Northwest Atlantic Fisheries Organization

Serial No. N6186 **NAFO SCR Doc. 13/032**

SCIENTIFIC COUNCIL MEETING – JUNE 2013

An index of retrospective pattern in VPA analysis

by

Mónica Mandado and Antonio Vázquez

Instituto de Investigaciones Marinas
Eduardo Cabello 6, 36208 Vigo, Spain

Abstract

KEYWORDS: Retrospective pattern, VPA analysis.

Several aggregation indices are proposed to measure the occurrence of retrospective pattern in VPA. Their behaviour is analysed by simulation. The sensibility of those indices to changes in natural mortality and in survey catchability point them as responsible for pattern, among other possible causes.

Introduction

Several indices have been proposed to quantify similarities in a retrospective analysis (e.g.: Vázquez and Mandado 2010). Retrospective analysis also shows the occurrence of patterns, that is to say, a systematic tendency to overestimate or underestimate previous results. Circumstances that produce a pattern and how to quantify it are two important issues, but they are not yet well understood. The indices for retrospective analysis that we introduce in this paper do not consider how much discrepancies are but how they are distributed: random or grouping. When two sets of the same variable estimates are compared, every position is examined to determine what set contains the highest in each case. The new indices try to quantify the occurrence of zones of the same type: “highest”/”lowest”. Random distribution may correspond with some kind of symmetry: equal power of both types, and should be considered the no pattern case in a retrospective analysis. Contrary to that, a clear pattern should be recognized by the aggregation of types: the occurrence of wide zones with the same type.

Material and methods

Aggregation indices

We introduce the aggregation indices **A1**, **A2** and **A3**, which are designed to measure how much discrepancies in retrospective analyses move away from a random distribution, indicating the presence of a pattern.

These indices are based in the comparison of two sets of data. Let us consider the simplest case, where both sets are p ordered real numbers, named X and Y. A third set Z of p ordered integers can be done with the following forming rule.

$Z_i = 1$, when $X_i > Y_i$
$Z_i = 2$, when $X_i < Y_i$

Groups of consecutive 1s or 2s can be identified. Let g be the number of groups and n_i the size of the i group. The algorithm **a** (in reference to asymmetry):

$$\mathbf{a} = \sum_{i=1,g} (n_i - 1)^2$$

will produce values between cero, when 1s and 2s are alternated and no grouping exist, and $(N-1)^2$ when all are 1s or all are 2s, and N is the set's size. By numerical simulation, we identified mean (m) and standard deviation (s) of this algorithm when 1s and 2s are distributed at random. These statistics are size dependent, let say m_p and s_p for a p sized set. Based on them, an **as** standardized index is proposed as:

$$\mathbf{as} = (\mathbf{a} - m_p) / s_p \qquad m_p = 1.5*p - 3.44 \qquad p = \text{set's size}$$

But this index has not a normal or symmetric distribution pattern, so an alternative index is proposed.

$$\mathbf{A} = \mathrm{sqrt}(\mathbf{a} - m_p) / s_p \qquad \mathrm{sqrt} = \text{squared root}$$

where s_p is now the standard deviation of sqrt(**a** – m_p):

$$s_p = [0.169*\log(p)+1.0591]**6.25$$

The m_p and s_p expressions were determined by simulation, so they only provide reasonable values for discrete cases.

This index has a close to normal distribution, N(0,1). Values higher than 2.5 may indicate some significant (99% prob.) depart from random distribution, and higher values indicate higher departure. Negative values are also possible, and they may indicate that 1s and 2s are more homogeneously mix than expected in a random process. We understand that such homogeneity is a different property that aggregation, and the limit between both is the random distribution. Based on that, negative values of the **A** index are reduced to cero:

$$\mathbf{A} = \max(\mathbf{A}, 0)$$

Because we are only interested in high values, this index is one tail text, and 99% limit is 2.33.

*Lineal aggregation index **A1***

Let us now consider two matrices p×q of real numbers, X and Y, and let us prepare a third one Z with the above forming rule in both dimensions.

$Z_{i,j} = 1$, when $X_{i,j} > Y_{i,j}$
$Z_{i,j} = 2$, when $X_{i,j} < Y_{i,j}$

The proposed **A1** index is calculate as the above linear index of the linearized matrix Z joining rows.

*Orthogonal aggregation index **A2***

A better measure of grouping in the Z matrix could be done in the following routine:

1- Calculate $Z_{i,j}$ as previously stated.
2- Forming all squares where elements $Z_{i,j}$, $Z_{i+1,j}$, $Z_{i,j+1}$, and $Z_{i+1,j+1}$ are all equal.
3- Making groups with the above squares which have at least one element in common.
4- Let g be the number of groups and n_i the size of the i group.
5- Calculate: $\mathbf{a} = \mathrm{sqrt}[\sum_{i=1,g} (n_i - 3)^2]$
6- Let m_{pq} and s_{pq} be the mean and standard deviation of **a** from Z matrices p×q build at random.
7- $\mathbf{A} = (\mathbf{a} - m_{pq}) / s_{pq}$ => we accept: A = N(0,1)
8- **A2** = MAX(**A**,0)

Expressions for m_{pq} and s_{pq} correspond to the random distributed situation, and they were deduced by simulation in discrete cases ($p, q \leq 100$):

$$m_{pq} = 1.6313 * \mathrm{sqrt}(p*q) + 0.721 * [\mathrm{sqrt}(p) + \mathrm{sqrt}(q)] + 1.2694$$

$$s_{pq} = 8.951 – 20 * (1/p + 1/q)$$

*Oblique aggregation index **A3***

Grouping could have an oblique component arising for some common factor for x and y, i.e.: age and years in a retrospective analysis. A new aggrupation index is proposed following the routine:

