SAW 47 Working Paper 4 (TOR 2b) – Zero Filling

8 November 2006

Simulation Studies of Issues Associated with Filling Zeros in VPA Tuning Indices

by

Chris Legault and Al Seaver

Introduction

Recently, the 2006 assessment of summer flounder (Terceiro 2006) was subject to a NMFS Office of Science and Technology (S&T) Peer Review (Methot 2006). Among the recommendations made by the S&T Peer Review panel was the following:

The Panel finds that one immediate modification of the VPA is justifiable and reduces the retrospective pattern in stock size during 2003-2005. The VPA model currently treats survey observations of zero as missing values. An observation of zero for a particular age of fish in a particular survey year does not mean that there are no fish of that age in the stock, only that the number of survey samples was not sufficient to detect any fish of that age. This VPA model, as with most assessment models, tunes to the logarithm of the survey observations so cannot explicitly deal with observations of zero. However, treating these zeroes as missing values can result in a bias because time periods of low abundance are underrepresented in the data input to the assessment model. In the case of summer flounder, the result may be an underestimate of the degree to which the stock has rebuilt since the low levels that occurred around 1990. The committee did not discuss this issue during the Sept 14-15 meeting, so is not prepared to present a definitive solution. An interim approach would use a small value in place of the zeroes. A value equal to one sixth of the smallest observed positive value would be reasonable until a more complete statistical solution can be developed.

This recommendation departs from the standard practice in NEFSC assessments of treating zero values in tuning indices as missing values. To more fully understand the implications of this suggested change, two types of simulation analyses were conducted. The first is a simple spreadsheet example of how a single time series is impacted by different levels of fish detection and the implications for a full VPA. The second is a full simulation that generates many random sets of data for VPA from a known case, creates zeros for some of the indices in some years, and compares different methods for dealing with these zeros, including treating them as missing values, replacing the zeros with a fixed small value, and the one sixth of the smallest observation rule.

First Study: Impact of Zeros on One Time Series

A population that declined and then increased was created artificially. A catchability coefficient was applied to generate a survey time series exactly from the data. The values in the time series were rounded to two, one, and zero decimal places creating observations of zero for 2, 4, and 7 years, respectively (Table 1). A series of constants was added to the time series ranging from 0.0001 to 10 so that the holes were filled. A new catchability coefficient was calculated that minimized the difference between the true population and the observed survey time series which had been modified to fill the holes. This was done to show how a model would need to change the predicted values to more closely match the observed series. In this study, treating the index values as missing results in an exact match between the observed and predicted values, due to the formulation of the problem and so are not considered further.

The differences between observed and predicted values depend strongly on the constant added to the time series (Figs 1-3). Adding a large value, such as 10, causes the survey time series to flatten relative to the true population. A model would try to reduce the change in the population in this case. Conversely, adding a very small value, such as 0.0001, causes the survey time series to exhibit a stronger decline and recovery than the true population. In this case, a model would try to increase the changes in the population. Adding one sixth of the minimum observed value appears to be an objective way to determine a value that is not too big or too small for the round 2 case where only two zeros are replaced.

However, the more disturbing result seen in these simulations is that the addition of a constant value to replace the zeros in a survey time series artificially imposes a pattern that may not match the actual pattern in the population. This is most clearly seen in the round 0 case where seven zeros are filled with the same value even though the true population declines then increases during the seven year period.

Second Study: Simulation Analysis of Different Methods of Treating Zeros

A comparative study was performed using the POPCOMP length based population simulator tool and VPA version 2.3.3. The objective was to examine the effects of using indices of abundance with some portion of the index data treated as missing or alternatively replaced with an imputed value. Four scenarios were examined. In each case the simulated data were sampled to create 100 realizations of VPA input data and the results of the multiple realizations were compared in their ability to recover the true stock numbers and fishing mortality at age. The test was performed in such a manner as the VPA files created for each realization would be the same for each scenario except in the specified removal and alternative replacement of index data based on an input cut point.

The simulated population was loosely based on the summer flounder assessment with the population initially declining due to high F (>2) and then rebuilding as F was lowered to <0.5). The simulated population spans 24 years starting in 1982 and consisting of 8 age classes with the last age class acting as a plus group. Natural mortality was 0.2 for all ages and years. Both recruitment and fishing mortality vary widely over the time series. The growth projection matrix was created using von Bertalanffy growth coefficients and length bins ranging from 10 to 84 cm. A logistic equation was used for fishery selectivity at length. Catch was removed from the population based on the true F but samples were collected from four market categories based on size (sample sizes 65-133 per 100 metric tons) to introduce variability in the catch at length. Age-length keys were created based on sampling 25% of the observed lengths and an ageing error matrix was included to introduce variability in the catch at age (mis-aged proportions ranged from 4% to 17%). The length-weight equation coefficients supplied to allow expansion of sampled catch to total landings, which had a small amount of variability relative to the true landings (CV = 0.01). Discards were not included in this simulation. This level of uncertainty in the catch at age is thought to be representative of the level associated with the summer flounder assessment. However, there is not a retrospective pattern observed when the simulated data are analyzed with VPA, so not all sources of uncertainty have been captured.

There was only one index generated for each age. The catchability for each index was chosen so that catchability increased with age (Table 2). The uncertainty was higher for the indices at younger ages than older ages (Table 2). The coefficients of variation were used to generate lognormally distributed error in the observed indices. The population trends, catchability coefficients, and coefficients of variation combined to produce different probabilities that a given index value would fall below 1.0 (Table 3). Index values below 1.0 were treated in four different ways:

- Case 1 - Actual values used
- Case 2 - Replaced with 0.0 and treated as missing
- Case 3 - Replaced with the arbitrary constant 0.01
- Case 4 - Replaced with 0.0 then a constant of 1/6 times the smallest non-zero element in the index vector added to all index vector elements including zeros.

