, as we know the value of $lnIEst_a^y$, we can generate a value of lnI_a^y generating a value of a normal of median $lnIEst_a^y$ and CV the one set in the OM.

In the case of the OM 5, the CV is estimated by the model via a prior. In this case, we use the posterior values of the CV generated by the model, so we have 1000 different values of CV, one for each group of ages (1, 2, 3+).

As our aim is to have the same structure for all the OMs, the way the simulation was made is the following, generating the errors of the indices:

We have that $sd = \sqrt{ln(1 + CV^2)}$. We generate, for each year and each age, 1000 values of a $N(0\,,1)$ and to get the error we multiply these values by the sd of the indices:

$$errorI\_2_{y}^{a} = N(0\,,1)_{y}^{a} * \sqrt{ln(1 + CV^2)}$$, y=1,…,20; a=1,…8, by iteration

The same 20 x 8 x 1000 values of the $N(0\,,1)$ have been being used for all the OMs (we just generate once the 1000 values of the standard Normal distribution).

3. Third method: As in the first method, we only calculate the total biomass of the survey but using an alternative way to calculate the index, calculating it from the total biomass obtained from the OM (B_y) via the total catchability of the biomass by year (q_y), in this way:

$$Iw\_B_y = q_y * B_y$$, y=1,…30

So, a regression over the Neperian logarithms in the 1988-2017 period of Iw_y over By is performed:

$$ln(Iw_y) \sim ln(q_y) + 1 * ln(B_y) + \varepsilon, \quad \text{being} \quad \varepsilon \sim N(0, \sigma)$$

With this regression, we estimate a catchability of the total biomass index (q_y) and an error and / or CV of the same. That q_y and that error are those that would be used in the future to estimate an index from the value of the total biomass in a year in the OM.

Results

Results of the first method

First, we fit a regression of $IwEst_y$ versus Iw_y, y=1988-2017, estimating the intercept and the slope ($ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$) in order to see if we can assume that they are 0 and 1, respectively. Fitting a "deterministic" regression (median of the estimated survey indices versus the "real" biomass) for all the OMs (Figure 1), the regressions have a good fit (R^2>0.9 in all cases) and the confidence intervals of the parameters include the 0 in the case of the intercept and the 1 in the case of the slope (Table 4) in almost all the cases (except OM 2 and OM 5), so we can conclude that we cannot say that a is different from 0 and b is different from 1.

Fitting the 1000 regressions of each iteration of $IwEst_y$ over Iw_y, the medians of the results are quite similar to fitting the median of the parameters.

The estimated error of the regression (sigma) for all the OMs are in Table 5, both for the deterministic fit (Det_1) and the median sigma of the 1000 iterations (Med_1). The value of sigma varies between 0.352 and 0.436 for this regression. Figure 2 shows the dispersion of sigma for each OM. We can see that the dispersion is low (CVs around 2.5% in most cases), although the value of sigma seems to be a bit high.

The autocorrelation of the indices is listed in Table 6 as Det_1 and Med_1, and the dispersion is in Figure 3. It is positive and around 0.5 in most cases but OM 4, so it exits a correlation between the indices.

Second, in light of the previous results, the regression fixing the parameters (a=0 and b=1) was made for all the historic years (1988-2017). The results of the deterministic regression for all the OMs are in Figure 4. Note that this is the fit of the logarithm of the "real" biomass over the median of the logarithm of the estimated index, so the black line of just the regression y=median($IwEst_y$).

The estimated error of the regression (sigma) for all the OMs are in Table 5, both for the deterministic fit (Det_2) and the median sigma of the 1000 iterations (Med_2). The value of sigma varies between 0.370 and 0.494. Note that for each OM, the value of sigma is quite similar to the one in the first case explained above.

In Figure 5 we can see, for all the OMs, the dispersion of sigma in the case of fitting the 1000 regressions of each iteration of $IwEst_y$ over Iw_y. The CV of sigma remains to be low, although a bit higher than in the previous case (CV around 3% in most cases).

The autocorrelation of the indices is listed in Table 6 as Det_2 and Med_2, and the dispersion is in Figure 6. As in the previous case, it is positive and around 0.5 in most cases but OM 4, so it exits a correlation between the indices.

Third, as the values of the sigma for the cases one and two is a bit high, in order to try to know from where the variation comes, the regression fixing the parameters (a=0 and b=1) was made only for the years in which SSB≥B_{lim} (1988-1994, 2007-2017). The results of the deterministic regression for all the OMs are in Figure 7.

The estimated error of the regression (sigma) for all the OMs are in Table 5, both for the deterministic fit (Det_3) and the median sigma of the 1000 iterations (Med_3). The value of sigma varies between 0.297 and 0.392. In this case the value of the sigma is lower than in the first and second case, which is an indication that most of the error in the regression comes from the years in which the biomass is low.

In Figure 8 we can see, for all the OMs, the dispersion of sigma in the case of fitting the 1000 regressions of each iteration of $IwEst_y$ over Iw_y. The CV of sigma remains to be low, although a bit higher than in the previous cases (CV around 4% in most cases).

The autocorrelation of the indices is listed in Table 6 as Det_3 and Med_3, and the dispersion is in Figure 9. In this case, the autocorrelation is negative and around 0.3 in most cases but OM 4, so the correlation is weaker than in the previous cases.

Forth, in order to contrast with the third regression, the regression fixing the parameters (a=0 and b=1) was made only for the years in which SSB<B_{lim} (1995-2006). The results of the deterministic regression for all the OMs are in Figure 10.

The estimated error of the regression (sigma) for all the OMs are in Table 5, both for the deterministic fit (Det_4) and the median sigma of the 1000 iterations (Med_4). The value of sigma varies between 0.336 and 0.680, given in all the OMs the highest variance of the four methods of regression, which corroborates the idea that most of the error in the regression comes from the years where the biomass was very low.

In Figure 11 we can see, for all the OMs, the dispersion of sigma in the case of fitting the 1000 regressions of each iteration of $IwEst_y$ over Iw_y. The CV of sigma remains to be low (CV around 4% in most cases).

The autocorrelation of the indices is listed in Table 6 as Det_4 and Med_4, and the dispersion is in Figure 12. In this case, the autocorrelation is positive and quite high, around 0.8 in all cases but OM 4.

In Figure 13, we can compare the results of the different fits applied to each OM.

As a conclusion, it was decided that, as in this moment the biomass is high (well above B_{lim}) and our aim is to maintain the SSB above B_{lim}, the most recommendable fit in this case is the third regression, fixing the parameters in 0 and 1 and taking only the years in which SSB≥B_{lim}. So, the value of sigma for this regression for each OM was multiplied by the values of the $N(0\,,1)$ generated for this case (common to all the OMs) to get the error of the index by this first method ($errorI\_1_y$), having for each OM a matrix of 1000 x 20 values (1000 iterations and 20 years to the future, 2018-2037). These errors are going to be multiplied by the index got in the future.

The median errors got by this method are shown in Figure 14. In general, the value of the errors is low and the variation between them in absolute numbers is small. There are more positive than negative errors, probably due to that we are taking only the years with high SSB. The median errors are quite similar between OMs; the most different values are the ones corresponding to the OM 4.

Results of the second method

The second method has been much easier to be developed. In this case, we just use the generating values of a $N(0\,,1)$ (that in this case are 1000 iterations x 8 ages x 20 years) and the value of the CV applied to the number-at-age in the survey in the OM in order to get the errors for the future ($errorI\_2_y^a$). This CV is the same in all OMs except for OMs 5. For this OM, the posterior of the CV was used. The value of the CVs of each OM is in Table 7, being the median of the posterior CV in the case of the OM 5. We can observe that if we leave the model to estimate the CV, the values are much higher than the 0.3 fix for the rest of the OMs, reaching levels around 0.8.

The median errors got by this method are shown in Figure 15. It can be seen that the median error is much higher in the case of the OM 5 (around three times more), as a consequence of the higher CV for the OM 5.

Results of the third method

The deterministic fit of the regression for all the OMs are in Figure 16. The value of q from the deterministic fit is around 0.7 for most of the OMs, being higher for the OM2 and lower for the OM4 (Table 8).

The estimated error of the regression (sigma) for all the OMs are in Table 5, both for the deterministic fit (Det_TM) and the median sigma of the 1000 iterations (Med_TM). The value of sigma varies between 0.414 and 0.452, given in all the OMs a sigma practically equal to the second case of the first method except for the OM4 in which the sigma of the third method is higher.

In Figure 17 we can see, for all the OMs, the dispersion of sigma in the case of fitting the 1000 regressions of each iteration of $IwEst_y$ over Iw_y, and in Figure 18 the dispersion of $logq$.

As the values of sigma are very similar in the chosen case for the first method and in the third method, it is clear that both methods are going to give the same results in the projected errors, so only one of them was implemented. The first method was chosen to fulfill this level of sigma just because the first method takes into account the q^a (the catchabilites by age).

Comparative results of the MSE

In order to see how the methods for estimating the errors of the survey in the future work, both errors, $errorI\_1_y$ and $errorI\_2_y^a$, were applied to (González-Costas *et al.*, 2019; González-Troncoso *et al.*, 2019):

-OM: base case
-Projection inputs: based on a bootstrap in years 2012-2017 (boot1) and on a random walk (RW)
-HCR: Model Free Trend

In the HCRs used, an estimation of the recruitment at age 1 in the survey, R_y^{surv}, is needed to calculate the catch in the next year. We obtain this index from the R_y of the OM and the catchability at age 1 in the OM, q^1, in this way:

$$ln(R_y^{surv}) = ln(R_y) + ln(q^1) + e^{errorI\_2_y^1}$$

being $errorI\_2_y^1$ the error obtained in the method 2 for age 1.

The comparisons of the results are in Figure 19 and 20, for boot1 and RW, respectively. It can be seen that the results are quite similar, both in the median and in the uncertainty, so used one or the other method does not make big differences in the projected results.

Discussion

We present three different methods for estimating the errors of the indices of the survey in the future, two based on the total biomass and other based on the numbers-at-age. Due to the problems of age readings in the case of 3M cod, it is possible that the methods that do not use ages are more appropriate. As the first and the third method give virtually the same results for all the OMs but 4, the first method was chosen to fulfill the uncertainty in this case as it takes into account the q^a (the catchabilites by age).

To develop the first method, four different regressions were fit. It seems quite reasonable, at least for most of the OMs, to fit a regression in which the parameters are fix, being 0 the intercept and 1 the slope, as the estimated total biomass by the model should be very similar to the “real” one. Fitting only the years in which the SSB ≥ B_{lim} seems to be the most rational case, since what is intended in the future is to maintain the biomass above B_{lim}. Most of the error in the regression comes from the years in which the biomass was very low (SSB < B_{lim}) and variable.