1- Calculate $Z_{i,j}$ as previously stated.
2- Forming all rhombuses where elements $Z_{i,j}$, $Z_{i,j+1}$, $Z_{i+1,j+1}$, and $Z_{i+1,j+2}$ are all equal.
3- Forming all rhombuses where elements $Z_{i,j}$, $Z_{i+1,j}$, $Z_{i+1,j+1}$, and $Z_{i+2,j+1}$ are all equal.
4- Making groups with the above rhombuses which have at least one element in common.
5- The **A3** index is calculated follow the same steps 4-7 as in the above **A2** index, but using:

$$m_{pq} = 2.9626 * \mathrm{sqrt}(p*q) + 2.28 * [\mathrm{sqrt}(p) + \mathrm{sqrt}(q)] – 1.1592$$

$$s_{pq} = 16.5 + 373.0 / p / q – 70.0 * (1/p + 1/q)$$

Retrospective analysis

In applying these indices to a retrospective analysis we proceeded as follows:

1- Let $N_{0,a,y}$ be a VPA estimation of abundance at age *a* and year *y* (*na* ages and *ny* years).
2- Calculate 5 retrospective peels by reducing the *p* last years, which produce $N_{p,a,y}$ estimations of abundance.
3- Reduce dimension of $N_{p,a,y}$ matrix from 3 to 2 by replacing numbers at age by their sum in a range of ages:

$$N_{p,y} = \sum_{a=a1,a2} N_{p,a,y}$$

4- Calculate Z matrix (ny×5) with the following forming rule:

$Z_{pa} = 1$, when $N_{p,y} > N_{p-1,y}$
$Z_{pa} = 2$, when $N_{p,y} < N_{p-1,y}$
$Z_{pa} = 0$, for non-existing positions

5- A modification to this rule was introduced: $Z_{pa} = 0$ when relative difference of $N_{p,y}$ and $N_{p-1,y}$ was less than 5%.
6- Calculate **A1**, **A2**, and **A3** by considering only 1s and 2s, but not 0s.

Inaccuracy and simulated population

While retrospective indices measure discrepancies among different VPA results, inaccuracy indices were introduce to measure differences between result and the original simulated population (Vázquez and Mandado 2010).

In applying these indices to measure inaccuracy between abundance of the simulated population and VPA abundance estimates, we proceeded as in the general case of comparing to matrices of the same size; we call the result indices **A1***, **A2***, and **A3***. There is a main difference between indices without and with asterisk: the former ones are based on “number of years × number of peels” matrices, while second ones are based on “number of years × number of ages” matrices.

Taking into account that last years are the most important, number of years used to calculate theses indices were reduced to the last 10.

Simulations

Simulations were done as described by Vázquez and Mandado (2010) and Mandado and Vázquez (2011):

Years	1975	2007												
Ages	1	14	(age 14 being a plus group)											
M	0.2													
ages	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*
Partial R	0.0	0.0	0.0	0.1	0.2	0.5	1.0	1.0	0.75	0.5	0.5	0.5	0.5	0.5
Stock weight	0.0	0.1	0.2	0.3	0.4	0.6	0.8	1.1	1.5	1.9	2.6	3.1	3.9	5.0
Maturity	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	1.0
F: min. max.	0.2	0.5												
Survey indices	1990	2007												
q survey 1	0.00	0.00	0.01	0.04	0.10	0.17	0.19	0.18	0.12	0.05	0.03	0.02	0.02	0.02
q survey 2	0.26	0.25	0.22	0.19	0.16	0.12	0.09	0.07	0.05	0.04	0.03	0.02	0.02	0.02

In those papers the effects of random variability of parameters and random error in otoliths ageing on retrospective and inaccuracy indices were analysed. In this paper new sources or variability were considered:

Natural mortality – M
Partial recruitment – pr
Survey catchability – q

A new simulation procedure was now introduced: till now random variability of parameters and random error in otoliths ageing were a constant situation along the simulated series of years; the new parameters (M, pr, q) change from one level to another in a certain year: there is a leap in their value along the time, but they remain constant (or with random variability) before and after the leap; two periods occur along the years in the each formulation. The leap only affects the simulated population, but not to the VPA methods in use.

The above values of natural mortality, partial recruitment and survey catchability were used unless others are indicated.

Natural mortality

In cases where natural mortality was modified, changes only affect the way the simulated population is generated: abundance, catch, and survey indices from one year to the end. Natural mortality remains equal 0.2 for the VPA routines analysing those data. Change was considered in five different cases, with values equal 0.2, 0.3, 0.4, 0.5, and 0.6, from the same year. The year when the leap is produced is indicated.

Partial recruitment

Partial recruitment of the fleet was simulated to allow easy modification and to explore this effect on as a dome shape the indices under study. A dome shape function was selected (see annex) and modification produced displacement of the whole curve to younger ages (Figure 1). Formulation 1 has not leap and the recruitment function is the red line in Figure 1. All other formulations leap from this recruitment function to any of the left side lines. Tested leap years were 2000 and 2003.

Survey catchability

The same dome shape function was used to simulate a catchability function for each simulated fleets (Figure 2). Changes in catchability were then explored as in the case of partial recruitment. Formulation 1 has not leap and the catchability function is the red line in Figure 2. All other formulations leap from this catchability function to any of the right side lines. Tested leap years were 1994, 1997, 2000 and 2003.

One thousand cases were considered in each formulation, with some set of defined parameters. Formulations are grouped in series where only one parameter changes. The behaviour on the indices is then analysed. Previously described retrospective and inaccuracy indices were also calculated jointly the new ones introduced here in order to check their behaviour in the new conditions.

The VPA methods in use were ADAPT and XSA. Two formulations of ADAPT were considered: ADAPT-10, where each cohort present in the last year had a parameter, the survivors, and ADAPT-9, with the same scheme, but last age is excluded (because mortality at that age is a function of the other old ages).

