

Local influence diagnostics for the retrospective problem in sequential population analysis

Noel G. Cadigan and Patrick J. Farrell

Cadigan, N. G., and Farrell, P. J. 2005. Local influence diagnostics for the retrospective problem in sequential population analysis. — ICES Journal of Marine Science, 62: 256–265.

The retrospective problem involves systematic differences in sequential population analysis (SPA) estimates of stock size or some other quantity in a reference year. The differences occur as successively more data are used for estimation, and they appear to be structural biases caused by a mis-specification of the SPA. In some cases the retrospective problem is so severe that the SPA is considered to be too unreliable for stock assessment purposes. There are many possible sources of retrospective patterns, and it is usually difficult in practice to determine which are more likely. We propose diagnostics to help determine the more likely causes. We use local influence diagnostics to investigate whether small changes or perturbations to SPA input components such as catches or natural mortalities can remove or reduce retrospective patterns. We show, for the fall-spawning herring stock in the southern Gulf of St. Lawrence SPA, that relatively small age- and year-specific changes to the SPA assumptions about the proportional relationship between an abundance index and stock size can result in greatly reduced retrospective patterns. We therefore conclude *prima facie* that these assumptions are a plausible source of the retrospective pattern. Larger changes to catches, natural mortality assumptions, or estimation weights are required to reduce the retrospective pattern. These other factors seem to be less plausible sources of the retrospective pattern, although this is best assessed by the herring stock experts who are more knowledgeable about the fishery and other scientific information for this stock.

© 2004 International Council for the Exploration of the Sea. Published by Elsevier Ltd. All rights reserved.

Keywords: fish stock assessment, herring, model mis-specification, perturbation analysis, virtual population analysis.

Received 17 May 2004; accepted 5 November 2004.

N. G. Cadigan: Northwest Atlantic Fisheries Center, Fisheries and Oceans Canada, PO Box 5667, St. John's, NL, Canada A1C 5X1. P. J. Farrell: School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada K1S 5B6. Correspondence to N. G. Cadigan: tel: +1 709 772 5028; fax: +1 709 772 4105; e-mail: cadiganN@dfo-mpo.gc.ca.

Introduction

The retrospective problem in sequential population analysis (SPA) has received considerable attention in fish stock assessments (e.g. Mohn, 1999). SPA is an analytical model of fishery catch data that can be used to estimate stock size. The model we investigate is described in the next paragraph. The retrospective problem involves a systematic pattern in stock estimates as catches and other stock data are updated. More specifically, let $S_{y,t}$ denote an SPA estimate of stock size in year y based on data up to year $t \geq y$. A retrospective problem is said to exist if $S_{y,t}$, $S_{y,t+1}$, $S_{y,t+2}$, and so on deviate systematically in a decreasing or increasing trend. This problem can be so severe that the SPA is considered to be too unreliable for stock assessment and management purposes.

We investigate the common SPA cohort model

$$N_{a+1,y+1} = N_{a,y} \exp(-M_{a,y}) - C_{a,y} \exp(-M_{a,y}/2), \quad (1)$$

where $N_{a,y}$ is the stock number at age a in year y , $C_{a,y}$ is the commercial catch, and $M_{a,y}$ is the natural mortality. If estimates of $N_{a,y}$ s in the last year (survivors) and at the oldest age in the SPA are available, along with time-series of $C_{a,y}$ s and $M_{a,y}$ s, then Equation (1) can be used to reconstruct stock size for all ages and years, and other derived quantities such as total biomass and spawner biomass. The survivors and other parameters can be estimated using stock size indices, although often an approximation is used to specify the numbers at the oldest age. A stock size index is a measurement that is

proportional to stock size; that is, if $R_{a,y}$ is an index from a stochastic measurement process then

$$E(R_{a,y}) = q_a N_{a,y}, \quad (2)$$

where the unknown constant of proportionality q_a is usually age-dependent and must be estimated. A common fit function used to estimate the SPA parameters, which we collectively refer to as θ , is the log error sum of squares,