As it is supposed that the indices of the survey must be correlated by time, as in general a stock increase or decrease by periods of time, the autocorrelation of the regression has been analyzed, concluding that in general the autocorrelation is quite high except for the chosen case. The best way to proceed in these cases is to fit the residuals of the regression to an AR(1) to get the index of the autocorrelation, but this exercise has been not developed yet. At the view of the results we do not know if it´s worth implementing an AR(1), as the autocorrelation in the chosen method is not too high.

It is remarkable that the autocorrelation is positive in all the scenarios studied except the chosen one (fitting just the years in which $SSB \geq B_{lim}$), and in this case we have the smaller autocorrelation, too. The highest positive autocorrelation corresponds to the case in which we fit only the years with SSB below B_{lim}. This suggest that, once the stock is in low levels, it is difficult to reach again high levels. Instead of that, as the correlation for the years with high biomass is negative, it seems that it is easier to reverse a condition of high biomass.

As the results for first and second methods seem to quite the same, we chose the errors estimated by the first method for the future indices survey indices to implement the MSE. The rationality of this choice is that the first method is over the total index of biomass of the survey rather than over the numbers-at-age, which implies on one hand that the ALK is not used for the dependent variable of the regression, an advantage due the problems in the age-reading for this species, and on the other hand a simpler way to fit the model is performed.

References

González-Costas, F., D. González-Troncoso, C. Fernández, A. Urtizberea, R. Alpoim, A. Avila de Melo, J. De Oliveira, P. Apostolaki, T. Brunel and D. García, 2019. Potential Operating Models, Harvest Control Rules and Performance Statistics for the NAFO 3M Cod MSE. NAFO SCR Doc. 19/001. Serial No. N6903.

González-Troncoso, D., F. González Costas, C. Fernández A. Urtizberea, R. Alpoim, A. Avila de Melo, T. Brunel, J. De Oliveira and P. Apostolaki, 2019. Specifications of the OMs and the projections for the 3M cod MSE. NAFO SCR Doc. 19/006. Serial No. N6908.

González-Troncoso, D., C. Fernández and F. González Costas, 2018. Assessment of the Cod Stock in NAFO Division 3M. NAFO SCR Doc. 18/042. Serial No. N6833.

Acknowledges

This research was partially funded by the European Union founds under the Framework Contract No EASME/EMFF/2016/008 - "provision of scientific advice for fisheries beyond EU waters”- Specific Contract No 03 (SI2.753135) “Support to a robust model assessment, benchmark and development of a Management Strategy Evaluation for cod in NAFO Division 3M”.

Table 1. EU bottom trawl survey numbers-at-age (I_a^y) and total (thousands) and total biomass, Iw_y (tons). These values are inputs of the OMs.

	1	2	3	4	5	6	7	8+	Total Abundance	Total Biomass
1988	4868	79905	49496	13448	1457	211	225	72	149683	40839
1989	19604	10800	91303	54613	20424	1336	143	139	198363	114050
1990	2303	12348	5121	16952	15834	4492	340	247	57637	59362
1991	129032	26220	16903	2125	6757	1731	299	113	183181	40248
1992	71533	41923	5578	2385	385	1398	244	21	123468	26719
1993	4075	138357	31096	1099	1317	173	489	87	176693	60963
1994	3017	4130	27756	5097	130	67	7	116	40319	26463
1995	1425	11901	1338	3892	928	33	23	26	19567	9695
1996	36	3121	6659	892	2407	192	8	5	13320	9013
1997	37	150	3478	4803	391	952	21	4	9837	9966
1998	23	83	95	1256	1572	78	146	6	3259	4986
1999	5	84	116	117	717	444	19	5	1507	2854
2000	178	16	327	198	96	446	172	38	1470	3062
2001	473	1990	13	122	79	15	142	117	2951	2695
2002	0	1330	641	29	70	33	26	130	2261	2496
2003	684	54	628	134	22	42	7	71	1642	1593
2004	14	3380	25	600	168	5	10	23	4226	4071
2005	8069	16	1118	78	709	136		41	10166	5242
2006	19709	3886	62	1481	85	592	115	35	25965	12505
2007	3917	11620	5022	21	1138	58	425	107	22308	23886
2008	6096	16671	12433	4530	72	946	56	320	41124	43676
2009	5139	7479	16150	14310	4154	26	1091	349	48697	75228
2010	66370	27689	8654	7633	4911	1780	8	766	117810	69295
2011	347674	142999	16993	6309	7739	3089	1191	304	526300	106151
2012	103494	128087	10942	11721	4967	4781	1630	1098	266720	113227
2013	5525	67521	32339	4776	4185	2782	1807	1346	120280	72289
2014	7282	2372	48564	43168	17861	6842	3447	4223	133760	159939
2015	1141	12952	7250	25614	14107	21854	3434	2812	89164	114807
2016	56	4485	14356	2230	14540	12375	4814	2716	55032	80583
2017	1714	484	9895	7051	12486	14741	8019	2850	57241	89414

Table 2. Weight-at-age (kg) in stock, $wstock_a^y$.

	1	2	3	4	5	6	7	8+
1988	0.032	0.106	0.308	0.664	1.970	3.500	5.742	6.954
1989	0.036	0.101	0.330	0.836	1.293	2.118	4.199	7.360
1990	0.043	0.181	0.354	0.868	1.566	2.507	4.132	6.572
1991	0.056	0.171	0.501	0.865	1.594	2.593	3.423	6.182
1992	0.056	0.247	0.485	1.394	1.723	2.578	3.068	9.406
1993	0.043	0.227	0.657	1.216	2.279	2.381	3.373	5.731
1994	0.063	0.214	0.599	1.321	2.132	4.054	4.119	6.555
1995	0.048	0.243	0.479	0.969	1.851	2.680	5.532	7.309
1996	0.044	0.260	0.544	0.813	1.331	2.252	4.079	5.118
1997	0.081	0.333	0.652	1.020	1.327	2.092	1.997	9.717
1998	0.073	0.371	0.773	1.206	1.684	2.015	3.070	7.525
1999	0.108	0.398	0.946	1.329	1.866	2.444	3.461	4.987
2000	0.106	0.606	0.971	1.638	1.940	2.860	3.461	7.985
2001	0.084	0.493	1.281	1.724	2.588	3.488	3.893	5.137
2002	0.071	0.440	1.191	1.540	2.661	3.916	5.302	5.672
2003	0.058	0.337	0.926	1.566	3.047	3.769	5.721	6.451
2004	0.004	0.620	1.488	2.098	3.332	4.808	6.207	7.886
2005	0.084	0.580	1.256	2.242	2.875	4.187	6.033	8.148
2006	0.096	0.720	1.096	2.549	3.644	4.777	5.858	9.691
2007	0.053	0.609	1.640	3.478	4.097	5.787	6.373	8.315
2008	0.068	0.382	1.344	2.695	3.191	5.015	6.324	7.938
2009	0.078	0.407	0.976	2.072	3.881	6.958	6.583	9.461
2010	0.061	0.384	1.089	1.677	2.956	5.379	7.616	9.144
2011	0.038	0.211	0.913	1.618	2.339	3.594	6.050	9.396
2012	0.074	0.369	0.726	1.349	1.988	2.656	4.933	7.812
2013	0.071	0.175	0.687	1.159	2.004	2.750	4.206	7.614
2014	0.048	0.169	0.354	1.059	1.623	2.536	3.846	8.444
2015	0.049	0.156	0.469	0.747	1.216	1.847	3.434	6.775
2016	0.044	0.169	0.412	0.783	1.304	2.024	2.883	6.905
2017	0.042	0.098	0.421	0.678	1.058	1.980	2.754	5.905

Table 3. Results of the OMs: median predicted EU bottom trawl survey total biomass $IwEst_y$ (tons) compared with the "real" biomass of the survey, Iw_y

OM	OM1	OM2	OM3	OM4	OM5	OM6	Original
1988	53033	60204	53284	58324	57670	53688	40839
1989	55691	61950	57851	63196	61673	57165	114050
1990	43701	46381	45487	51444	49478	44457	59362
1991	36990	41271	35445	37192	42008	37563	40248
1992	32334	38115	30212	29037	33777	32627	26719
1993	33153	40342	31879	28541	34155	33737	60963
1994	22076	26013	21882	20886	22545	22883	26463
1995	7402	8157	7578	8663	7487	7696	9695
1996	4573	4919	4573	5662	4361	4826	9013
1997	2917	2688	3024	4684	2752	3251	9966
1998	2019	1528	2096	3606	1772	2303	4986
1999	1800	1262	1805	2699	1445	2145	2854
2000	1907	1422	1885	2271	1545	2043	3062
2001	2108	1727	2071	2131	1826	1991	2695
2002	2769	2486	2766	2608	2543	2693	2496
2003	3021	2896	3116	2686	2915	3019	1593
2004	6006	5526	6235	4770	5868	6107	4071
2005	7614	7179	7822	5368	7410	7988	5242
2006	15450	13999	15791	13153	15551	15882	12505
2007	28336	25887	28953	18576	29301	28536	23886
2008	40949	37653	43228	26765	43696	42416	43676
2009	55824	52495	59951	46810	59351	56640	75228
2010	70394	66931	77867	79400	73589	73993	69295
2011	72272	68717	78943	120666	74206	74790	106151
2012	96671	92300	105296	166900	97992	98912	113227
2013	107553	104349	111715	93490	109740	110613	72289
2014	107886	106882	107435	72838	110356	111737	159939
2015	93877	97279	88060	73568	94330	97116	114807
2016	98074	109567	85269	85998	98091	100761	80583
2017	87169	108254	71226	87927	88169	82706	89414

Table 4. Confidence interval for all the OMs for the parameters of the regression $ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$.

OM	CI	2.5%	97.5%
1	Intercept	-0.305	1.869
	Slope	0.828	1.046
2	Intercept	0.311	2.430
	Slope	0.772	0.985
3	Intercept	-0.354	1.827
	Slope	0.831	1.050
4	Intercept	-0.644	1.247
	Slope	0.889	1.079
5	Intercept	0.149	2.266
	Slope	0.788	1.000
6	Intercept	-0.436	1.675
	Slope	0.844	1.056

Table 5. Sigma of the fits for all the OMs. Det: deterministic; Med: median of the 1000 iterations. 1: First regression: $ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$; 2: Second regression: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$; 3: Third regression: as second regression only fitted in the years in which SSB≥B_{lim}; 4: Forth regression: as second regression only fitted in the years in which SSB<B_{lim}. TM: third method.