Figure 3 shows the agreement between some aggregation indices.

Results

1- On the behaviour of former retrospective and inaccuracy indices and the new aggregation indices in a random environment.

Simulated cases are generated with a partial recruitment that varies at random, and catch at age and survey indices differ from their simulated ‘exact’ figures in a random way. Standard deviation of partial recruitment (sdPR), and coefficients of variation of catch at age (cvCA) and survey indices (cvSI) have the same value en each of the five cases: 0.0001, 0.1, 0.2, 0.5, and 1.0 (Figure 4). The figures are equivalent to Figure 1 in Vázquez and Mandado (2010).

All retrospective indices have a good behaviour. The aggregation indices **A** are insensible to the proposed random variability.

Inaccuracy indices also have a good behaviour, and XSA produce the best fits between the results and the simulated population. The aggregation indices **An*** reach an almost constant high value from the 0.2 level of random variability.

2- On the behaviour of indices with otoliths sample size.

Simulation includes a “stratified strategy sampling otoliths” with 5, 10, 20, 50, and 100 otoliths for length class, and 5% ageing mistake level (Mandado and Vázquez 2011), random variability was introduced by making sdPR, cvCA and cvSI equal 0.2, and the same variability was included as a background in all formulations. Results are presented in Figure 5, which are equivalent to Figure 2 in Mandado and Vázquez (2011).

All retrospective indices have a good behaviour, and XSA produce the smallest values. The **A1** index produces an unexplainable figure, and indices **A2** and **A3** are insensible to the proposed random variability plus otoliths sample size.

Inaccuracy indices also have a good behaviour; it means that they start in cero, when variability is null, and they grow as variability increases. Besides, XSA produce the best fits between the results and the simulated population. The aggregation indices have a high level due to random variability but they do no change with otoliths sample size.

3- On the behaviour of indices with a leap in natural mortality.

First formulation does not contain change in M. All other formulation have a leap that occurs in 2002, changing from M=0.2 to 0.3, 0.4, 0.5, and 0.6 (Figure 6).

All retrospective are sensible to the leap in M, with small differences among VPA formulations. The aggregation indices **A** are sensible when M changes to more than 0.3, and XSA are affected the most.

All inaccuracy indices are sensible to the leap, and XSA produces small values (best fit). Aggregation indices **A1***, **A2***, and **A3*** increase as new M increase, and XSA produce the highest values (highest aggregation, highest retrospective).

Because of the good behaviour of the aggregation indices in this case, the relationship among them was explored. Figure 3 shows that A2 and A3 measure the same, as well A2* and A3*.

4- On the behaviour of indices with a leap in partial recruitment.

The five formulations contain the five partial recruitment relationships presented in Figure 1, from right to left. Leap years tested were 2000 (Figure 7) and 2003; results are quite similar. No sensibility was observed in anyone of the indices: changes in partial recruitment were undetectable with them.

5- On the behaviour of indices with a leap in survey catchability.

The five formulations contain the five catchability relationships presented in Figure 2 for each survey, from left to right. Tested leap years were: 1994, 1997, 2000 (Figure 8), and 2003, and results are quite similar in all cases.

Retrospective indices are sensible to these changes, and maximum effect occurs in 1997 and 2000, but it is more reduced in 1994 and 2003. XSA produces the less sensible indices.

Aggregation indices **A** are quite sensible with ADAPT, but not with XSA.

Inaccuracy indices are also sensible to these changes, and XSA is the less affected. Maximum occurs in 2000 in all formulations.

Aggregation indices A1*, A2* and A3* show a common behaviour, with an almost uniform increase with the importance of the leap.

Results could be summarized in the following table:

Source of disagreement	retrospective		inaccuracy	
	degree	aggregation	degree	aggregation
random	+	no	+	no
sample size	+	no	+	no
M	+	+	+	+
pr	no	no	no	no
q	+	+ ADAPT no XSA	+	+

“degree” refers to the retrospective or inaccuracy indices. “aggregation” is **An** or **An*** depending of the case.

Results from a retrospective analysis of Greenland halibut if 2+2JKLMN (Healey *et al.* 2010) were used to test these new indices on a real case. Figure 9 shows the results’ plot (they are no the final version). Aggregation indices were calculated as:

	A1	**A2**	**A3**
5+ biomass	0	0	0
recruits	0.05	1.58	0

They indicate a limited departure from the proposed reference: aggregation produced at random. Departure is low in recruits, but it is null in 5+ biomass. This is coherent with the shape of the plots in Figure 9, and also with the fact that 5+biomass, being the sum of several results, are more stable that single magnitudes. As it was pointed out in describing the retrospective analysis, only relative differences greater than 5% were considered. When this limit is reduced to 0.5% the results were:

	A1	**A2**	**A3**
5+ biomass	3.28	3.26	2.43
recruits	1.69	3.69	2.14

We conclude that the limit of 5% or 0.5% is crucial, and it requires further analyses.

Discussion

The indices we proposed are an honest attempt to quantify some immensurable property, with is the occurrence of permanent trends in consecutive retrospective peels to overestimate or underestimate abundance figures. The validity of these indices cannot be judged *a priori*, but observing their behaviour, and that is the reason of all described simulations. The indices are proposed as a measure of aggregation, while retrospective and inaccuracy indices measure the importance of the deviates, so they try to measure complementary properties.

In general, the three aggregation indices **A** have a similar behaviour, but **A2** and **A3** produced values higher than **A1**. All of them have a sound behaviour: they are not sensible to random errors or changes in partial recruitment, but react to changes in natural mortality and survey catchability. The use of XSA instead ADAPT can prevent the effect of changes in survey catchability.