$$l(\theta) = \sum_{a,y} \{ \log(r_{a,y}) - \log[q_a N_{a,y}(t)] \}^2.$$

A substantial retrospective problem exists in the SPA for the fall spawning herring stock in the southern Gulf of St. Lawrence (see Figure 41 in LeBlanc *et al.*, 2003). This stock, found off the north coast of Nova Scotia, Canada, in NAFO Division 4T, is commonly referred to as 4T fall herring. Some details about the SPA for this stock are provided in the Results section. Additional information is available in LeBlanc *et al.* (2003) and Cadigan and Farrell (2003). In the 2002 assessment of this stock the SPA was rejected because of the retrospective problem (LeBlanc *et al.*, 2002). We investigate an SPA structure that closely resembles the formulation in LeBlanc *et al.* (2003); our retrospective patterns are also similar. For example, estimates of total biomass for ages 5–10 (B_+) are shown in Figure 1. The estimates cover the years $y = 1978, \dots, t$, for $t = 1998, \dots, 2002$, where t indicates the last year of catches and other fishery data used to estimate population size. The estimates for year y usually decrease as t increases and more data are used. For example, the estimate of B_+ in

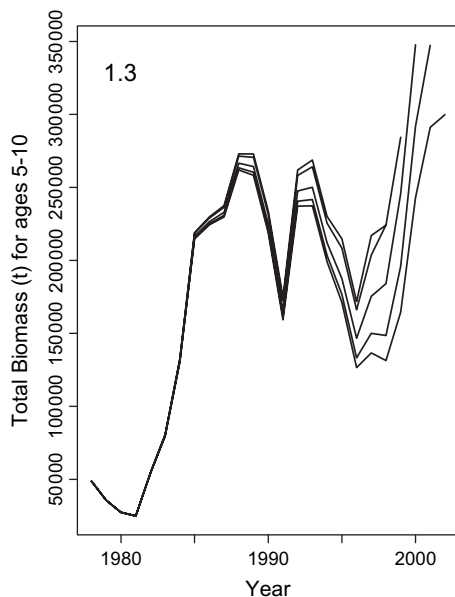


Figure 1. Retrospective estimates of total biomass (B_+) for the 4T fall herring stock. The retrospective ρ statistic is shown in the top left-hand corner.

1998 based on data up to 1998 is 224 000 t; if data up to 2002 are used to obtain an updated 1998 estimate, the result is 131 000 t, which is almost 50% lower. This consistently decreasing trend indicates a structural bias in the population size estimators caused by model mis-specification (Evans, 1996). Similar retrospective patterns exist in the SPA estimates of stock numbers-at-age and fishing mortalities (see Cadigan and Farrell, 2003). The patterns in Figure 1 are severe, but not uniquely so. For example, a haddock stock considered by Sinclair *et al.* (1991) exhibited similar symptoms, while another stock considered in Mohn (1999) possessed a more severe retrospective pattern.

A common perception in stock assessments is that historic estimates of stock size are more accurate than recent ones. This will tend to be true for stocks that are heavily exploited, and for which accurate catch data exist (Pope, 1972). Hence, when a retrospective problem as in Figure 1 is present, the common perception is that current stock size is overestimated. This is important because current stock size and trends in recent stock size are usually required for fishery management decisions such as setting the total allowable catch (TAC) for next year. If current stock size is overestimated then this may mean that the TAC will be set too high and be unsustainable. However, as pointed out by Sinclair *et al.* (1991), for certain types of SPA model mis-specifications the historic estimates may be less accurate than the current estimates. Mohn (1999) also demonstrated this phenomenon using simulated data with model mis-specifications. He showed that certain types of model mis-specifications tended to compound over time such that adding more data yielded estimates of historic stock size with larger biases. However, even in these situations the current trends in stock size estimated by SPA may be overly optimistic and lead to a TAC that is too high.