OM	Det_1	Med_1	Det_2	Med_2	Det_3	Med_3	Det_4	Med_4	Det_TM	Med_TM
1	0.420	0.420	0.444	0.446	0.317	0.320	0.585	0.588	0.449	0.452
2	0.435	0.436	0.492	0.494	0.310	0.312	0.679	0.680	0.446	0.447
3	0.419	0.420	0.441	0.443	0.313	0.315	0.582	0.584	0.412	0.414
4	0.352	0.353	0.370	0.372	0.390	0.392	0.336	0.338	0.531	0.532
5	0.427	0.432	0.471	0.481	0.297	0.311	0.650	0.655	0.500	0.506
6	0.402	0.403	0.414	0.417	0.308	0.310	0.536	0.538	0.445	0.448

Table 6. Autocorrelation of the residuals (ϒ) of the fits for all the OMs. Det: deterministic; Med: median of the 1000 iterations. 1: First regression: $ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$; 2: Second regression: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$; 3: Third regression: as second regression only fitted in the years in which SSB>B_{lim}; 4: Forth regression: as second regression only fitted in the years in which SSB<B_{lim}.

OM	Det_1	Med_1	Det_2	Med_2	Det_3	Med_3	Det_4	Med_4
1	0.482	0.483	0.508	0.511	-0.297	-0.288	0.813	0.812
2	0.487	0.488	0.562	0.563	-0.249	-0.241	0.803	0.803
3	0.495	0.494	0.519	0.521	-0.309	-0.299	0.815	0.814
4	0.224	0.228	0.224	0.231	0.044	0.048	0.577	0.575
5	0.498	0.503	0.563	0.571	-0.343	-0.293	0.825	0.821
6	0.432	0.436	0.451	0.456	-0.314	-0.305	0.784	0.784

Table 7. Value of the CV for calculating the errors of the index in the future for each OM. For OMs 1 to 4 and 6, is the same for all ages and equal to the CV of the prior of the survey numbers-at-age in the assessment. For OM 5 is median the posterior of the CV by groups of ages (1, 2, 3+).

OM	1-4, 6	5		
Age	1-8+	1	2	3+
CV	0.3	0.87	0.85	0.74

Table 8. Value of the q from the regression $ln\left(Iw_y\right) \sim ln(q_y) + 1 * ln\left(B_y\right) + \varepsilon$. Det: deterministic; median: median of the 1000 iterations

OM	OM1	OM2	OM3	OM4	OM5	OM6
Det	0.708	1.095	0.735	0.462	0.745	0.685
Median	0.707	1.096	0.736	0.463	0.744	0.685

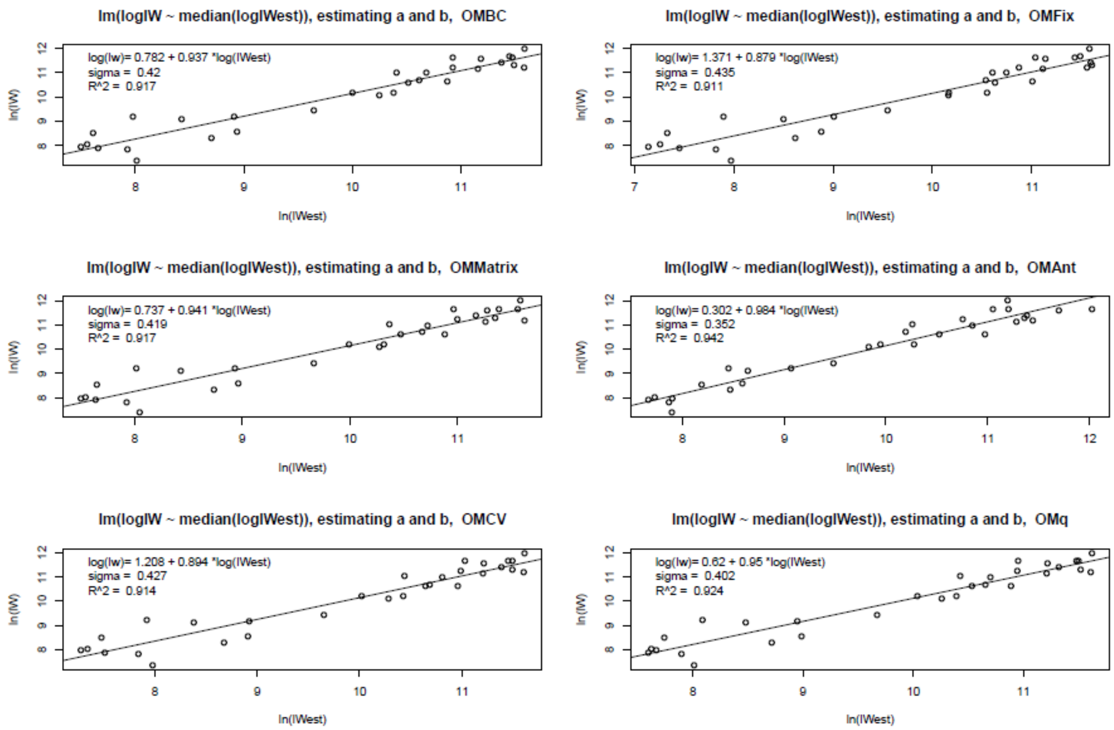


Figure 1. Regression of $IwEst_y$ versus Iw_y for all the OMs (method 1) estimating the intercept and the slope: $ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$

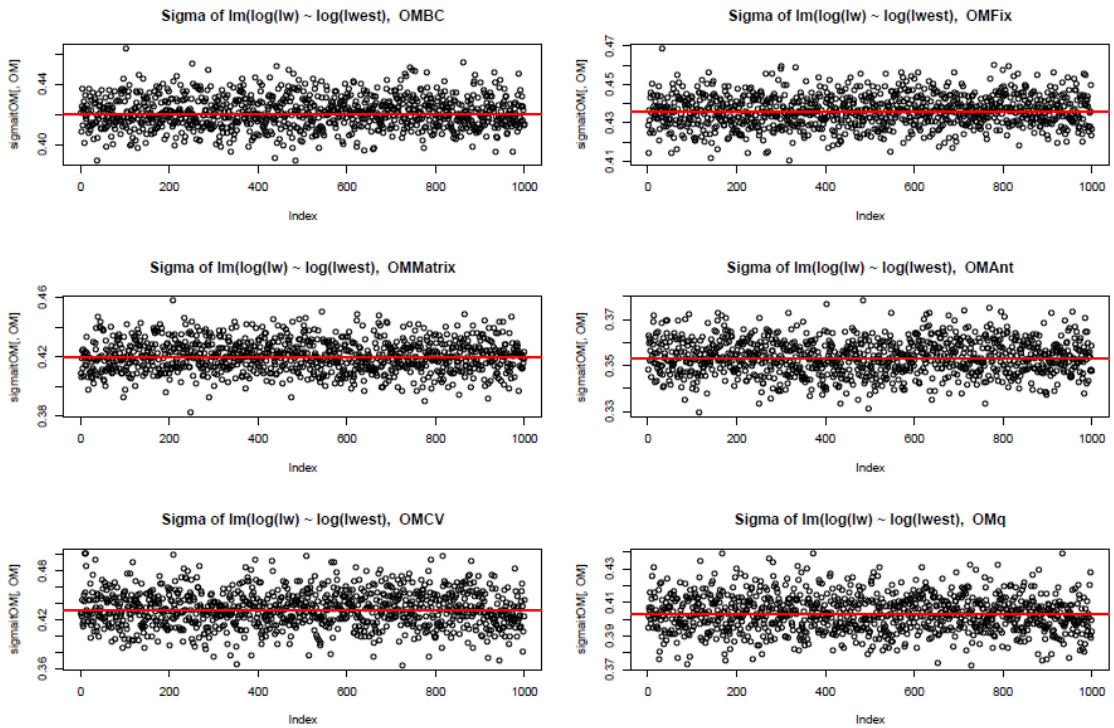


Figure 2. Dispersion of sigma for all the OMs in the first regression model, regression of $IwEst_y$ versus Iw_y estimating the parameters: $l\,n(Iw_y) \sim a + b * l\,n(IwEst_y) + \varepsilon$

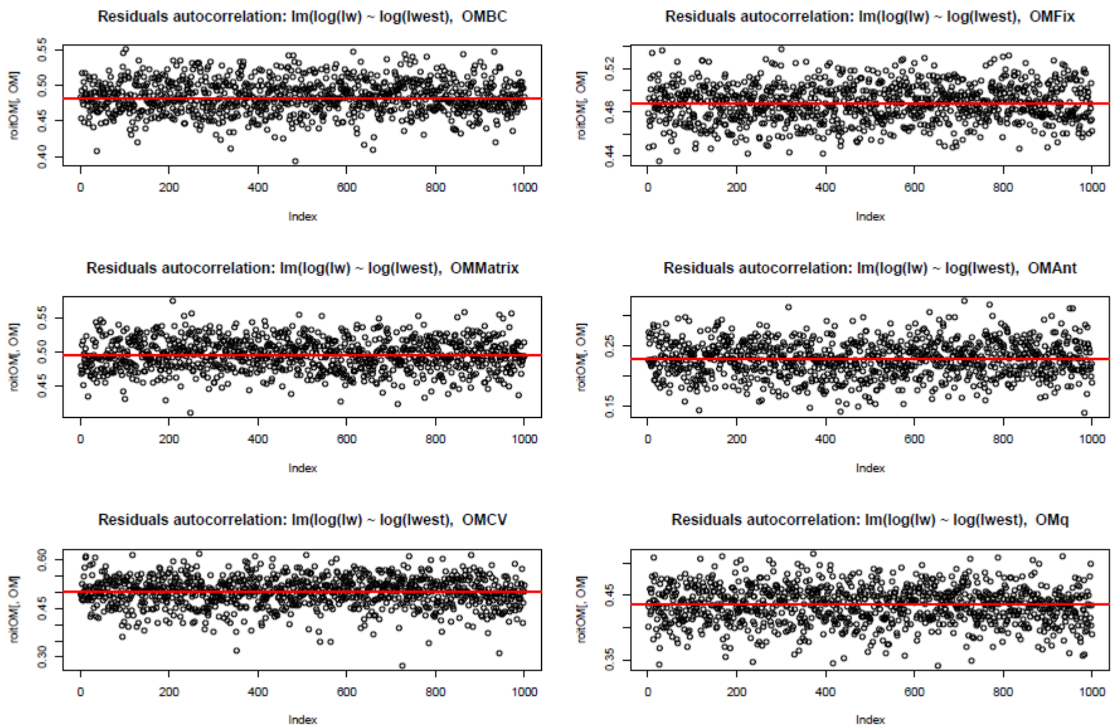


Figure 3. Dispersion of the autocorrelation of the residuals (ϒ) for all the OMs in the first regression model, regression of $IwEst_y$ versus Iw_y estimating the parameters: $ln(Iw_y) \sim a + b * ln(IwEst_y) + \varepsilon$