Aggregation indices **A1***, **A2*** and **A3*** have also a similar behaviour, even A1* produce lower values. **A1*** reaches a certain level around 2.0 and **A3*** and **A4*** round 4.0 due to random errors and, after that, they are unaffected by changes in partial recruitment. All of them are positively affected by changes in natural mortality and survey catchability.

If aggregation indices are real indicators of a retrospective pattern, changes in natural mortality and changes in survey catchability would be two circumstances producing a pattern. Other circumstances could be also important, so this issue requires further analysis.

References

Healey, B.P., J.-C. Mahé and M.J. Morgan – 2010. An Assessment of Greenland Halibut *(Reinhardtius hippoglossoides)* in NAFO Subarea 2 and Divisions 3KLMNO. *NAFO SCR Doc.* 10/40.

Mandado, M. and A. Vázquez – 2011. On otoliths sampling. *NAFO SCR Doc.* 11/23.

Vázquez, A. and M. Mandado – 2010. Random Retrospective Pattern in Fish Stock Assessment. *NAFO SCR Doc.* 10/36.

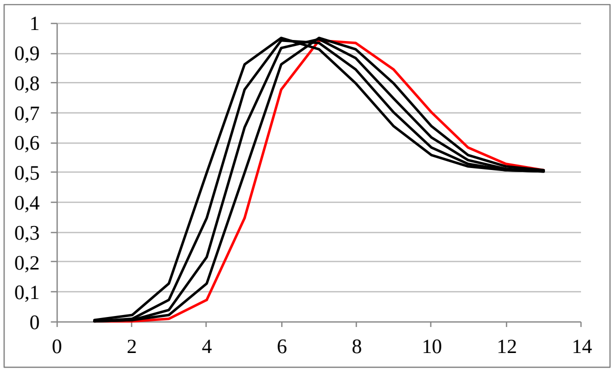


Figure 1 – Partial recruitment at age. The original values are those of the red line. Change occurs once in each case, and partial recruitment takes the values of any of the left side lines.

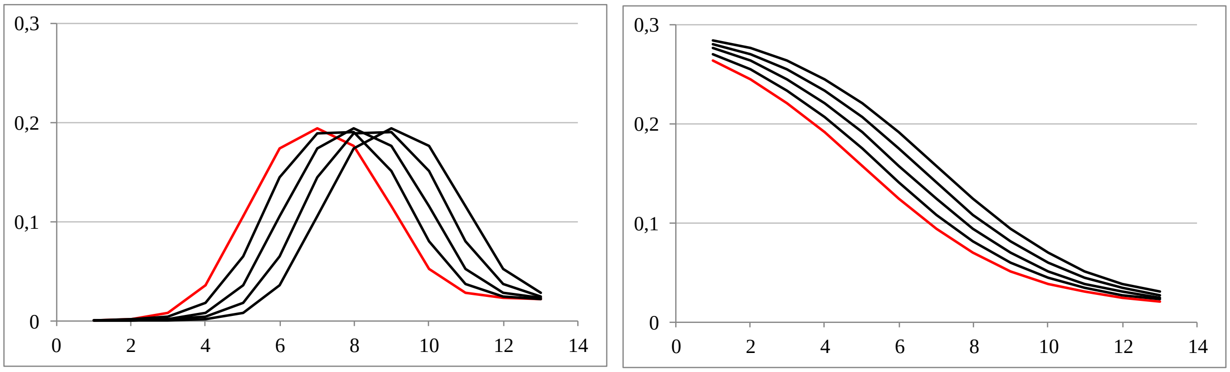


Figure 2 – Catchability at age of the two surveys. The original values are those of the red line of each survey. Change occurs once in each case, and catchability takes the values of any of the right side lines.

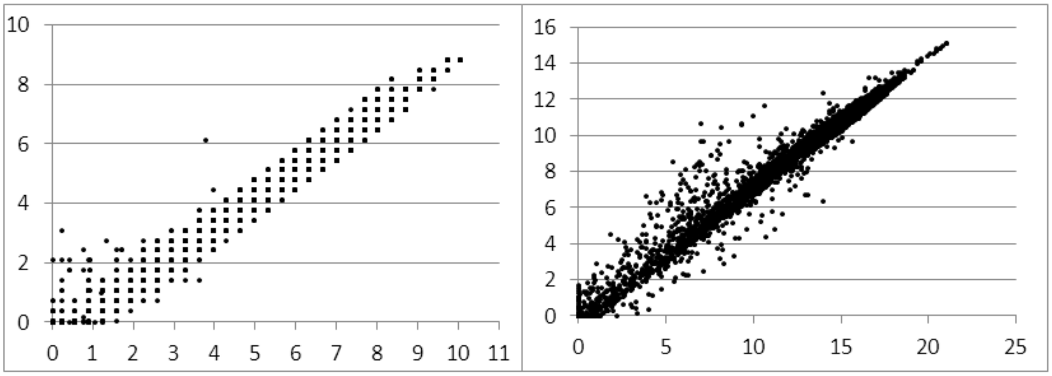


Figure 3 – Relationship between **A2-A3** (left) and **A2*-A3*** (right) when changes in natural mortality are simulated.

Figure 4 – Retrospective (left) and inaccuracy (right) indices: random variability in   sdPR, cvCA and cvSI at the level indicated on x axis.