The VPA input files generated for each realization were identical excepting that indices of abundance were altered by case.

The median values of F and N at age from the 100 realizations of the VPA model under the four cases of treating index values below 1.0 were compared with the true values from the simulated population (Tables 4-5 and Figs 4-5). Due to the convergence properties of VPA, the medians from the 4 cases are essentially identical for years 1982-1994, as seen in Figures 4-5, and so are not shown in Tables 4-5. The most striking feature seen in the tables and figures is the poor performance of Case 3 (replacing zeros with the arbitrary constant 0.01). The fishing mortality rates in Case 3 were well below the true values while the estimated population abundances were well above the true values. Case 3 clearly demonstrates the potential for introducing bias by replacing zeros in tuning index time series with an arbitrary constant. While not as clear, generally the Case 4 (add 1/6 of smallest non-zero element) estimates were more biased than the Case 2 (treat zeros as missing) estimates. The exception to this generality is seen in age 1 results where the VPA formulation had to be modified slightly to estimate only ages 3-8 in the terminal year +1 due to the lack of information for age 2 in the terminal year +1 when the index was zero. For older ages, Case 2 actually outperformed Case 1 (all data used) relative to the truth. It is not clear why this happened and may be an artifact of the bias introduced by the mis-ageing matrix used to generate the catch data. However, even if Case 1 is used as the basis for comparison, instead of the true values, Case 2 performs at least as well as Case 4 for all ages except age 1.

Discussion

An alternative method to determining the constant to use in place of zeros that was not considered in this exercise is provided by Berry (1987). The Berry approach consists of finding the constant that minimizes a function of the skewness plus kurtosis of the raw data. This approach is not appropriate for use with tuning index data because the residuals are assumed to follow a lognormal distribution, not the raw observations.

While the 1/6 of the smallest non-zero approach appears to provide reasonable results in some cases, it is an arbitrary rule. In some situations, 1/5 or 1/7 of the smallest non-zero index value would perform better than 1/6. The main problem remains however. Filling zeros with a constant value, no matter how that constant is selected, creates a pattern that may not match reality. These simulations show that this approach can produce results further from the truth than treating zeros as missing values.

Of course, in reality the zeros do have information. Results should be checked to ensure that predicted values are not high when index is zero. If an assessment model predicts high abundance for a year-age combination that had a zero index, the model results should be questioned. However, adding incorrect information arbitrarily has the potential to bias the results, as demonstrated in these simulations.

Conclusions

The two simulation studies have demonstrated problems that can arise when tuning indices with zero values are replaced with arbitrary constants. This practice assumes that the correct magnitude can be chosen to fill the zeros and that it is better to provide the model with information that the index is low rather than treat the data as missing. Results demonstrate that this premise is not always correct. Thus, we recommend the NEFSC standard approach of treating zero values in tuning indices for VPA as missing values.

References Cited

Berry DA. 1987. Logarithmic transformations in ANOVA. Biometrics 43:439-456.

Methot R. 2006. Review of the 2006 Summer Flounder Assessment Update. Chair's Report. NMFS Office of Science and Technology. 6 p.

Terceiro M. 2006. Stock assessment of summer flounder for 2006. NEFSC Ref Doc. 06-17. 119 p.

Table 1. Artificial time series for a population and the associated time series of indices given a catchability of 0.000002 when the values are rounded to two, one, and zero decimal places. Highlighted cells are years when the tuning index has an observed zero.

		Index		
Year	Population	Round 2	Round 1	Round 0
1980	2000000	4.00	4.0	4
1981	1500000	3.00	3.0	3
1982	1300000	2.60	2.6	3
1983	1000000	2.00	2.0	2
1984	500000	1.00	1.0	1
1985	300000	0.60	0.6	1
1986	200000	0.40	0.4	0
1987	10000	0.02	0.0	0
1988	5000	0.01	0.0	0
1989	1000	0.00	0.0	0
1990	2000	0.00	0.0	0
1991	50000	0.10	0.1	0
1992	100000	0.20	0.2	0
1993	300000	0.60	0.6	1
1994	400000	0.80	0.8	1
1995	700000	1.40	1.4	1
1996	1200000	2.40	2.4	2
1997	1500000	3.00	3.0	3
1998	1100000	2.20	2.2	2
1999	1200000	2.40	2.4	2
2000	1700000	3.40	3.4	3

Table 2. Catchability coefficients (q) and coefficients of variation (CV) by age for the tuning indices used in the second study. The q values multiplied the true populations at age to generate the expected values for the indices by year. The CV values describe the amount of lognormally distributed error used to create the random VPA input data.

Param	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
q	0.0000000 3	0.000000 1	0.000000 1	0.0000 1	0.0000 1	0.0000 1	0.0000 1	0.0000 1
CV	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3

Table 3. Probability that an index value will be below 1.0 and thus set to zero given the true population, catchability coefficient, and uncertainty associated with each index and year.