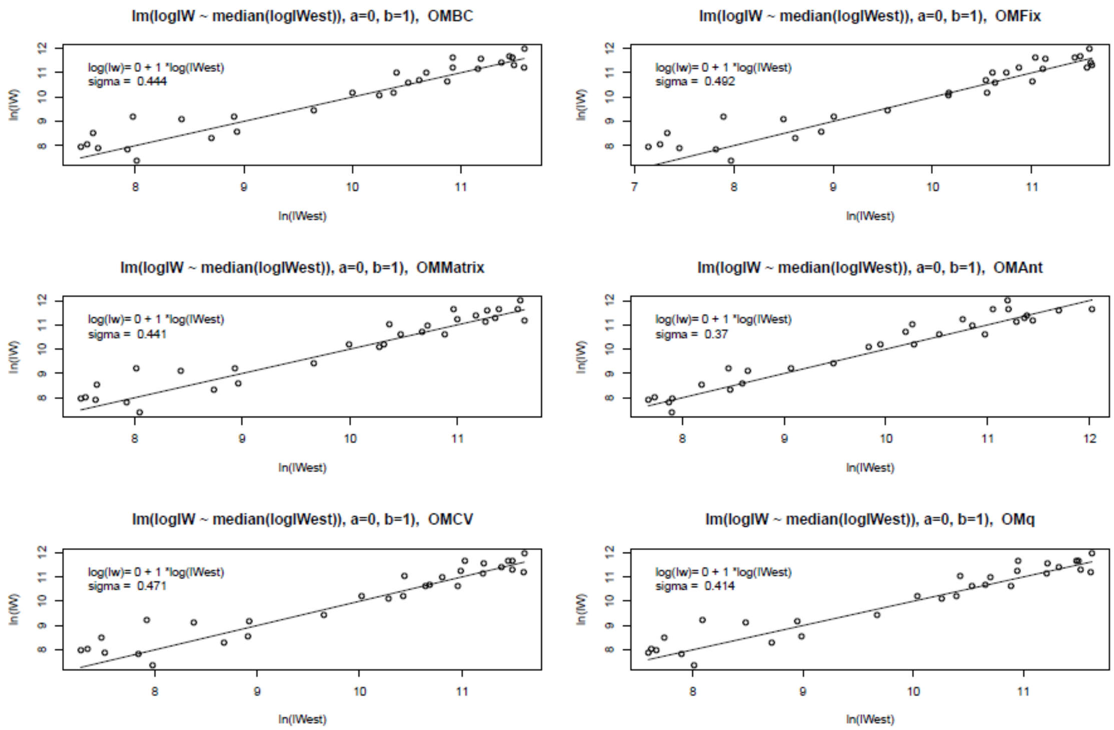


Figure 4. Regression of $IwEst_y$ versus Iw_y for all the OMs assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$

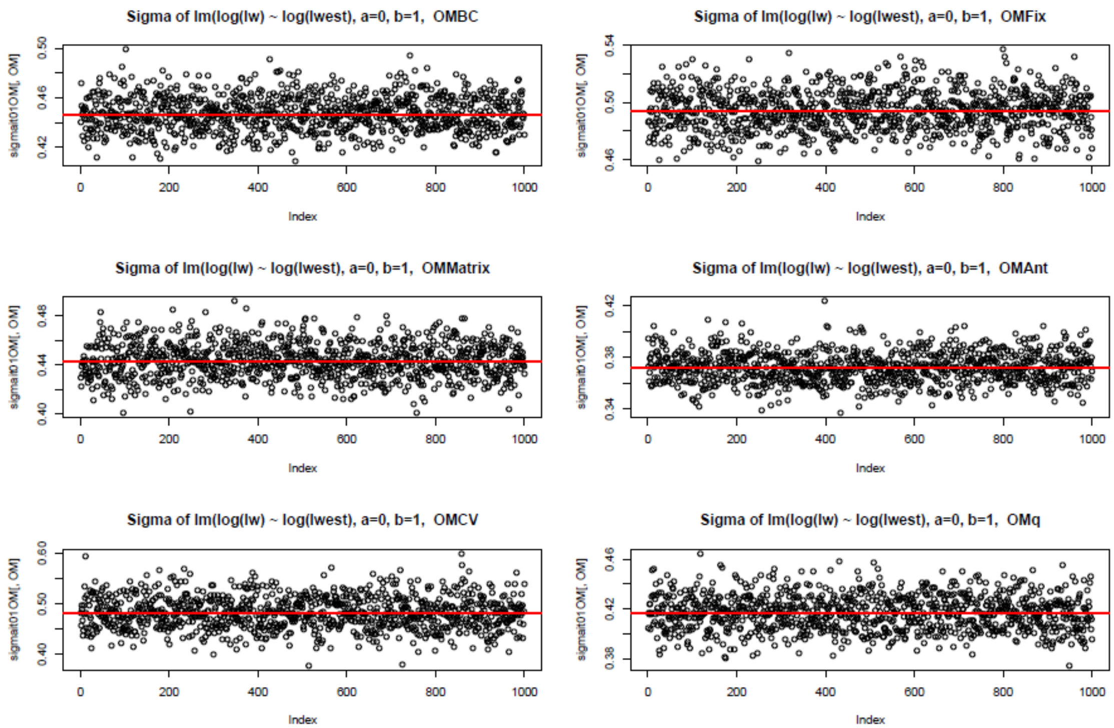


Figure 5. Dispersion of sigma for all the OMs in the second regression model, regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $l\,n(Iw_y) \sim 0 + 1 * l\,n(IwEst_y) + \varepsilon$

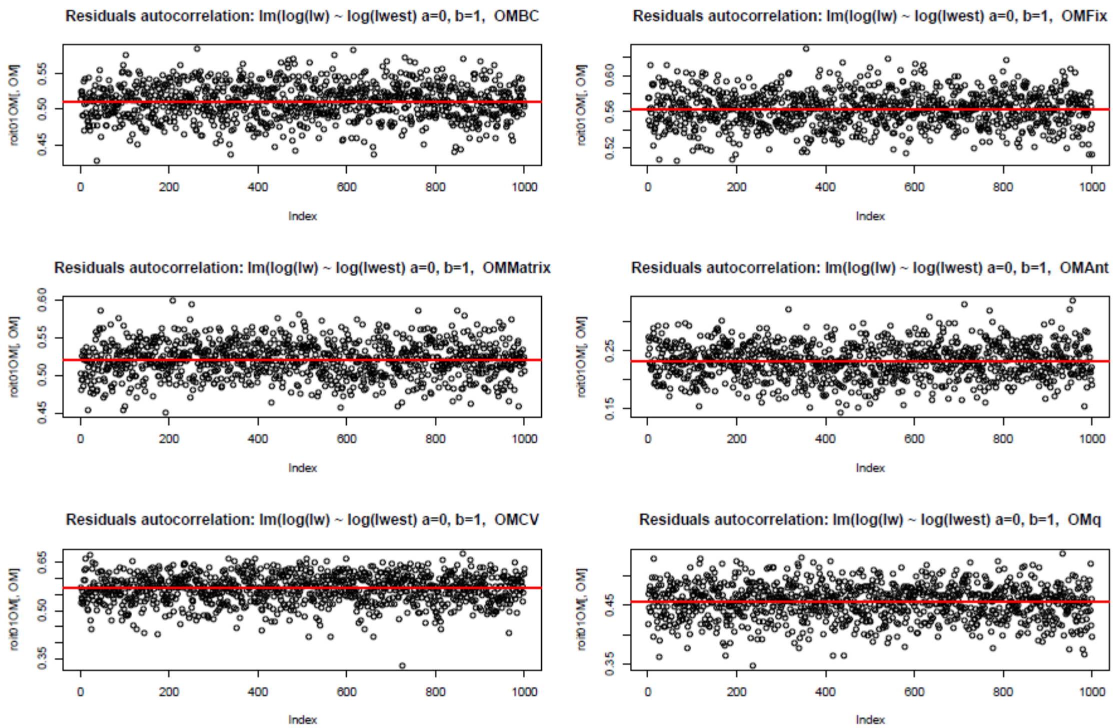


Figure 6. Dispersion of the autocorrelation of the residuals (ϒ) for all the OMs in the second regression model, regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$

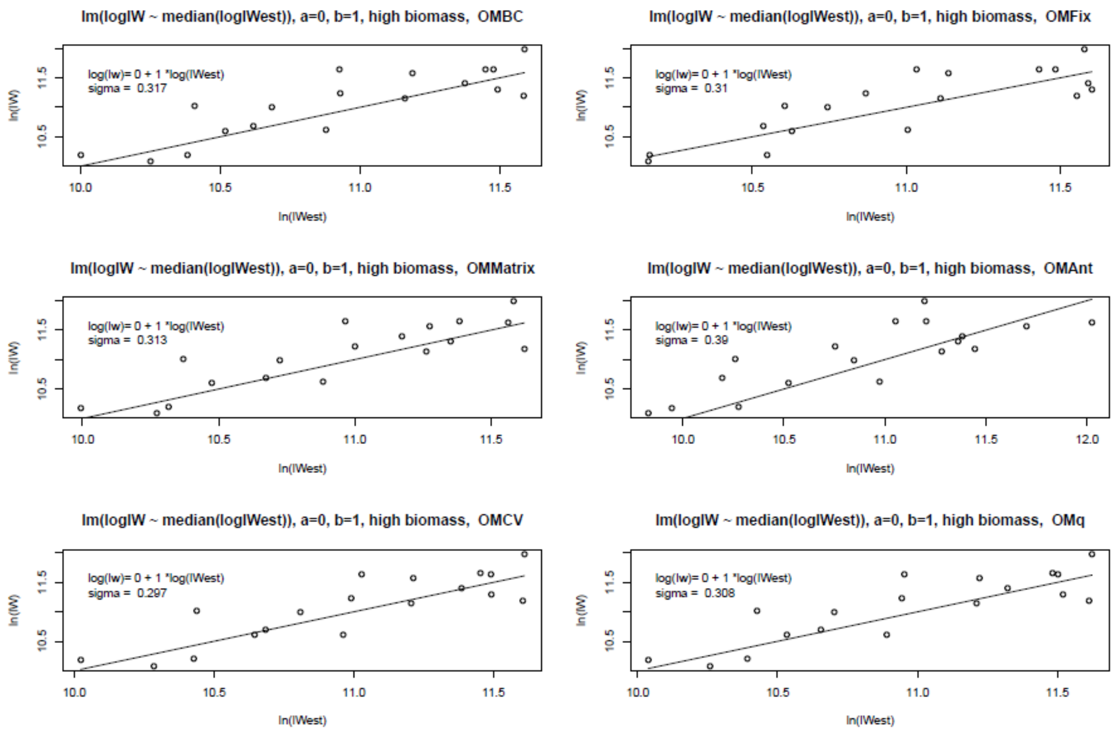


Figure 7. Regression of $IwEst_y$ versus Iw_y for all the OMs assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$ fitting the years in which SSB≥B_{lim}.