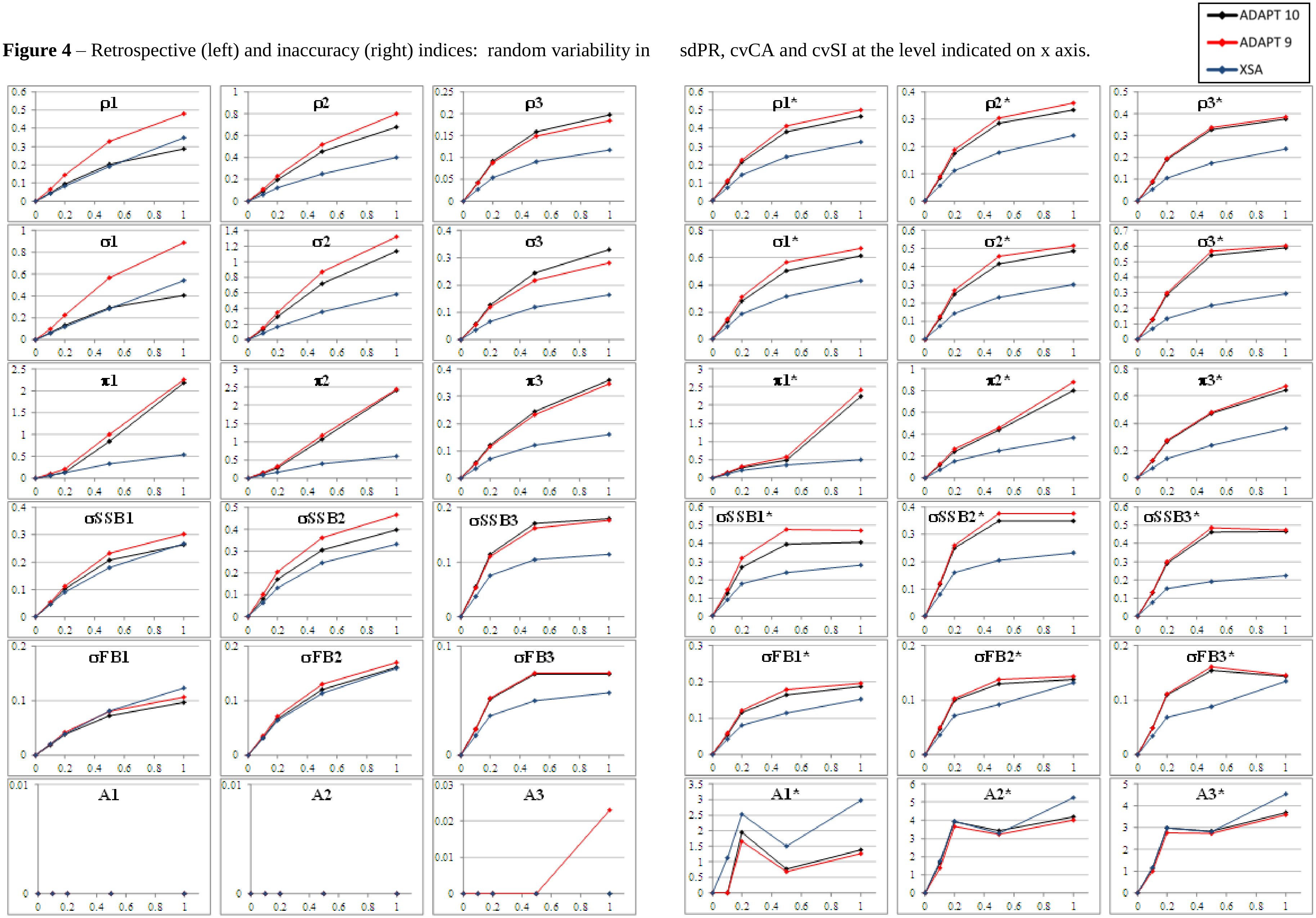


Figure 5 – Retrospective (left) and inaccuracy (right) indices: otoliths sample size (as indicated in x axis) and 5% ageing mistakes.