Year	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8
1982	6.9	<1.0	16.4	<1.0	<1.0	6.4	9.5	94.6
1983	4.9	<1.0	2.1	<1.0	<1.0	<1.0	98.5	89.6
1984	28.0	<1.0	4.4	<1.0	<1.0	46.6	>99.0	>99.0
1985	27.8	<1.0	1.2	<1.0	<1.0	<1.0	>99.0	>99.0
1986	21.6	<1.0	10.9	<1.0	<1.0	<1.0	85.8	>99.0
1987	35.5	<1.0	14.0	<1.0	<1.0	<1.0	63.4	>99.0
1988	98.6	1.0	5.8	<1.0	<1.0	<1.0	6.6	>99.0
1989	73.8	64.1	26.5	<1.0	<1.0	<1.0	58.5	>99.0
1990	65.9	11.2	97.6	<1.0	<1.0	1.5	>99.0	>99.0
1991	70.0	6.9	59.9	<1.0	<1.0	0.8	>99.0	>99.0
1992	60.7	8.5	47.8	<1.0	<1.0	1.5	>99.0	>99.0
1993	58.4	4.9	48.1	<1.0	<1.0	84.2	>99.0	>99.0
1994	53.2	3.8	27.8	<1.0	<1.0	<1.0	>99.0	>99.0
1995	45.3	3.2	31.0	<1.0	<1.0	<1.0	92.5	>99.0
1996	70.8	1.9	26.3	<1.0	<1.0	<1.0	64.5	>99.0
1997	68.8	7.9	18.7	<1.0	<1.0	<1.0	51.2	>99.0
1998	64.2	7.3	42.7	<1.0	<1.0	<1.0	36.4	>99.0
1999	69.5	5.3	34.4	<1.0	<1.0	<1.0	10.8	>99.0
2000	53.7	7.0	29.1	<1.0	<1.0	<1.0	1.7	96.7
2001	69.0	2.8	31.7	<1.0	<1.0	<1.0	<1.0	71.5
2002	56.2	6.3	15.0	<1.0	<1.0	<1.0	<1.0	8.4
2003	80.3	2.9	21.5	<1.0	<1.0	<1.0	<1.0	<1.0
2004	55.1	11.4	12.2	<1.0	<1.0	<1.0	<1.0	<1.0
2005	97.5	2.7	31.9	<1.0	<1.0	<1.0	<1.0	<1.0
2006	97.5	46.6	12.4	<1.0	<1.0	<1.0	<1.0	<1.0

Table 4. Comparison of true fishing mortality at age with medians from 100 realizations under the four cases of treating index values less than 1.0.

		F at Age					Percent bias in Medians vs Truth			
Age	Year	True	Case 1	Case 2	Case 3	Case 4	Case 1	Case 2	Case 3	Case 4
1	1995	0.084	0.084	0.084	0.082	0.084	0	0	-3	0
1	1996	0.079	0.078	0.078	0.073	0.078	-1	-1	-7	-1
1	1997	0.085	0.082	0.082	0.073	0.082	-3	-3	-14	-3
1	1998	0.066	0.065	0.066	0.049	0.064	-1	0	-25	-3
1	1999	0.067	0.064	0.065	0.039	0.061	-5	-3	-42	-9
1	2000	0.060	0.058	0.060	0.031	0.055	-3	0	-49	-8
1	2001	0.046	0.044	0.045	0.038	0.044	-3	-1	-17	-3
1	2002	0.033	0.031	0.032	0.026	0.032	-4	-3	-20	-3
1	2003	0.032	0.031	0.031	0.029	0.032	-1	-2	-10	0
1	2004	0.032	0.031	0.033	0.025	0.029	-3	3	-21	-10
1	2005	0.036	0.030	0.024	0.057	0.034	-16	-34	59	-5
2	1995	0.378	0.380	0.381	0.377	0.380	1	1	0	1
2	1996	0.356	0.370	0.370	0.361	0.369	4	4	1	4
2	1997	0.380	0.385	0.385	0.358	0.382	1	1	-6	1
2	1998	0.299	0.299	0.301	0.256	0.297	0	1	-14	-1
2	1999	0.306	0.305	0.308	0.223	0.298	0	1	-27	-3
2	2000	0.272	0.266	0.270	0.155	0.253	-2	-1	-43	-7
2	2001	0.210	0.208	0.213	0.104	0.196	-1	1	-51	-7
2	2002	0.151	0.150	0.151	0.129	0.150	-1	0	-15	-1
2	2003	0.147	0.144	0.146	0.119	0.142	-2	0	-19	-3
2	2004	0.147	0.152	0.151	0.137	0.154	3	3	-6	5
2	2005	0.167	0.170	0.173	0.135	0.156	2	4	-19	-6
3	1995	0.730	0.709	0.710	0.705	0.709	-3	-3	-3	-3
3	1996	0.688	0.669	0.669	0.657	0.668	-3	-3	-4	-3
3	1997	0.737	0.732	0.733	0.700	0.729	-1	-1	-5	-1
3	1998	0.578	0.562	0.564	0.502	0.558	-3	-2	-13	-3
3	1999	0.597	0.573	0.577	0.457	0.563	-4	-3	-23	-6
3	2000	0.529	0.506	0.509	0.329	0.486	-4	-4	-38	-8
3	2001	0.410	0.378	0.388	0.188	0.357	-8	-5	-54	-13
3	2002	0.297	0.281	0.295	0.124	0.261	-5	-1	-58	-12
3	2003	0.289	0.275	0.278	0.227	0.275	-5	-4	-22	-5
3	2004	0.290	0.287	0.284	0.226	0.277	-1	-2	-22	-4
3	2005	0.329	0.336	0.335	0.291	0.338	2	2	-11	3
4	1995	0.973	0.913	0.913	0.909	0.913	-6	-6	-7	-6
4	1996	0.913	0.848	0.849	0.836	0.848	-7	-7	-8	-7
4	1997	0.980	0.910	0.911	0.877	0.908	-7	-7	-11	-7
4	1998	0.765	0.726	0.727	0.662	0.721	-5	-5	-13	-6
4	1999	0.790	0.725	0.729	0.595	0.714	-8	-8	-25	-10
4	2000	0.701	0.637	0.643	0.431	0.617	-9	-8	-39	-12
4	2001	0.542	0.488	0.499	0.262	0.465	-10	-8	-52	-14
4	2002	0.390	0.337	0.351	0.138	0.310	-14	-10	-65	-21
4	2003	0.381	0.340	0.361	0.125	0.308	-11	-5	-67	-19
4	2004	0.383	0.349	0.355	0.271	0.351	-9	-7	-29	-8