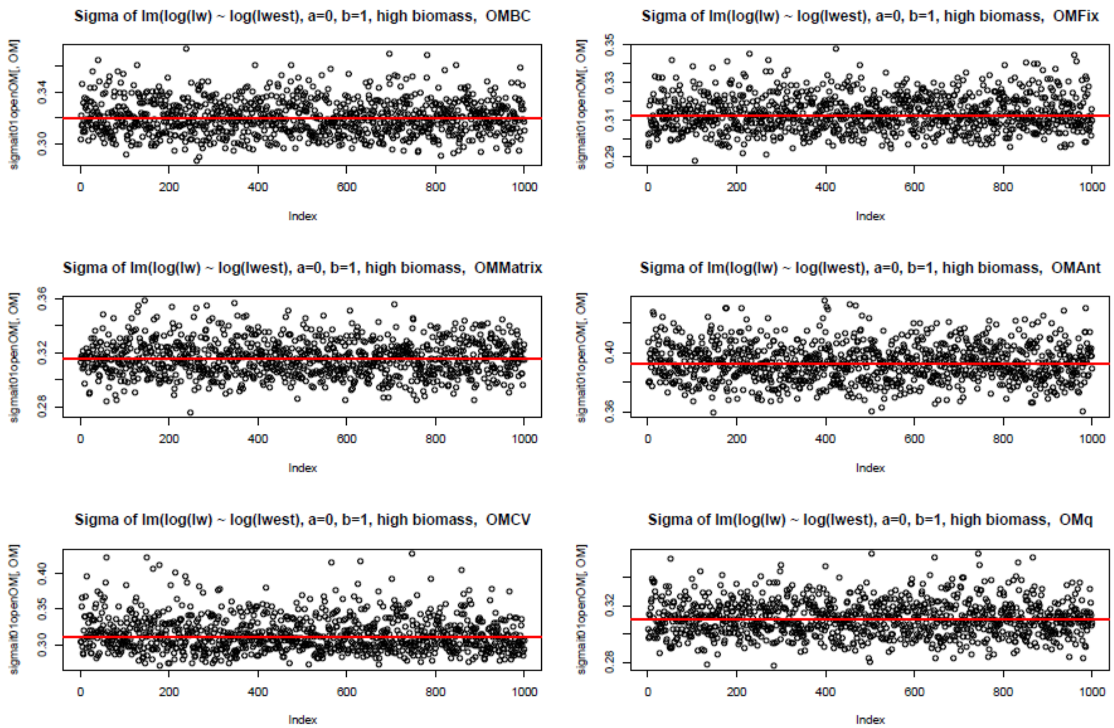


Figure 8. Dispersion of sigma for all the OMs in the second regression model, regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$ fitting the years in which SSB>B_{lim}.

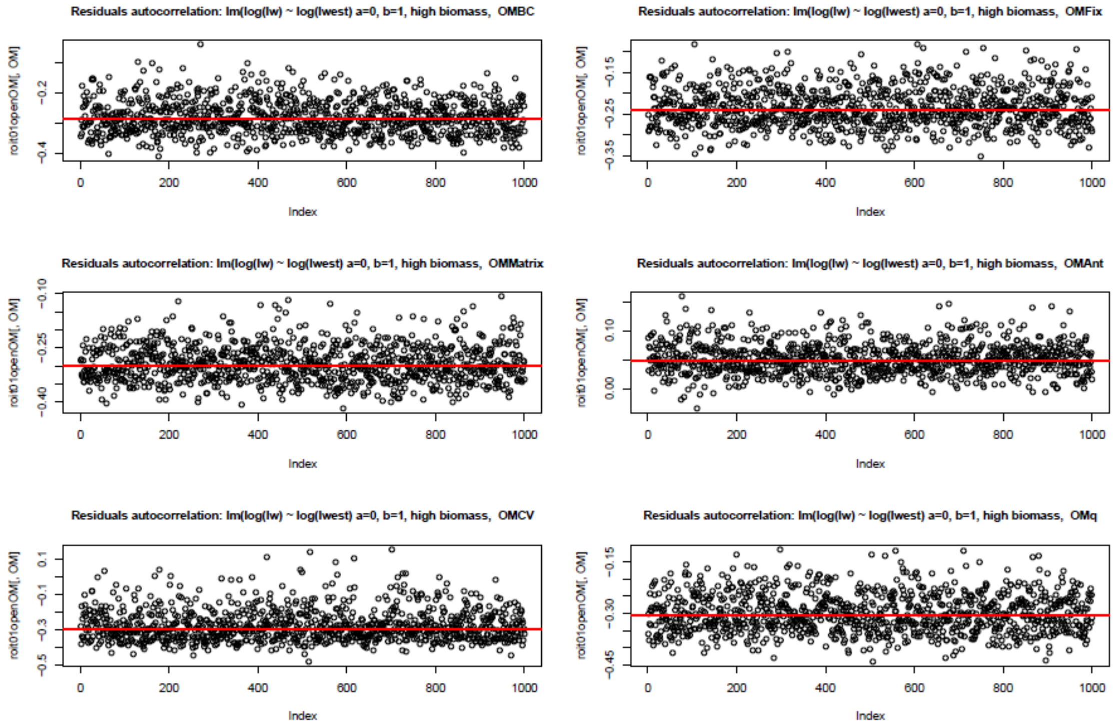


Figure 9. Dispersion of the autocorrelation of the residuals (Υ) for all the OMs in the second regression model, regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$ fitting the years in which SSB>B_{lim}.

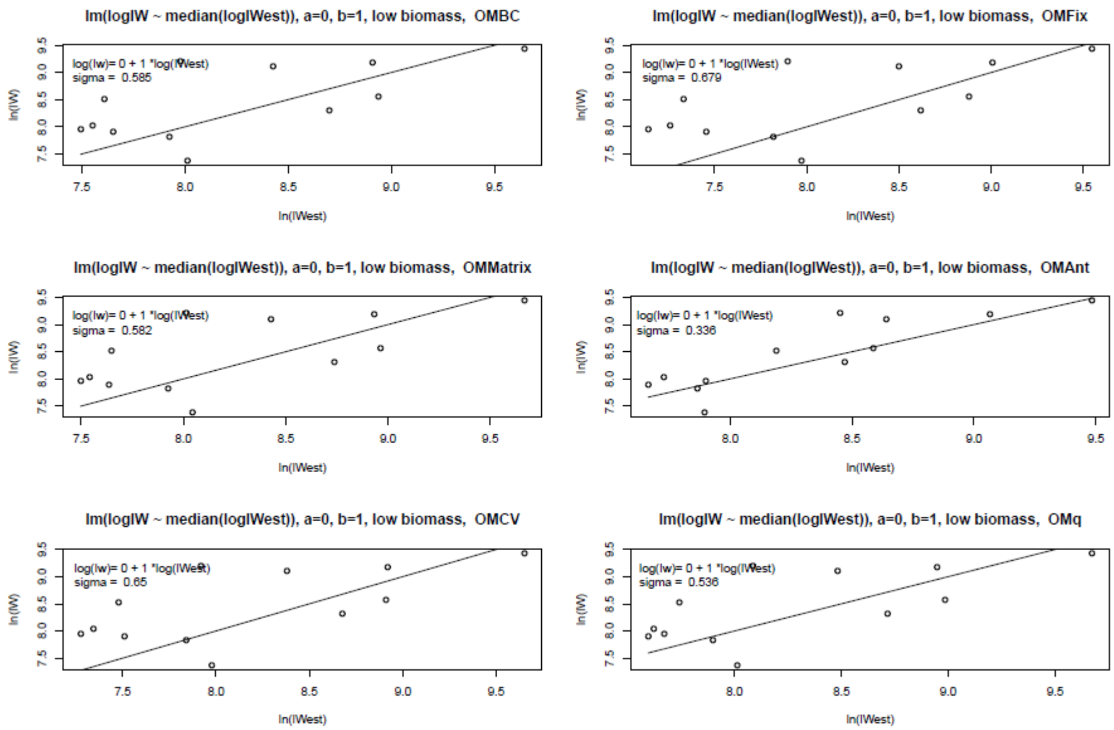


Figure 10. Regression of $IwEst_y$ versus Iw_y for all the OMs assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$ fitting the years in which SSB<B_{lim}.

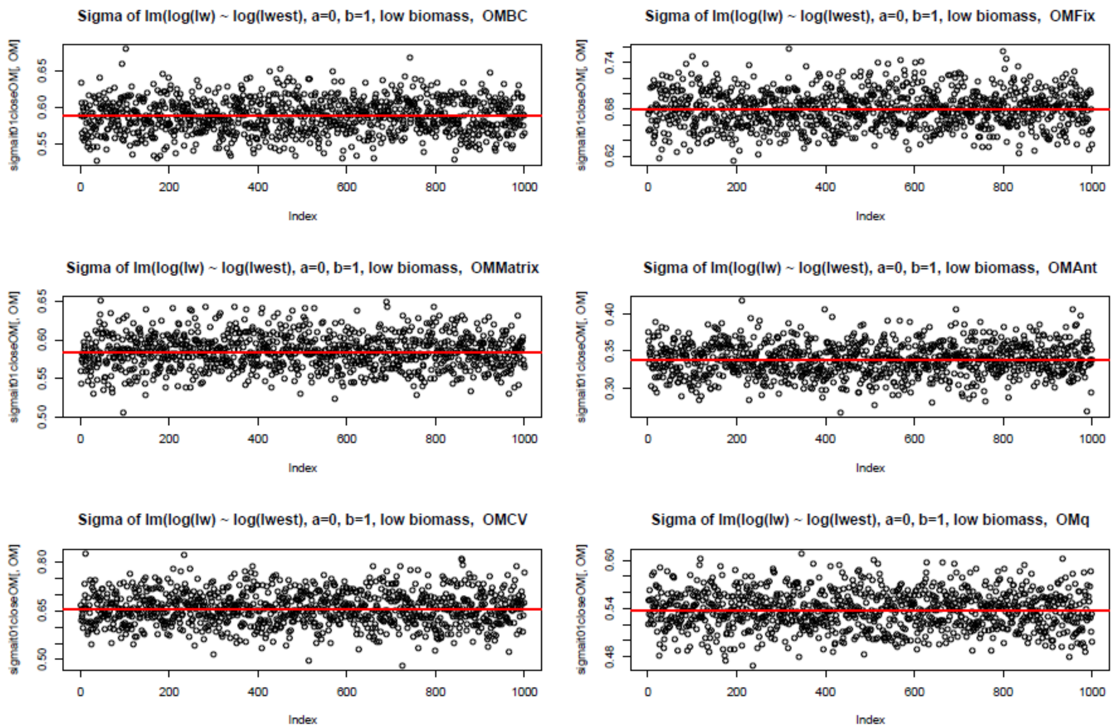


Figure 11. Dispersion of sigma for all the OMs in the second regression model, regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$ fitting the years in which SSB<B_{lim}.