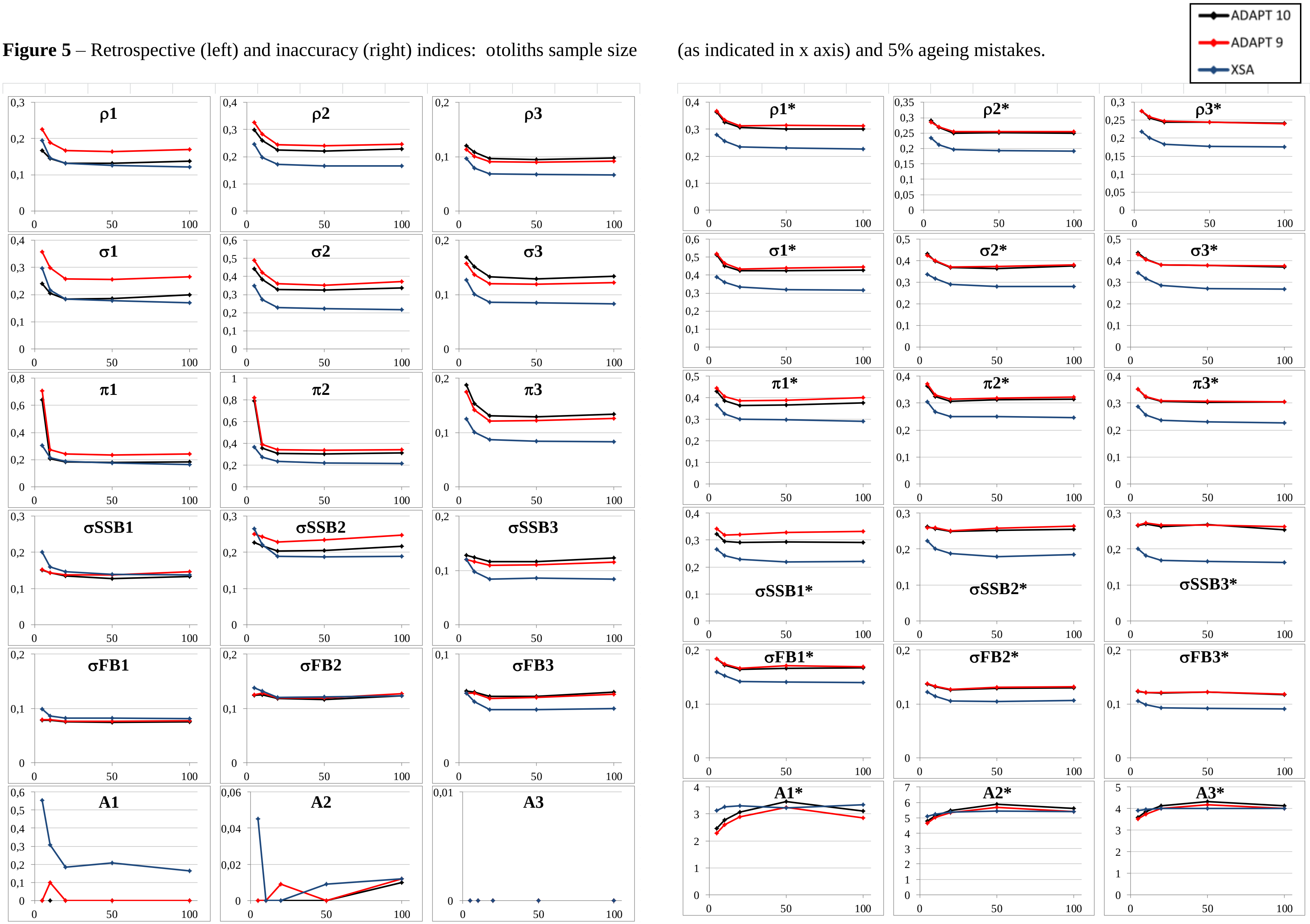


Figure 6 – Retrospective (left) and inaccuracy (right) indices: leap of natural mortality in 2002.

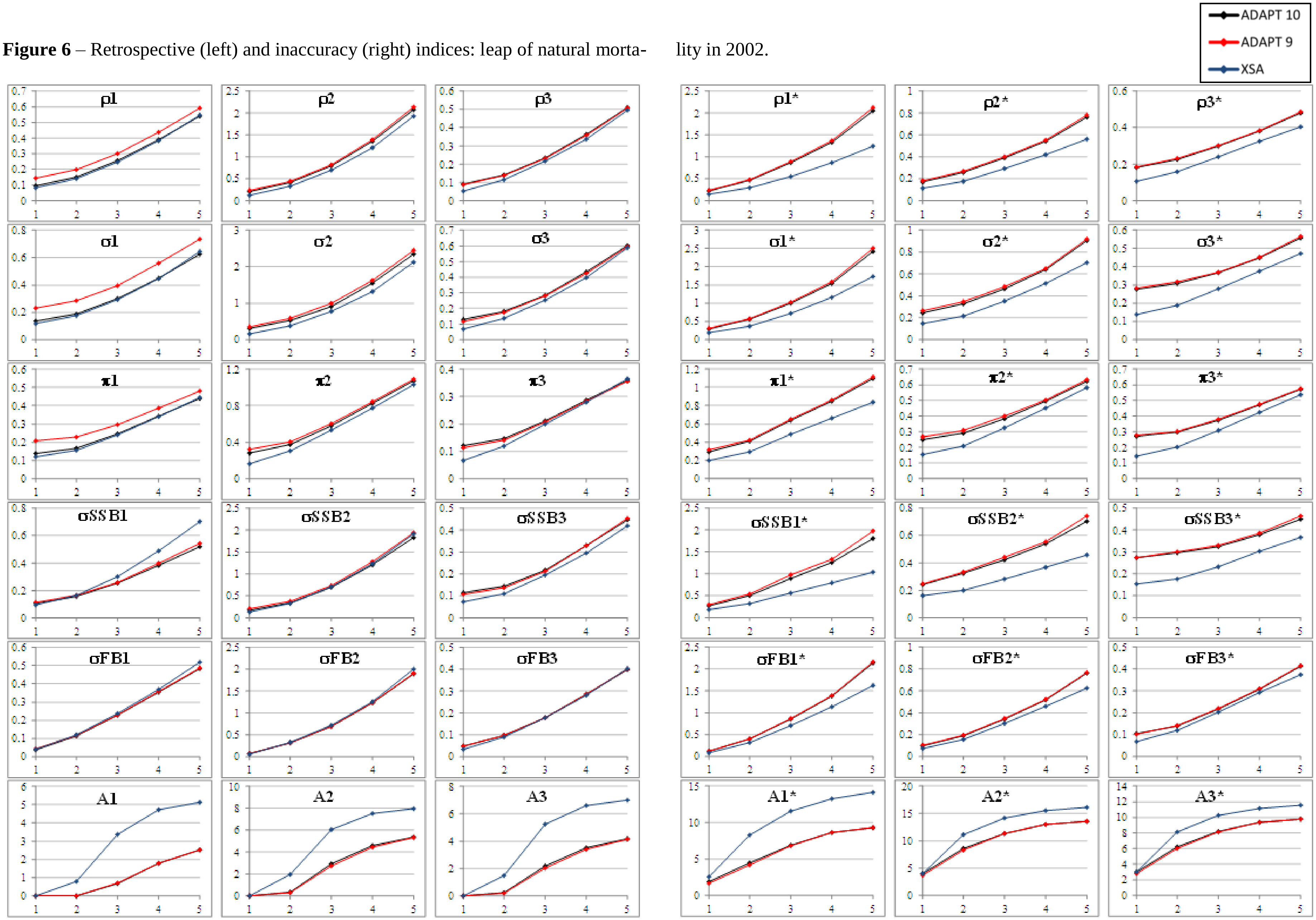


Figure 7 – Retrospective (left) and inaccuracy (right) indices: leap of partial recruitment in 2000.