4	2005	0.435	0.410	0.408	0.303	0.392	-6	-6	-30	-10
5	1995	1.110	1.051	1.051	1.046	1.051	-5	-5	-6	-5
5	1996	1.041	0.974	0.975	0.960	0.973	-6	-6	-8	-7
5	1997	1.115	1.041	1.042	1.004	1.038	-7	-7	-10	-7
5	1998	0.870	0.820	0.822	0.750	0.814	-6	-5	-14	-6
5	1999	0.894	0.849	0.855	0.710	0.838	-5	-4	-21	-6
5	2000	0.793	0.730	0.737	0.502	0.706	-8	-7	-37	-11
5	2001	0.611	0.548	0.564	0.292	0.522	-10	-8	-52	-15
5	2002	0.438	0.390	0.407	0.166	0.361	-11	-7	-62	-18
5	2003	0.427	0.374	0.396	0.122	0.333	-12	-7	-71	-22
5	2004	0.429	0.375	0.405	0.108	0.328	-13	-6	-75	-23
5	2005	0.488	0.449	0.453	0.312	0.442	-8	-7	-36	-9
6	1995	1.177	1.133	1.133	1.128	1.133	-4	-4	-4	-4
6	1996	1.106	1.028	1.028	1.017	1.027	-7	-7	-8	-7
6	1997	1.185	1.107	1.107	1.069	1.102	-7	-7	-10	-7
6	1998	0.923	0.848	0.851	0.777	0.844	-8	-8	-16	-8
6	1999	0.947	0.867	0.875	0.727	0.856	-8	-8	-23	-10
6	2000	0.838	0.778	0.788	0.535	0.742	-7	-6	-36	-11
6	2001	0.644	0.573	0.591	0.308	0.542	-11	-8	-52	-16
6	2002	0.461	0.400	0.415	0.169	0.370	-13	-10	-63	-20
6	2003	0.447	0.385	0.403	0.126	0.341	-14	-10	-72	-24
6	2004	0.449	0.374	0.413	0.093	0.319	-17	-8	-79	-29
6	2005	0.511	0.424	0.474	0.090	0.350	-17	-7	-82	-32
7	1995	1.209	1.133	1.133	1.128	1.133	-6	-6	-7	-6
7	1996	1.136	1.028	1.028	1.017	1.027	-10	-10	-10	-10
7	1997	1.218	1.107	1.107	1.069	1.102	-9	-9	-12	-10
7	1998	0.948	0.848	0.851	0.777	0.844	-11	-10	-18	-11
7	1999	0.971	0.867	0.875	0.727	0.856	-11	-10	-25	-12
7	2000	0.859	0.778	0.788	0.535	0.742	-9	-8	-38	-14
7	2001	0.660	0.573	0.591	0.308	0.542	-13	-10	-53	-18
7	2002	0.471	0.400	0.415	0.169	0.370	-15	-12	-64	-21
7	2003	0.456	0.385	0.403	0.126	0.341	-16	-12	-72	-25
7	2004	0.458	0.374	0.413	0.093	0.319	-18	-10	-80	-30
7	2005	0.520	0.490	0.503	0.369	0.474	-6	-3	-29	-9
8	1995	1.227	1.133	1.133	1.128	1.133	-8	-8	-8	-8
8	1996	1.150	1.028	1.028	1.017	1.027	-11	-11	-12	-11
8	1997	1.234	1.107	1.107	1.069	1.102	-10	-10	-13	-11
8	1998	0.961	0.848	0.851	0.777	0.844	-12	-11	-19	-12
8	1999	0.984	0.867	0.875	0.727	0.856	-12	-11	-26	-13
8	2000	0.870	0.778	0.788	0.535	0.742	-11	-9	-38	-15
8	2001	0.667	0.573	0.591	0.308	0.542	-14	-11	-54	-19
8	2002	0.476	0.400	0.415	0.169	0.370	-16	-13	-64	-22
8	2003	0.461	0.385	0.403	0.126	0.341	-17	-12	-73	-26
8	2004	0.462	0.374	0.413	0.093	0.319	-19	-11	-80	-31
8	2005	0.525	0.490	0.503	0.369	0.474	-7	-4	-30	-10

Table 5. Comparison of true population numbers at age (thousands) with medians from 100 realizations under the four cases of treating index values less than 1.0.