Mohn (1999) presented diagnostics to help discriminate between the possible sources of mis-specification that cause the retrospective problem. One of his diagnostics involved examining the magnitude of perturbations to model components required to remove the retrospective pattern. Mohn (1999) considered simple perturbations that involved adding a common “effect” to part of a model component. For example, he considered catch perturbations of the form

$$C_{a,y}(\omega) = \begin{cases} C_{a,y}, & y < y_0, \\ C_{a,y}\omega\phi_a, & y \geq y_0, \end{cases}$$

where $C_{a,y}$ and $C_{a,y}(\omega)$ were the observed and perturbed catch at age a in year y . The perturbation $\omega\phi_a$ was applied only after some specified year y_0 . The magnitude of the perturbation was controlled by ω , while ϕ_a was an age effect that was fixed in all perturbations. This analysis was used to explore whether unreported and age-dependent discarding of catches starting in year y_0 could be the source of the retrospective problem. The magnitude and timing of the perturbation required to remove the pattern was used to assess whether discarding was a plausible causal mechanism.

Mohn (1999) considered perturbations to other model components as well. The “parameters” of his perturbation analyses were ω and y_o , which he profiled over to find values that removed the retrospective problem.

Mohn (1999) proposed a statistic to measure the retrospective pattern,

$$\rho = \sum_{y=y_o}^{Y-1} \frac{S_{y,Y} - S_{y,Y}}{S_{y,Y}}, \quad (3)$$

where y_o and Y are the first and last years for which the retrospective pattern was assessed. Recall that $S_{y,Y}$ is the estimate of stock size in year y based on all catch and other stock data up to year Y . If the retrospective estimates for year y based on data only up to year y (i.e. $S_{y,y}$) for $y = y_o, \dots, Y-1$ fluctuate randomly about $S_{y,Y}$, then ρ is approximately zero. For 4T fall herring we assessed the retrospective pattern for $y_o = 1998$ and $Y = 2002$. The value of ρ is shown in the upper left-hand corner of Figure 1. The positive value indicates that retrospective estimates of stock size decrease with time.

In this paper, we also use perturbation analyses to diagnose the more likely sources of the retrospective problem. We use less constrained perturbations than Mohn (1999); for example, we investigate catch perturbations of the form

$$C_{a,y}(\omega) = C_{a,y} \times \omega_{a,y}. \quad (4)$$

In Equation (4) we perturb each $C_{a,y}$ separately, whereas Mohn (1999) considered more simple perturbations. We also find perturbations that remove the retrospective pattern. The advantage of searching over a higher dimensional and less restricted perturbation space is the potential of finding more realistic perturbations to remove the retrospective pattern than those presented in Mohn (1999).

We use the local influence approach for perturbation analyses. This method is briefly described in the next section, and more fully discussed in Cadigan and Farrell (2002). It is based on the influence graph of $\rho(\omega)$ vs. ω , where $\rho(\omega)$ is the value of ρ we find after re-estimating the SPA parameters using the perturbation ω . Similar to Cook (1986), we use basic concepts in differential geometry to study the effect of a perturbation ω on $\rho(\omega)$. Our approach is based on the slope of the perturbation (influence) surface near the origin, which is the point of no perturbation. This is why it is referred to as local influence. When $\rho(\omega)$ is fairly linear around a relevant neighborhood of the origin then our approach provides a good description of influence. We find the direction of maximum slope for $\rho(\omega)$ at the perturbation origin. This is the local direction of greatest change in $\rho(\omega)$. We then examine perturbations ω in this direction that reduce $\rho(\omega)$ to near zero. This involves re-estimating the SPA a small number of times, but the directions themselves only require the unperturbed SPA parameter estimates. When the perturbation surface of ρ is

linear, our method finds the smallest perturbation that removes the retrospective pattern, as measured by ρ .

In the Results section, we consider influence for four distinct perturbation schemes on catches, natural mortality, survey catchability, and case weights in order to determine if the retrospective pattern in the 4T fall herring SPA is more likely caused by any of these components, or at least if the pattern is unlikely to be caused by some of the components. Our diagnostics do not provide corrections for retrospective patterns. Such methods are beyond the scope of this paper, although we feel that they will likely be stock-specific and involve additional information about data and stock dynamics.