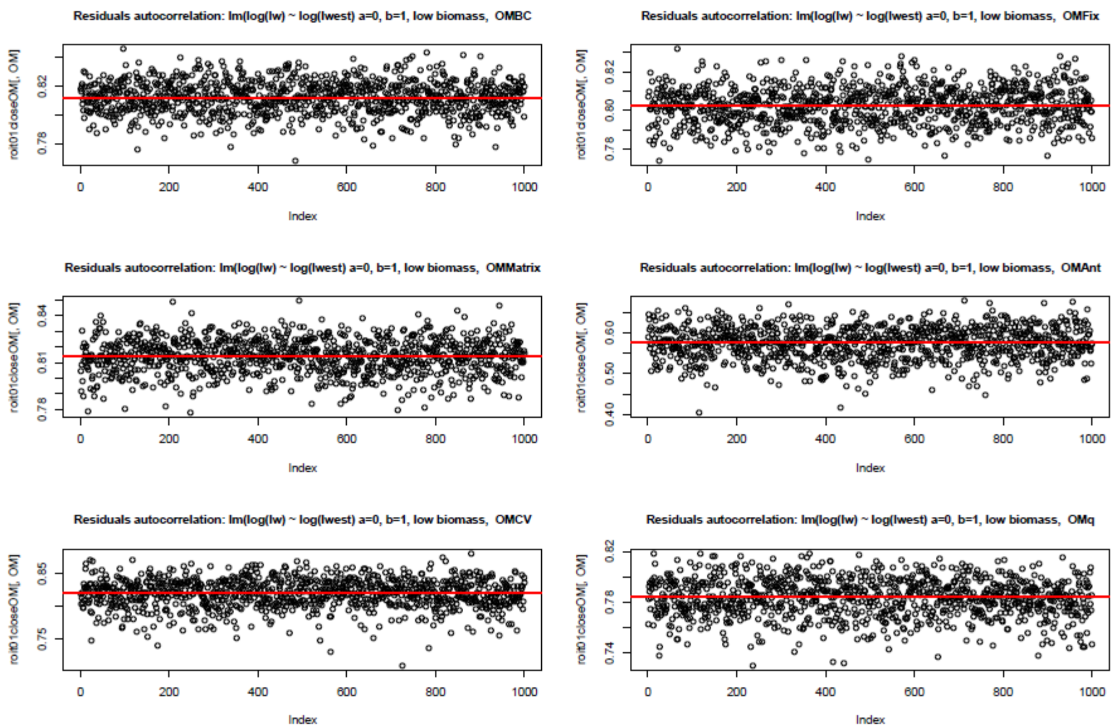
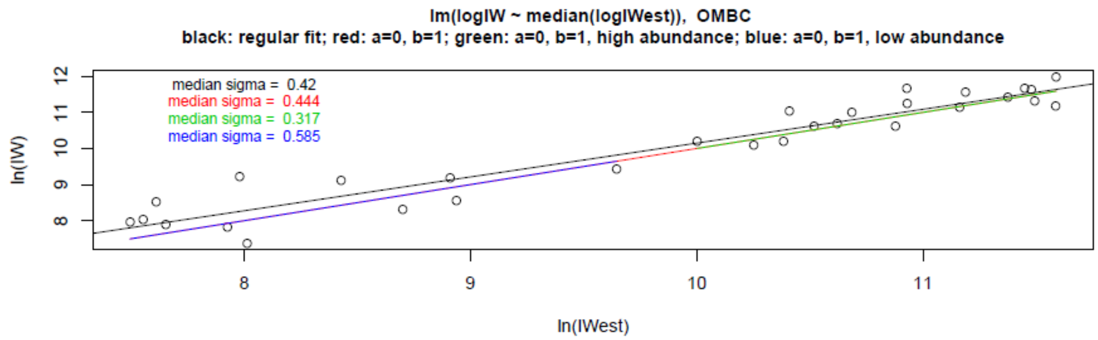
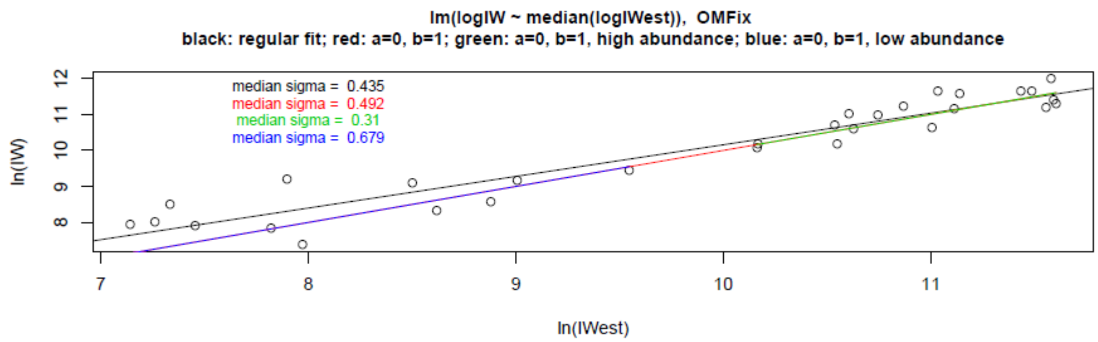
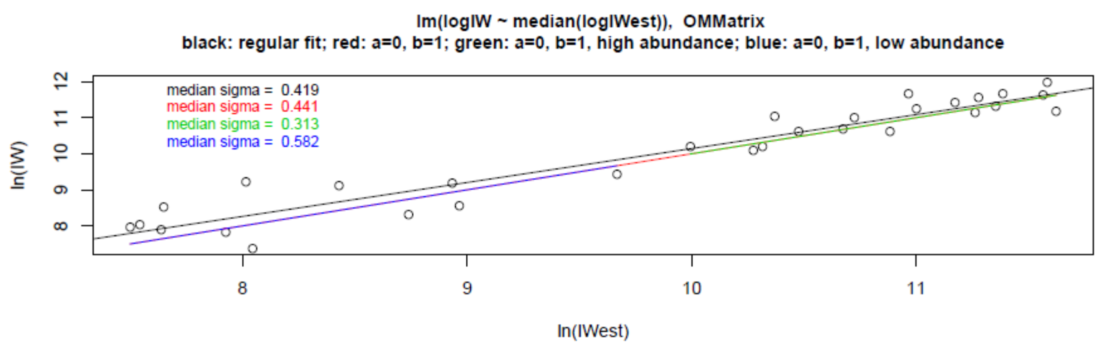
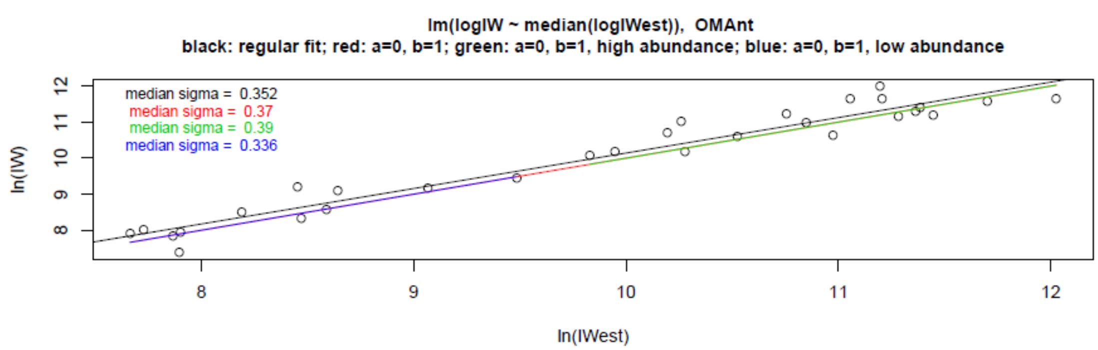


Figure 12. Dispersion of the autocorrelation of the residuals (ϒ) for all the OMs in the second regression model, regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $l\,n(Iw_y) \sim 0 + 1 * l\,n(IwEst_y) + \varepsilon$ fitting the years in which SSB<B_{lim}.


lm(logIW ~ median(logIWest)), OMBC
black: regular fit; red: a=0, b=1; green: a=0, b=1, high abundance; blue: a=0, b=1, low abundance
median sigma = 0.42
median sigma = 0.444
median sigma = 0.317
median sigma = 0.585
ln(IW)
8 9 10 11 12
8 9 10 11
ln(IWest)


lm(logIW ~ median(logIWest)), OMFix
black: regular fit; red: a=0, b=1; green: a=0, b=1, high abundance; blue: a=0, b=1, low abundance
median sigma = 0.435
median sigma = 0.492
median sigma = 0.31
median sigma = 0.679
ln(IW)
8 9 10 11 12
7 8 9 10 11
ln(IWest)


lm(logIW ~ median(logIWest)), OMMatrix
black: regular fit; red: a=0, b=1; green: a=0, b=1, high abundance; blue: a=0, b=1, low abundance
median sigma = 0.419
median sigma = 0.441
median sigma = 0.313
median sigma = 0.582
ln(IW)
8 9 10 11 12
8 9 10 11
ln(IWest)


lm(logIW ~ median(logIWest)), OMAnt
black: regular fit; red: a=0, b=1; green: a=0, b=1, high abundance; blue: a=0, b=1, low abundance
median sigma = 0.352
median sigma = 0.37
median sigma = 0.39
median sigma = 0.336
ln(IW)
8 9 10 11 12
8 9 10 11 12
ln(IWest)