		F at Age					Percent bias in Medians vs Truth			
Age	Year	True	Case 1	Case 2	Case 3	Case 4	Case 1	Case 2	Case 3	Case 4
1	1995	35236	35204	35180	35943	35249	0	0	2	0
1	1996	25724	26354	26310	27872	26458	2	2	8	3
1	1997	26449	26916	26811	30592	27112	2	1	16	3
1	1998	28054	28554	28367	36913	29142	2	1	32	4
1	1999	26207	27566	27102	44094	28562	5	3	68	9
1	2000	31907	32711	31866	61032	34452	3	0	91	8
1	2001	26383	27216	27231	31739	27606	3	3	20	5
1	2002	30976	31174	31460	37429	31911	1	2	21	3
1	2003	22272	21688	22217	24795	21886	-3	0	11	-2
1	2004	31379	31477	30866	38880	33836	0	-2	24	8
1	2005	13176	14426	20116	8650	13929	9	53	-34	6
1	2006	13176	13461	27312	3659	11610	2	107	-72	-12
2	1995	24025	24166	24162	24367	24182	1	1	1	1
2	1996	26523	26475	26454	27088	26520	0	0	2	0
2	1997	19463	19941	19918	21154	20006	2	2	9	3
2	1998	19900	20228	20157	23206	20440	2	1	17	3
2	1999	21504	21908	21761	28759	22399	2	1	34	4
2	2000	20057	21123	20803	34764	22028	5	4	73	10
2	2001	24611	25247	24559	48407	26662	3	0	97	8
2	2002	20634	21314	21398	24997	21670	3	4	21	5
2	2003	24546	24731	24986	29795	25386	1	2	21	3
2	2004	17667	17198	17604	19741	17318	-3	0	12	-2
2	2005	24890	24958	24448	31007	26906	0	-2	25	8
2	2006	10407	11427	16081	6674	11061	10	55	-36	6
3	1995	12645	12600	12599	12651	12604	0	0	0	0
3	1996	13490	13521	13517	13659	13529	0	0	1	0
3	1997	15224	14967	14959	15492	15024	-2	-2	2	-1
3	1998	10904	11082	11050	12111	11133	2	1	11	2
3	1999	12087	12282	12214	14715	12447	2	1	22	3
3	2000	12966	13234	13057	18672	13585	2	1	44	5
3	2001	12517	13289	12992	24479	13977	6	4	96	12
3	2002	16330	16799	16227	35675	17965	3	-1	118	10
3	2003	14523	15014	15026	17976	15280	3	3	24	5
3	2004	17356	17479	17544	21764	18001	1	1	25	4
3	2005	12489	12140	12403	14081	12114	-3	-1	13	-3
3	2006	17254	17285	16870	22139	18867	0	-2	28	9
4	1995	4938	4959	4959	4970	4960	0	0	1	0
4	1996	5000	5099	5100	5141	5102	2	2	3	2
4	1997	5559	5670	5668	5796	5677	2	2	4	2
4	1998	5974	5904	5899	6279	5933	-1	-1	5	-1
4	1999	5015	5191	5169	5967	5241	4	3	19	5
4	2000	5454	5672	5621	7670	5803	4	3	41	6
4	2001	6259	6555	6417	11078	6850	5	3	77	9
4	2002	6802	7408	7157	16538	7980	9	5	143	17
4	2003	9941	10365	9909	25846	11260	4	0	160	13
4	2004	8905	9327	9348	11763	9506	5	5	32	7

4	2005	10635	10737	10873	14259	11289	1	2	34	6
4	2006	7361	7048	7273	8649	7071	-4	-1	17	-4
5	1995	1354	1488	1488	1490	1488	10	10	10	10
5	1996	1529	1639	1639	1649	1640	7	7	8	7
5	1997	1645	1772	1771	1809	1775	8	8	10	8
5	1998	1710	1867	1865	1977	1875	9	9	16	10
5	1999	2277	2343	2335	2659	2368	3	3	17	4
5	2000	1865	2047	2037	2688	2098	10	9	44	12
5	2001	2216	2451	2418	4064	2576	11	9	83	16
5	2002	2982	3290	3203	6966	3519	10	7	134	18
5	2003	3770	4345	4133	11817	4791	15	10	213	27
5	2004	5560	6003	5660	18664	6770	8	2	236	22
5	2005	4971	5379	5373	7295	5482	8	8	47	10
5	2006	5636	5852	5934	8603	6258	4	5	53	11
6	1995	355	394	394	395	394	11	11	11	11
6	1996	366	427	427	429	427	17	17	17	17
6	1997	442	505	505	515	506	14	14	16	14
6	1998	442	510	509	542	512	15	15	23	16
6	1999	587	679	676	760	684	16	15	30	16
6	2000	762	815	808	1076	833	7	6	41	9
6	2001	691	812	798	1342	852	17	16	94	23
6	2002	985	1157	1136	2464	1258	18	15	150	28
6	2003	1575	1825	1750	4824	2010	16	11	206	28
6	2004	2014	2445	2271	8516	2823	21	13	323	40
6	2005	2964	3397	3073	13723	3991	15	4	363	35
6	2006	2498	2812	2784	4372	2903	13	11	75	16
7	1995	66	83	83	83	83	27	27	27	27
7	1996	90	106	106	106	106	18	18	19	18
7	1997	99	122	122	125	122	23	23	26	24
7	1998	111	135	135	145	136	22	22	31	23
7	1999	144	180	179	201	181	25	25	40	26
7	2000	186	232	230	305	237	24	23	64	27
7	2001	270	307	303	514	322	14	12	90	19
7	2002	297	369	358	804	405	24	20	171	36
7	2003	509	633	609	1692	703	24	20	233	38
7	2004	825	1010	963	3467	1175	23	17	320	42
7	2005	1053	1376	1234	6359	1679	31	17	504	60
7	2006	1456	1809	1562	10263	2309	24	7	605	59
8	1995	9	13	13	13	13	48	48	48	48
8	1996	18	24	24	24	24	34	33	34	34
8	1997	28	36	36	36	36	27	27	29	27
8	1998	31	41	41	44	41	34	34	42	35
8	1999	45	60	59	68	60	33	32	52	34
8	2000	58	73	73	95	75	26	25	64	29
8	2001	85	110	108	182	113	30	28	115	34
8	2002	150	184	180	388	198	23	20	159	32
8	2003	228	287	277	803	321	26	22	252	41
8	2004	382	487	456	1705	559	28	19	347	46
8	2005	624	690	666	845	710	11	7	35	14
8	2006	814	1057	937	14806	997	30	15	1718	22