Local influence approach

Cadigan and Farrell (2002) considered local influence diagnostics for problems that involved estimating a $p \times 1$ parameter vector θ by maximizing a fit function $l(\theta)$ that had basic smoothness properties. The estimate of θ , denoted as $\hat{\theta}$, was the solution to

$$\dot{l}(\hat{\theta}) = \frac{\partial l(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = 0.$$

They perturbed model components with a $k \times 1$ perturbation vector ω and studied the influence of the perturbations on key model results. The perturbation ω was of the form $\omega = \omega(h) = \omega_o + h\mathbf{d}$, where ω_o was the null perturbation vector, \mathbf{d} was a fixed direction vector of length one, and h was a scalar that determined the magnitude of the perturbation. For multiplicative perturbations ω_o is usually a $k \times 1$ vector of ones, while for additive perturbations ω_o is usually a vector of zeros. The dimension of ω (i.e. k) could be large; for example, in one of the SPA applications in Cadigan and Farrell (2002), k was equal to 520.

Cadigan and Farrell (2002) measured influence on $g(\hat{\theta})$, an important but arbitrarily specified scalar SPA result. They considered some basic geometric properties of the influence surface of $g_\omega(\hat{\theta}_\omega)$ vs. ω near the origin, ω_o . The primary diagnostic they used was the local slope in the direction \mathbf{d} ,

$$S(\mathbf{d}) = \frac{\partial g_\omega(\hat{\theta}_\omega)}{\partial h} \Big|_{h=0} = \mathbf{d} \cdot \frac{\partial g_\omega(\hat{\theta}_\omega)}{\partial \omega} \Big|_{\omega=\omega_o} = \mathbf{d}' \dot{g}_o.$$

A particularly interesting diagnostic was the direction of maximum slope,

$$s_{\max} = \dot{g}_o / \sqrt{\dot{g}_o' \dot{g}_o}, \quad (5)$$

which is a $k \times 1$ vector. Further information and computational details about these diagnostics are given in Section 2.1 of Cadigan and Farrell (2002).

Some minor modifications to the methods in Cadigan and Farrell (2002) are required to study local influence for ρ . The influence measure $g_{\omega}(\hat{\theta}_{\omega})$ they used involved parameter estimates from the optimization of a single fit function, whereas ρ is based on parameter estimates from multiple optimizations. For example, ρ in Figure 1 is based on five optimizations. Let $S_{y,t}(\omega)$ denote a perturbed estimate of stock size. Using $S_{y,t}(\omega)$ s in Equation (3) gives the perturbed value, $\rho(\omega)$. Based on their notation, $S_{y,t}(\omega)$ is a particular form for $g(\hat{\theta})$. Let

$$\dot{S}_{o,y,t} = d \left. \frac{\partial S_{y,t}(\omega)}{\partial \omega} \right|_{\omega=\omega_o}$$

be a vector of local derivatives. Each $\dot{S}_{o,y,t}$ can be computed using the methods in Cadigan and Farrell (2002). The local slope of $\rho(\omega)$ in the direction d is given by

$$\dot{\rho}_o = d \left. \frac{\partial \rho(\omega)}{\partial \omega} \right|_{\omega=\omega_o} = d \left[\sum_y (S_{o,y,Y} S_{o,y,Y}^{-1} \dot{S}_{o,y,Y} - \dot{S}_{o,y,Y}) S_{o,y,Y}^{-1} \right], \quad (6)$$

where Y is the last year under consideration. Let $\dot{\rho}_{\max}$ denote the maximum local slope. The direction of maximum slope, s_{\max} , can be found using Equation (5) with \dot{g}_o equal to the $[\cdot]$ term in Equation (6). Additional computational information is provided in Cadigan and Farrell (2003).