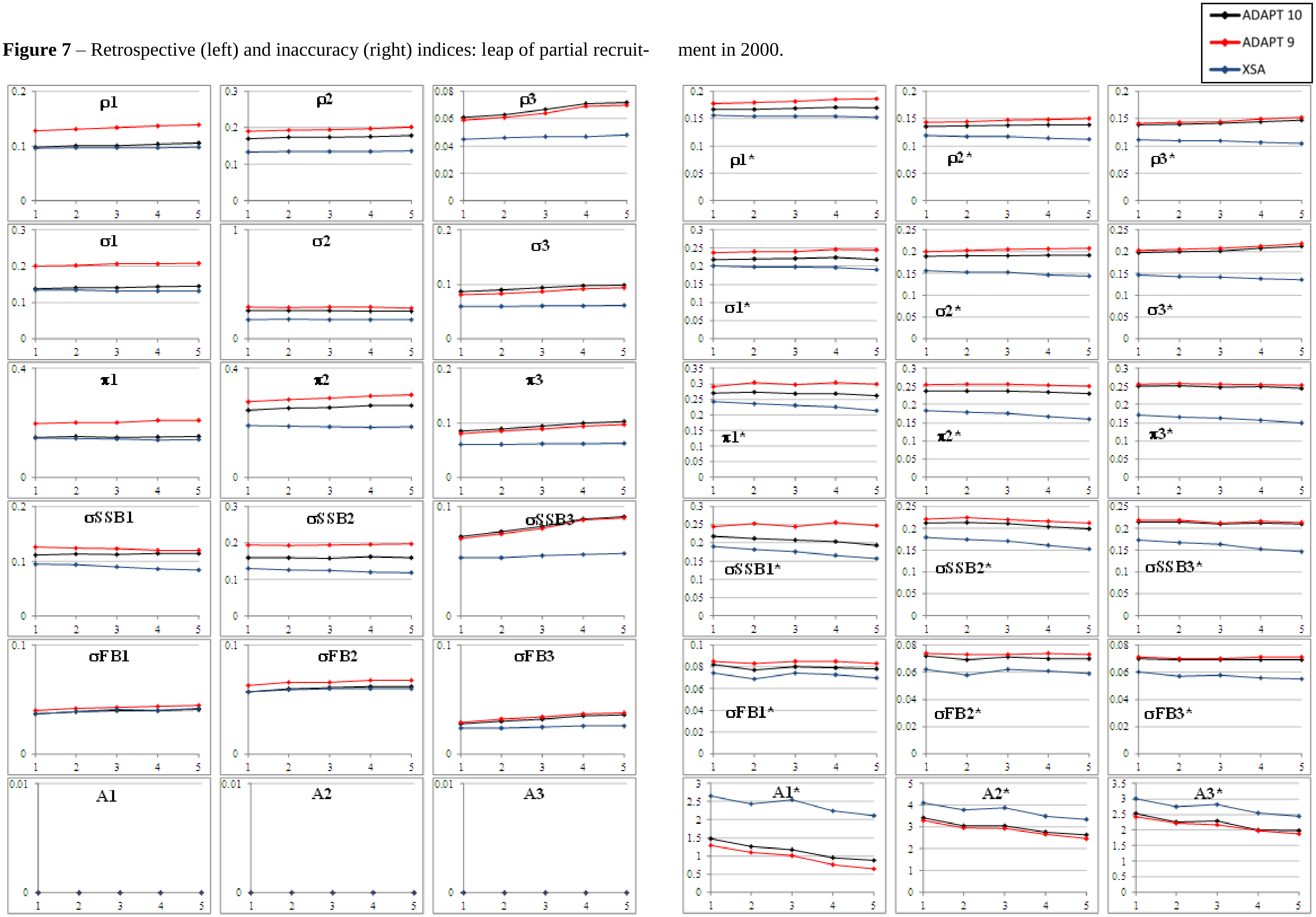


Figure 8 – Retrospective (left) and inaccuracy (right) indices: leap of survey catchability in 2000.