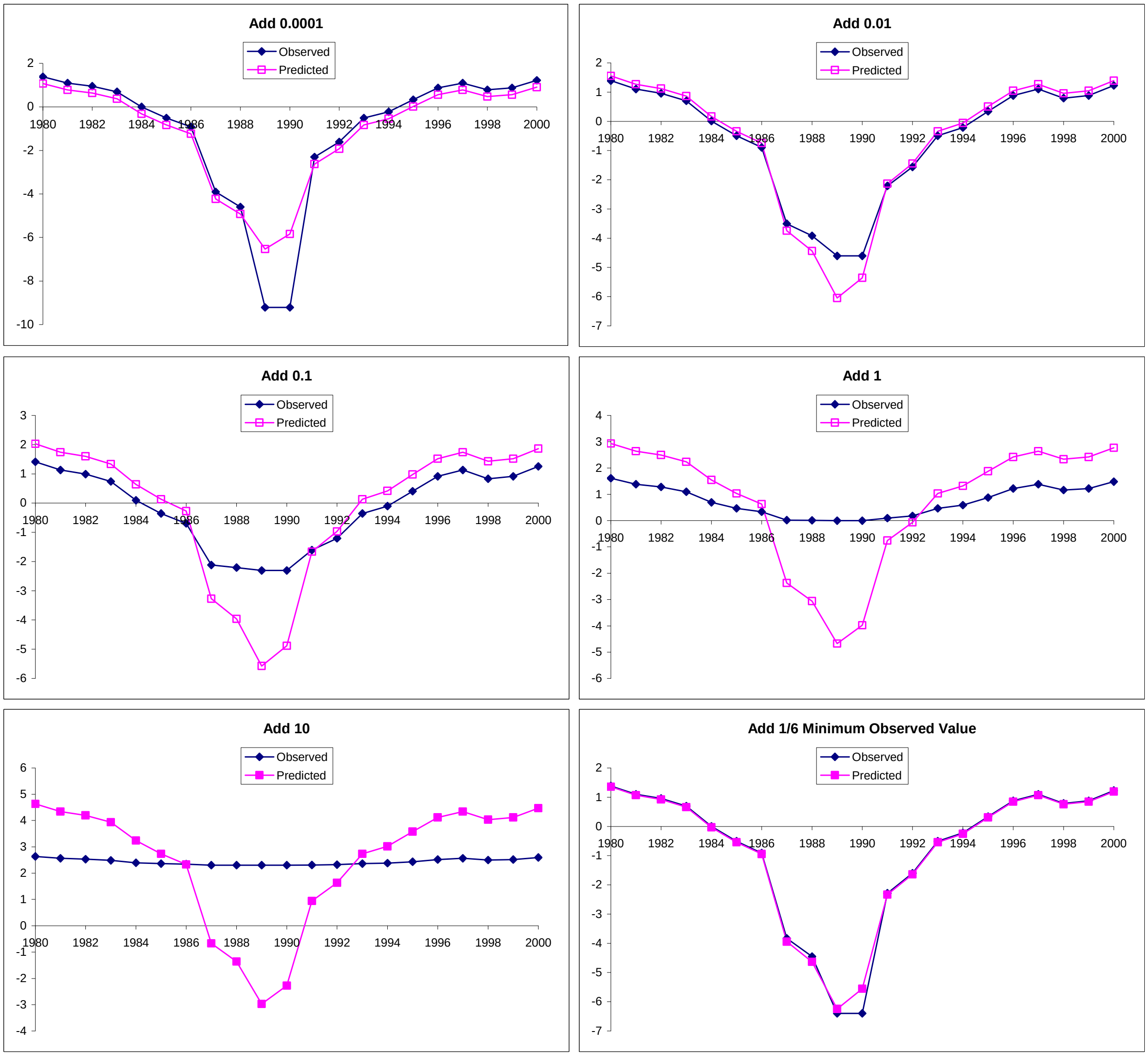

Figure 1. Comparison of observed and predicted indices when observed values are rounded to two decimal places and resulting zeros are replaced by different constants. The predicted indices follow the true population pattern and are scaled by a catchability coefficient to minimize the natural logarithm of the squared residuals. Note the y-axes are log scale.