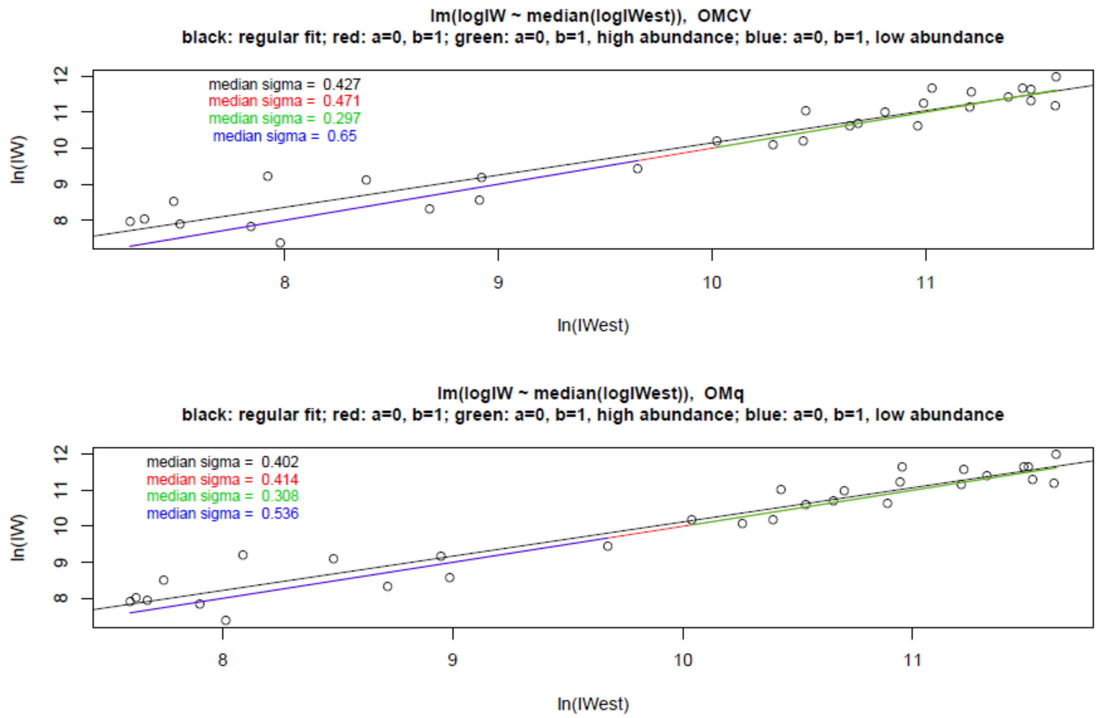


Figure 13. Results of the three regression of $IwEst_y$ versus Iw_y assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(IwEst_y) + \varepsilon$ (regressions 2-4).

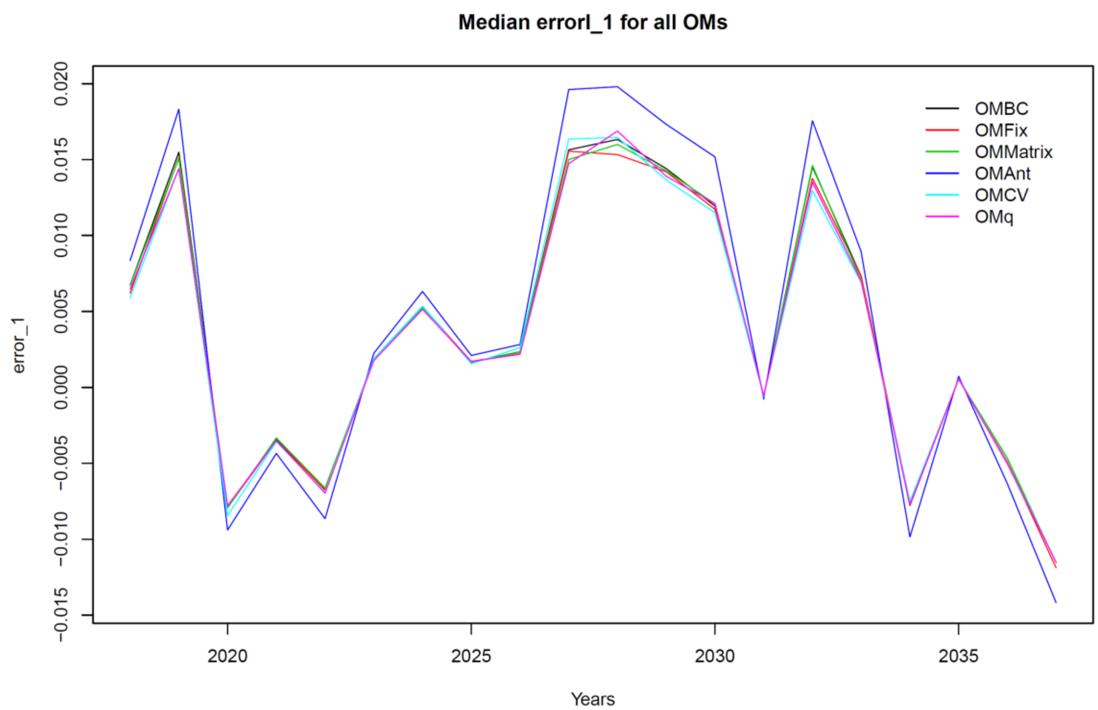


Figure 14. Median of the errors of the indices of the surveys for the method 1: *errorI_1*, by year and OM.

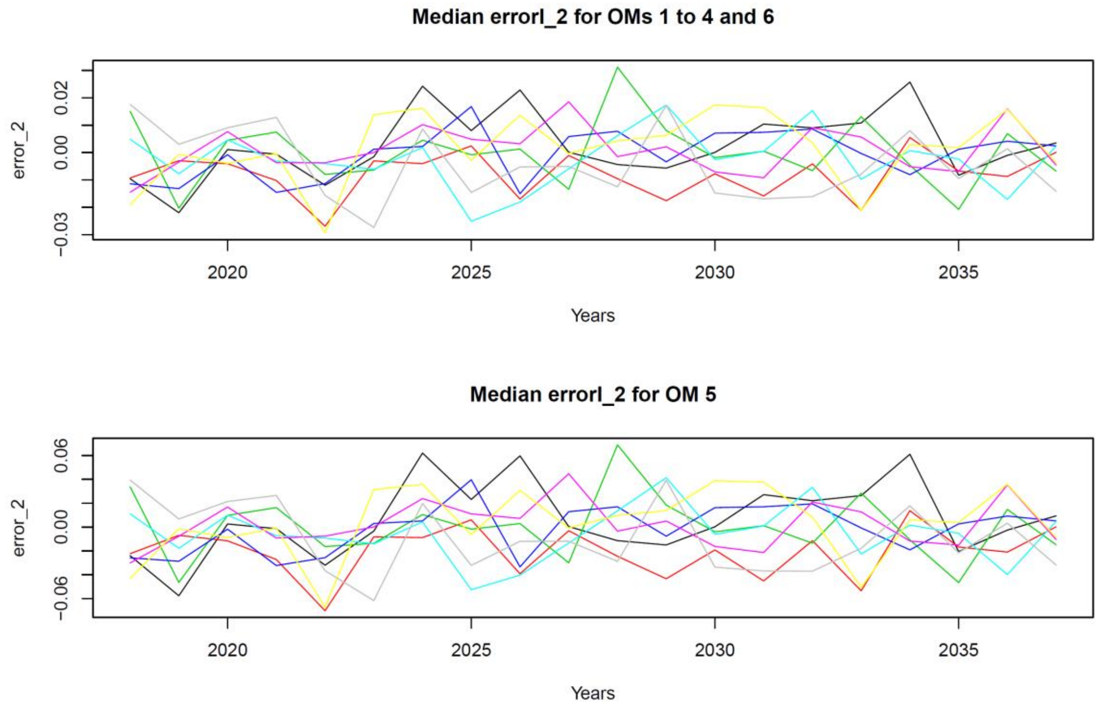


Figure 15. Median of the errors of the indices of the surveys for the method 2: errorI_2, by year, age and OM. For OMs 1 to 4 and 6, the error is the same.

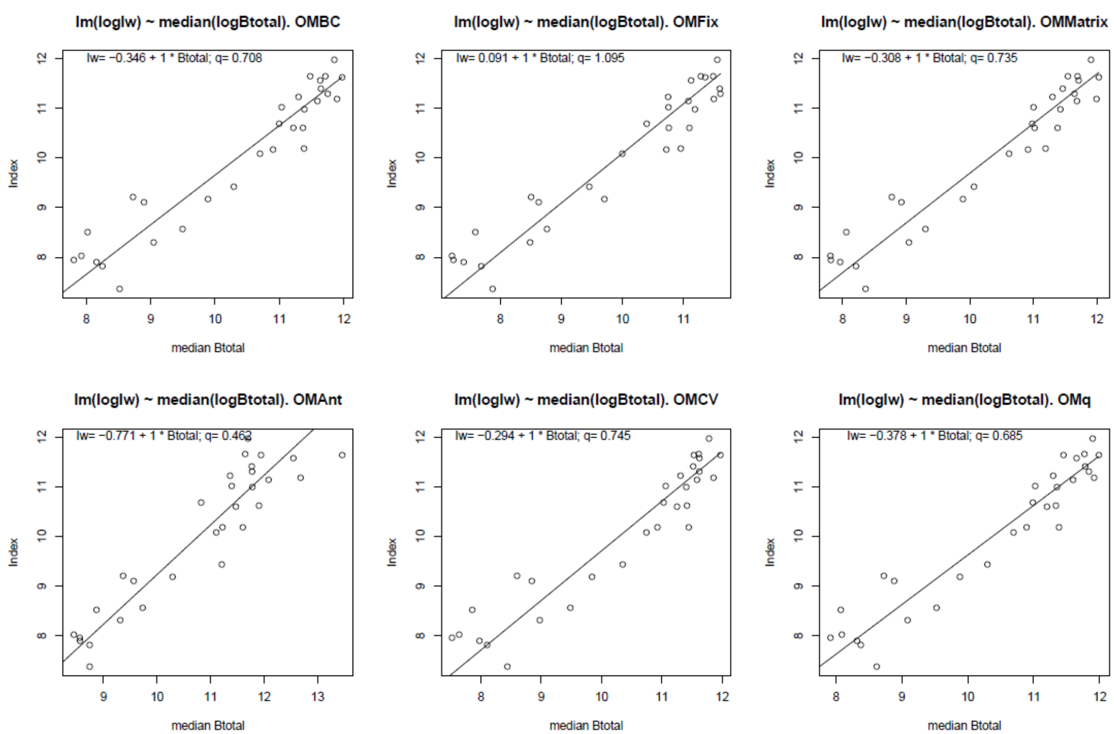


Figure 16. Regression of Iw_y versus $Btotal_y$ for all the OMs assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(Btotal_y) + \varepsilon$.