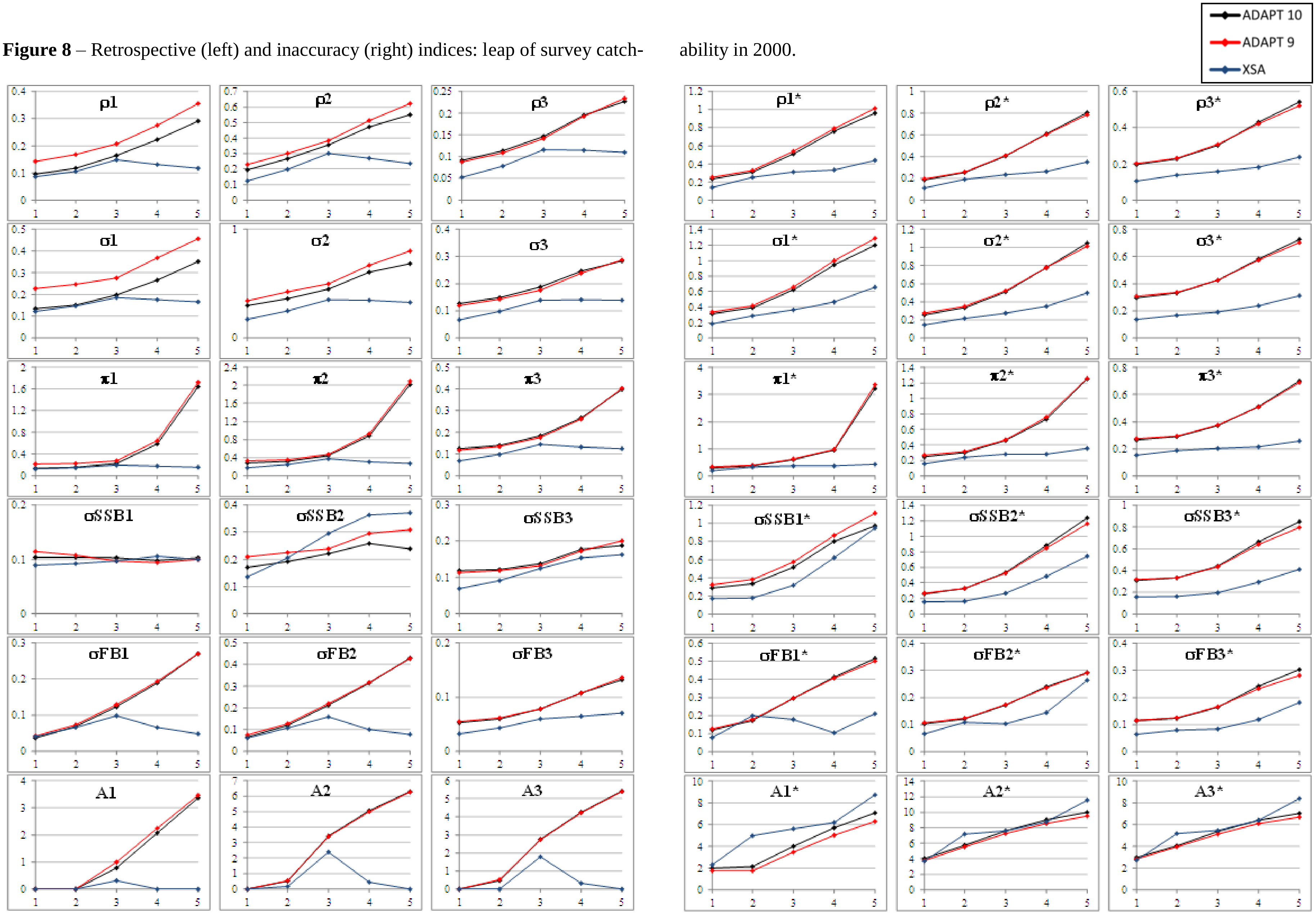
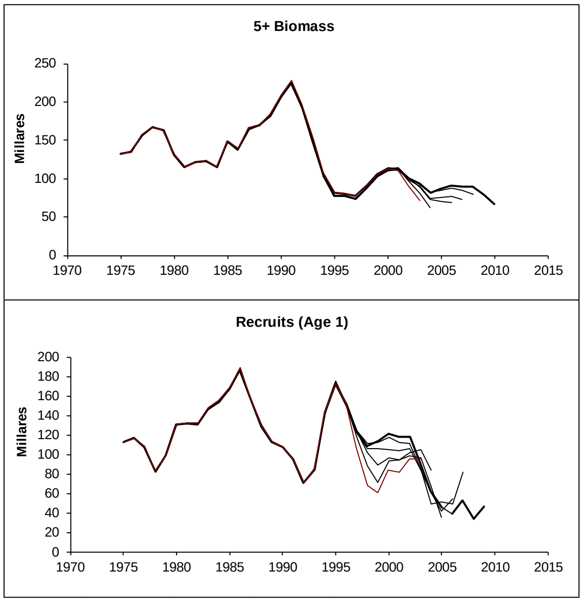


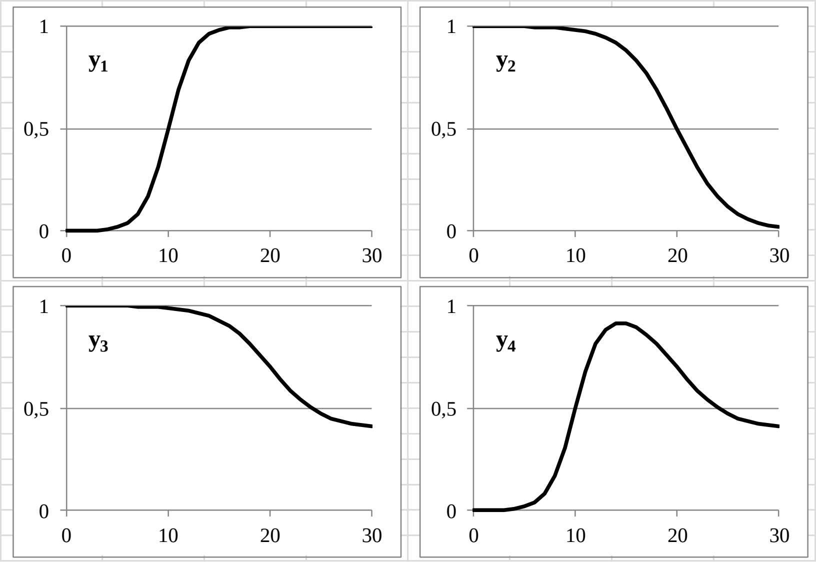
Figure 9 – Retrospective results of the 2+3JKLMNO Greenland halibut (Healey *et al*. 2010) (current plots are the final ones).


5+ Biomass

Recruits (Age 1)

ANNEX - From a logistic function to a dome shape in three steps:

0: logistic equation	$y_1 = k / [1 + EXP(4*p_1*(B-x))]$	$k=1, B=10, p_1=0.2$
1: symmetry	$y_2 = k / [1 + EXP(4* p_2*(x-B-C))]$	$k=1, B=10, C=10, p_2=0.1$
2: reduced	$y_3 = D * y_2 – D + 1$	$D = 0.6$
3: dome shape curve	$y_4 = y_1* y_3$	



Note that: k = the highest value (equal 1 in this case)
B = x value for y=0.5
C = x difference between y=0.5 of $\mathbf{y_1}$ and $\mathbf{y_2}$
p_1 and p_2 = maximum slope of $\mathbf{y_1}$ and $\mathbf{y_2}$
D = shrink factor for $\mathbf{y_3}$