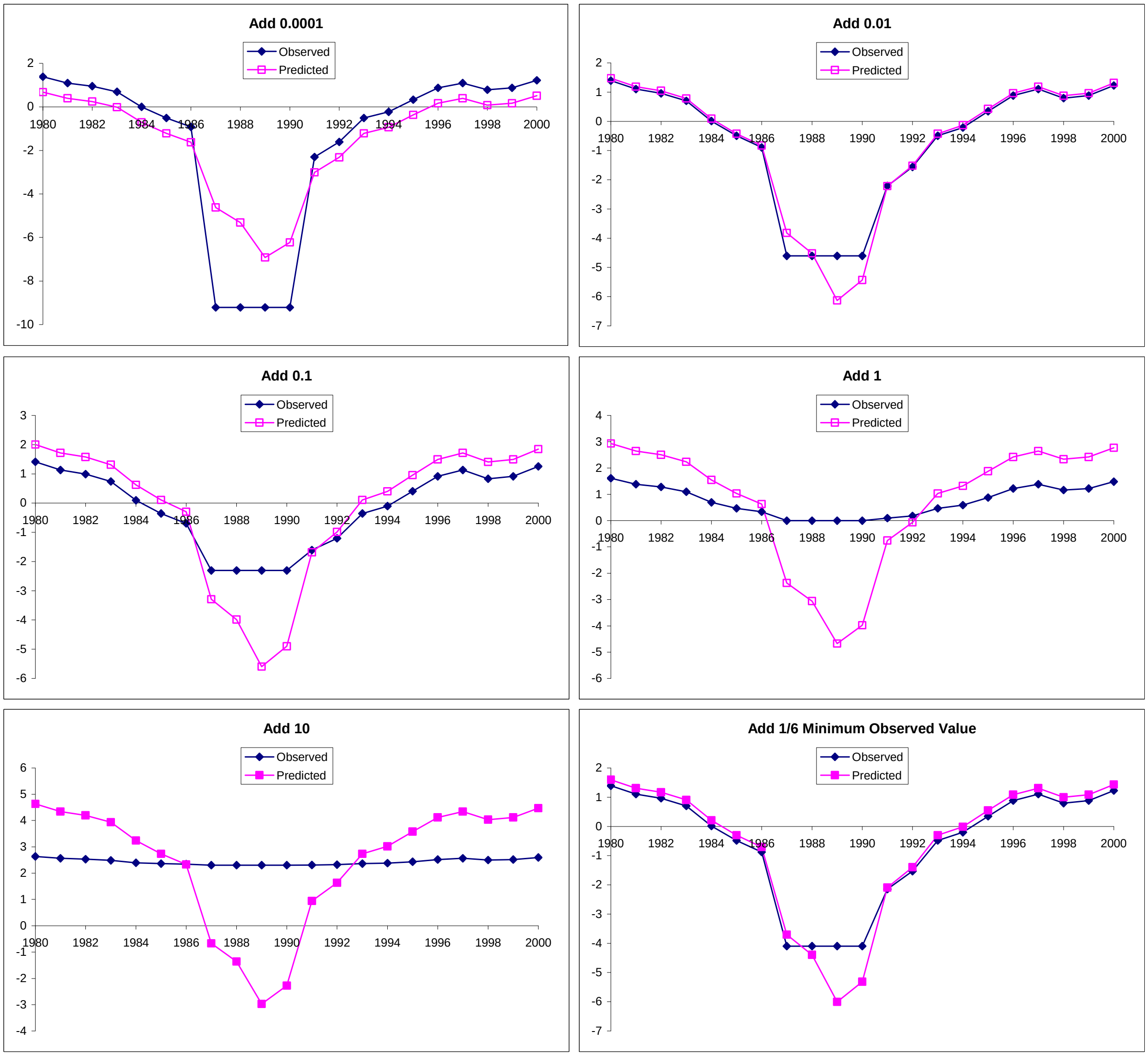

Figure 2. Comparison of observed and predicted indices when observed values are rounded to one decimal place and resulting zeros are replaced by different constants. The predicted indices follow the true population pattern and are scaled by a catchability coefficient to minimize the natural logarithm of the squared residuals. Note the y-axes are log scale.