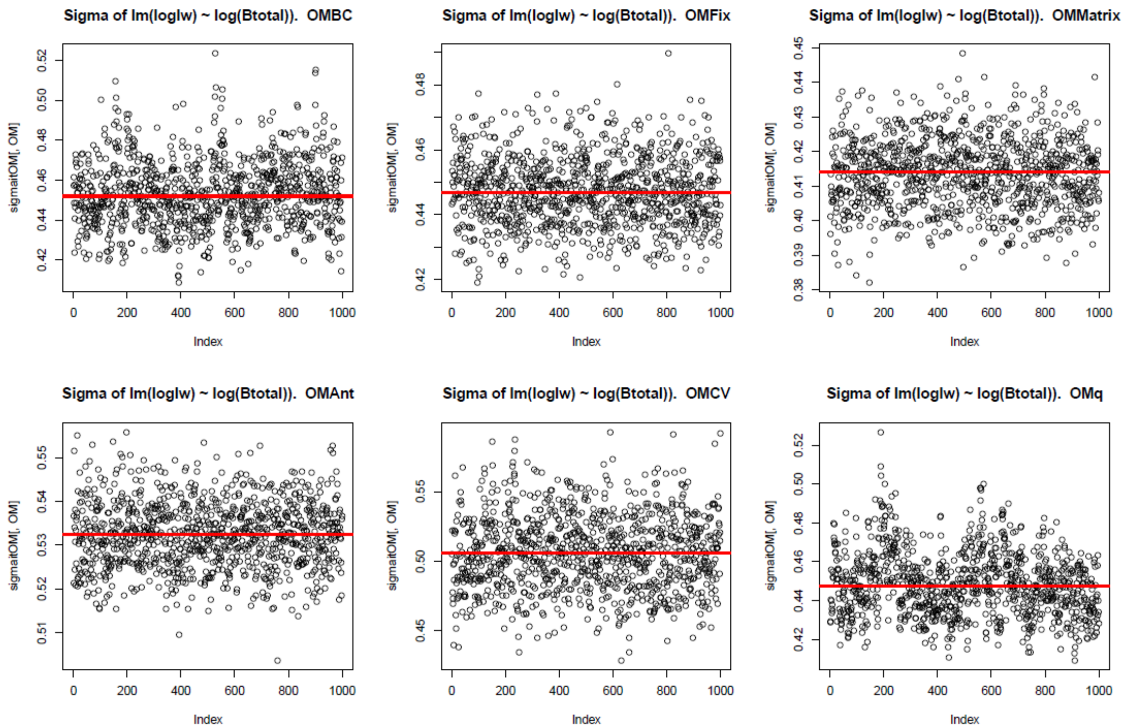


Figure 17. Dispersion of sigma for all the OMs with the third method, regression of Iw_y versus $Btotal_y$ assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(Btotal_y) + \varepsilon$.

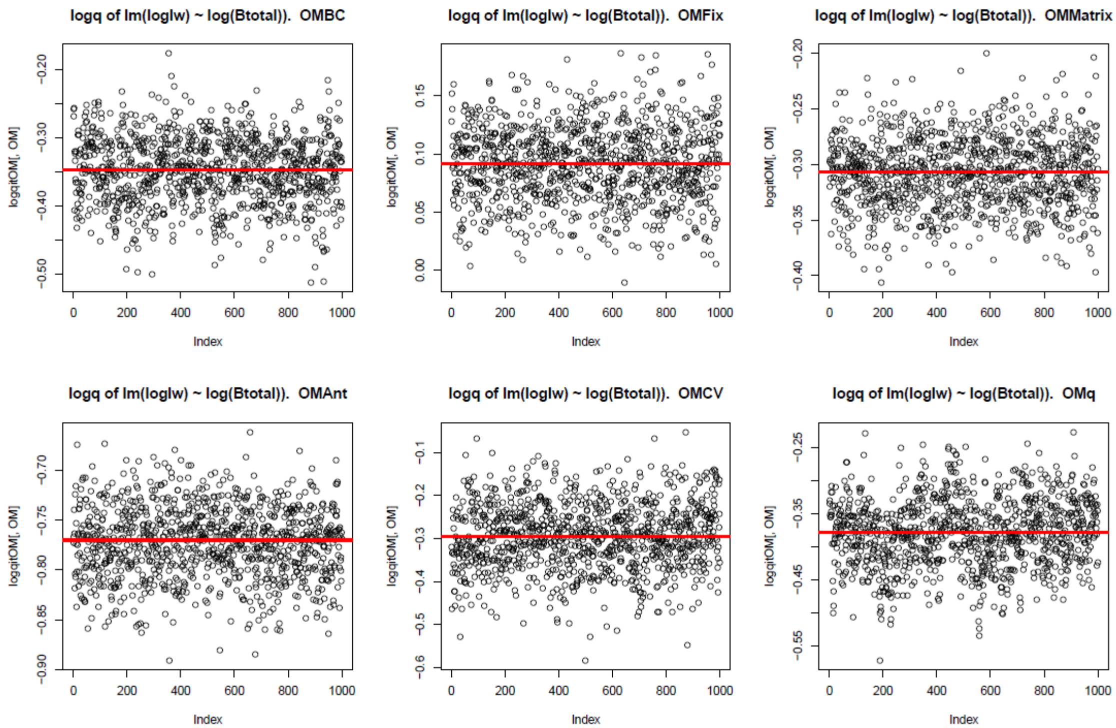


Figure 18. Dispersion of the logarithm of the catchability ($logq$) for all the OMs with the third method, regression of Iw_y versus $Btotal_y$ assuming that the intercept is 0 and the slope is 1: $ln(Iw_y) \sim 0 + 1 * ln(Btotal_y) + \varepsilon$.

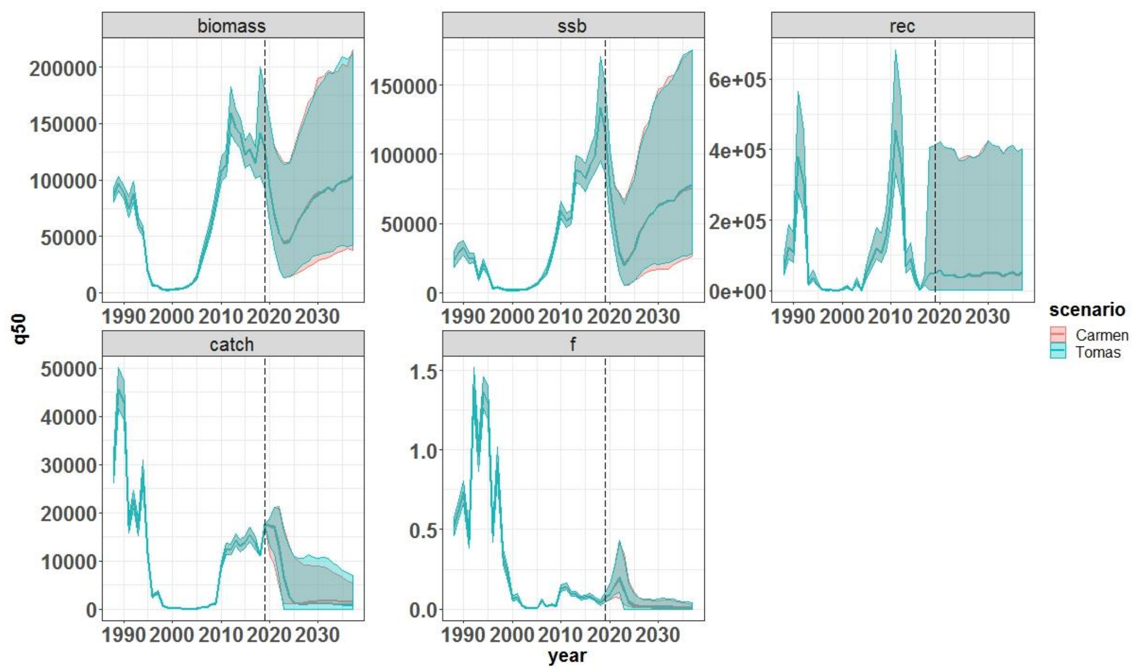


Figure 19. Results of the projection taken both errors **(**$errorI\_1_y$ and $errorI\_2_y^a$) for OM base case, taken as projection bootstrap over 2012-2017 (boot1) and as HCR the Trend-Based one.

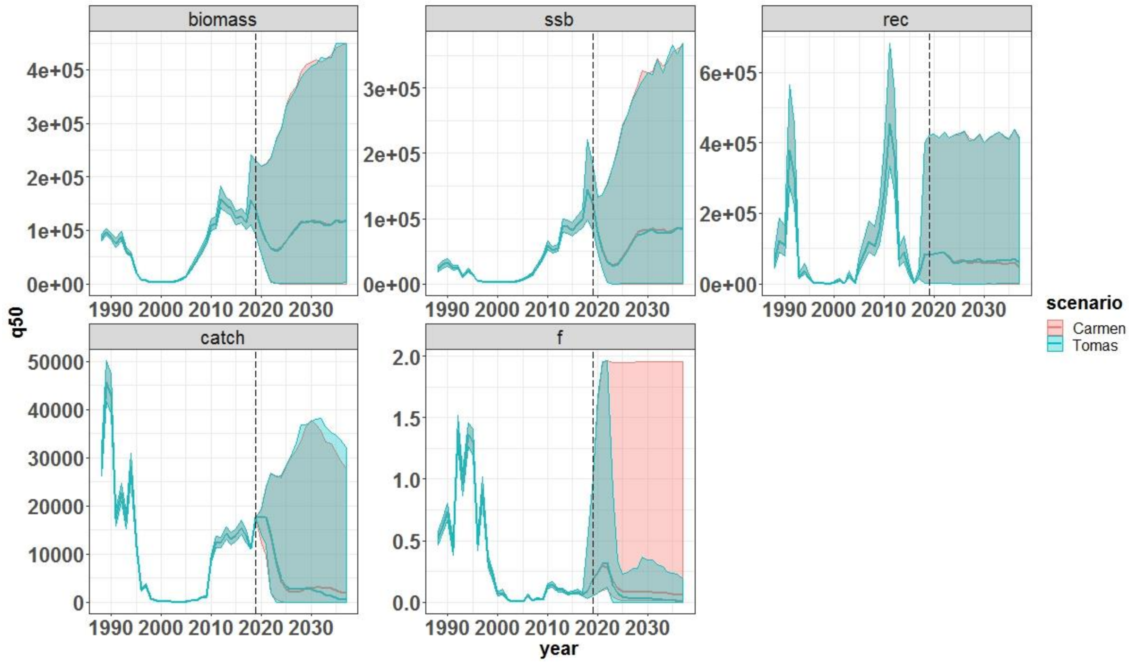


Figure 20. Results of the projection taken both errors ($errorI\_1_y$ and $errorI\_2^a_y$) for OM base case, taken as projection Random Walk (RW) and as HCR the Trend-Based one.