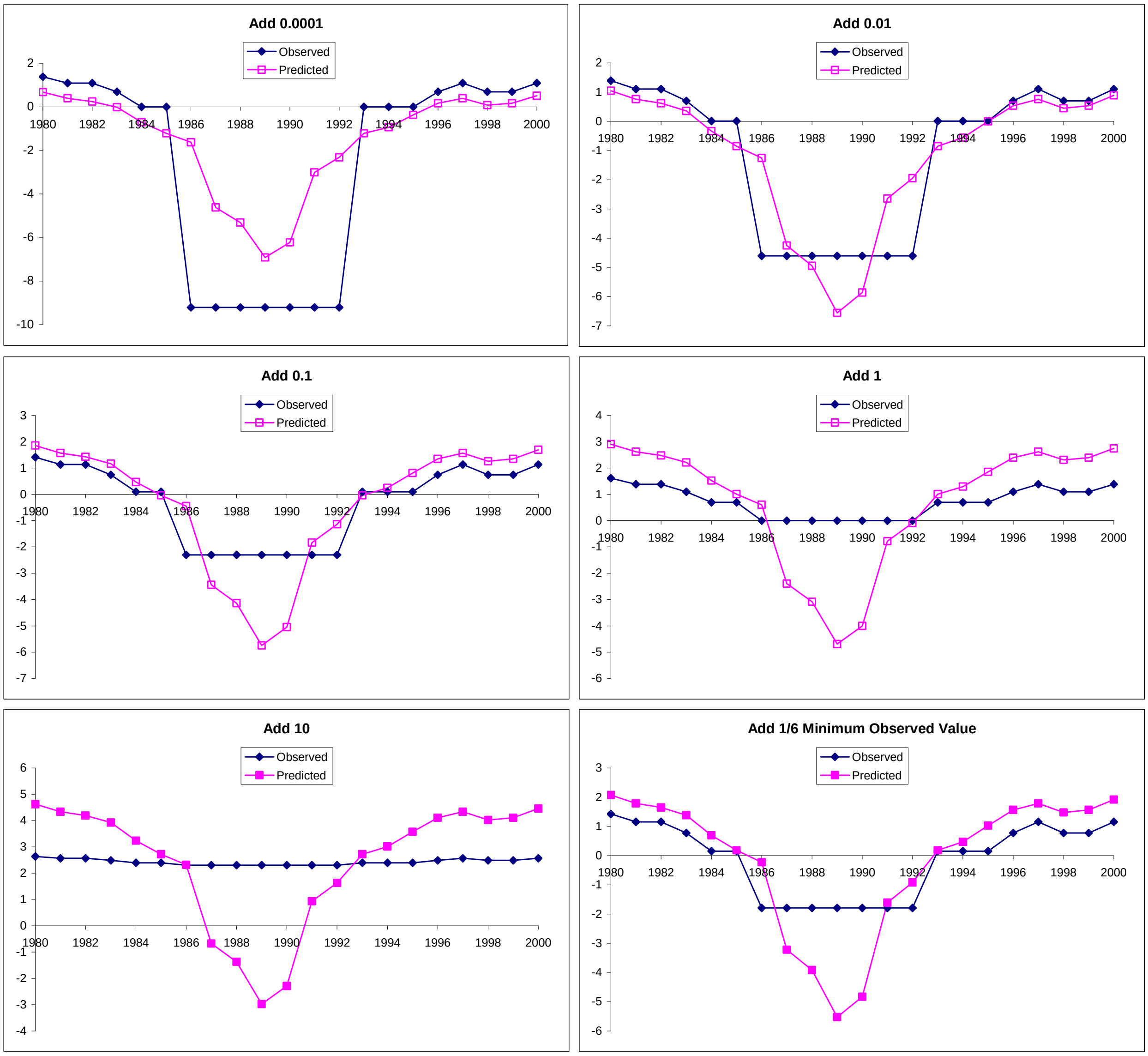

Figure 3. Comparison of observed and predicted indices when observed values are rounded to zero decimal places and resulting zeros are replaced by different constants. The predicted indices follow the true population pattern and are scaled by a catchability coefficient to minimize the natural logarithm of the squared residuals. Note the y-axes are log scale

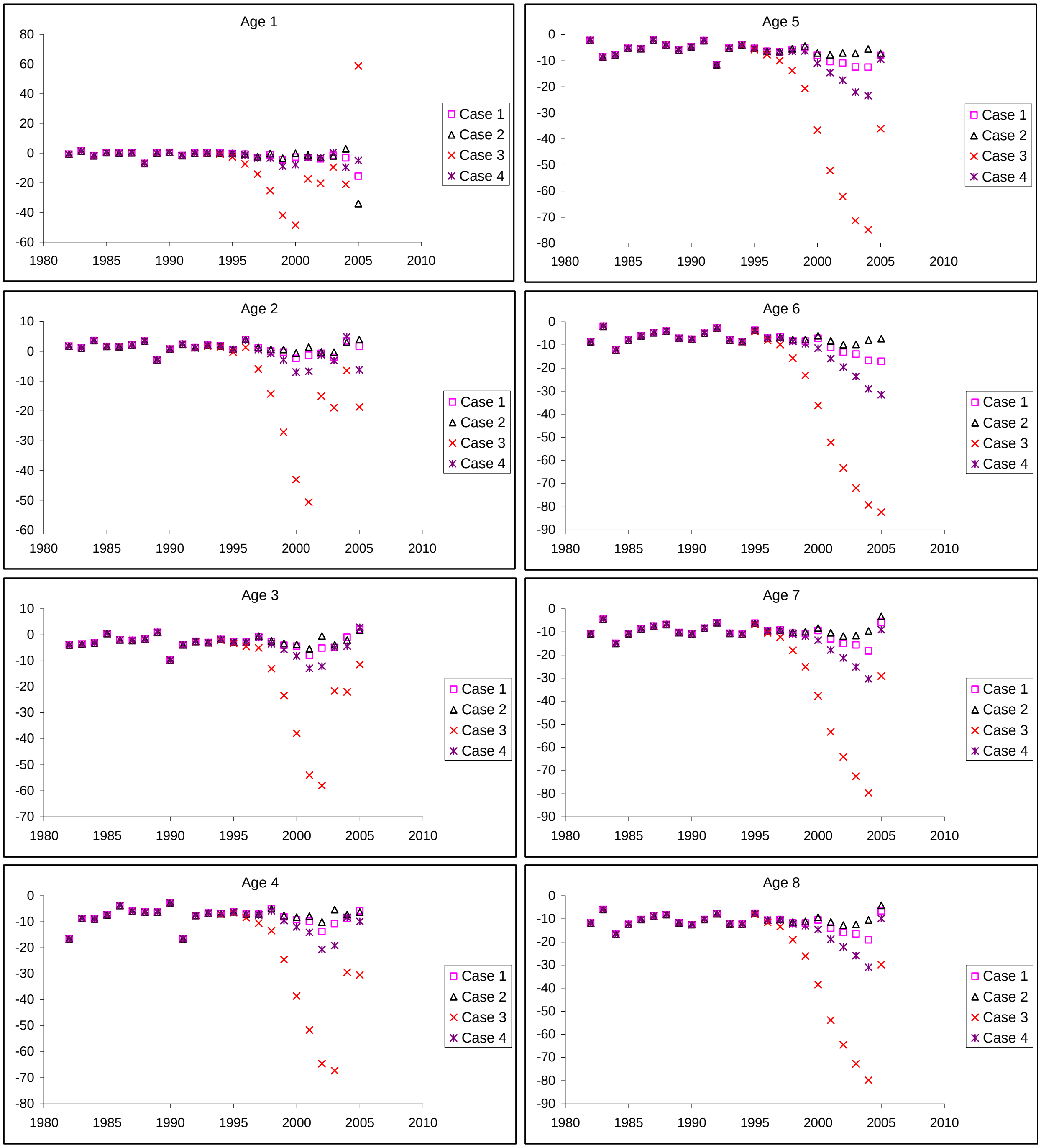

Figure 4. Percent bias in the medians of fishing mortality by age and year for the four cases of how index values less than one are treated

Figure 5. Percent bias in the medians of population numbers by age and year for the four cases of how index values less than one are treated. Note a number of case 3 values are too large to be shown on the plots, values are given in Table 5.