We can use s_{\max} to find perturbed SPA inputs that result in reduced retrospective patterns. We illustrate this using catch perturbations. Let $\omega_{\max}(h) = \omega_o + h s_{\max}$ and let $\rho_{\max}(h) = \rho\{\omega_{\max}(h)\}$ denote the ρ value based on $\omega_{\max}(h)$ perturbed catches, $C_{a,y}(h) = C_{a,y} \times \{1 + h s_{a,y}\}$, where $s_{a,y}$ is the element of the vector s_{\max} that corresponds to $C_{a,y}$. Note that the influence graph of $\rho(\omega)$ is a $k+1$ dimensional surface where $k = \dim(\omega)$, and $\rho_{\max}(h)$ is called the lifted line that is the intersection of $\rho(\omega)$ and the vertical plane at s_{\max} (or the plane containing all vectors orthogonal to s_{\max}). If the influence surface in the direction of s_{\max} is linear, then $\rho_{\max}(h) = \rho + \dot{\rho}_{\max} h$, which implies that $\rho_{\max}(h) = 0$ when $h = h_{\max} = -\rho / \dot{\rho}_{\max}$. In this case using $\omega_{\max}(h_{\max})$ gives the smallest perturbation that removes the retrospective pattern. The same procedure can be used to find perturbations to other SPA inputs (i.e. M in Equation 1) and assumptions (i.e. constant catchability at age, q_a , in Equation 2) that result in a SPA with reduced retrospective patterns. When the influence graph has some nonlinearity, $\omega_{\max}(h_{\max})$ may not completely remove the retrospective patterns. In this situation a value of h close to h_{\max} , determined through trial and error, will usually be a better choice. This is illustrated in the following sections.

Results

We used essentially the same SPA formulation for 4T fall herring as outlined in LeBlanc *et al.* (2003). The SPA is based on catches for 1978–2002 and for ages 4–10.

Natural mortality, $M_{a,y}$, was assumed to be 0.2 for all ages and years. A catch per unit effort (cpue) index of stock abundance was used to estimate SPA parameters. The $N_{10,y}$ s were approximated using assumptions about fishing mortalities; that is, $F_{10,y} = (F_{8,y} + F_{9,y})/2$ where

$$F_{a,y} = \log(N_{a,y}/N_{a+1,y+1}) - M_{a,y}.$$

Note that estimates at age 4 in the last year are more imprecise than at other ages; therefore, it is common to exclude this age when examining retrospective patterns. It is for this reason that B_+ in Figure 1 is based on ages 5–10. Further details regarding the SPA used here are given in Cadigan and Farrell (2003). To save space, results for SPA estimates of stock size, residuals, etc., are not presented here, but are given as Figures 3–8 in Cadigan and Farrell (2003).

Pronounced age-dependent time trends in SPA residuals are apparent for this stock (see Figure 7 in Cadigan and Farrell, 2003), which suggests that assumptions are mis-specified in the SPA. This is consistent with the results in Figure 1 presented here. Retrospective patterns in abundance-at-age estimates ($\hat{N}_{a,y}$) are similar for all ages except age four (see Figure 9 in Cadigan and Farrell, 2003), where the retrospective pattern is less pronounced. However, these graphical analyses of residuals provided little insight about the source of model mis-specification. In the next four subsections, we present diagnostics that can assist in identifying this source.

Catch perturbations

In this section we present the direction of maximum slope for ρ , i.e. s_{\max} , based on catch perturbations described by Equation (4). The elements of s_{\max} are shown in panel a of Figure 2. At the top of each panel, $-h_{\max}^{-1} = \dot{\rho}_{\max}/\rho$ is given in per cent. Values for $\dot{\rho}_{\max}$ can be derived from h_{\max} and ρ in Figure 1. By construction the $\dot{\rho}_{\max}$ s are always positive, which indicates that perturbations in these directions will increase ρ . Hence, these results suggest that increasing reported catches from the 1993–1996 cohorts, increasing catches at age 10 prior to 2001, and decreasing most other catches will result in relatively large reductions in retrospective patterns. The relatively high sensitivity to catches at age 10 is related to the approximation used to compute $N_{10,y}$ (see Results). This is a problem because it indicates that the model is sensitive to this somewhat artificial constraint. Potentially better methods such as F-shrinkage are available to estimate the $N_{10,y}$ s; however, this is beyond the scope of this paper. Cadigan and Farrell (2003) also computed s_{\max} for ps based on total abundance and fishing mortality. Their results (see their Figure 11) were similar to those in Figure 2 here; however, the elements of s_{\max} for ρ based on fishing mortalities tended to be opposite in sign since the retrospective trends for fishing mortality are the reverse of the trends for B_+ .

If we reduce catches in the direction of s_{\max} in Figure 2a by $h_{\max} = -5.6$, then ρ should be near zero for this

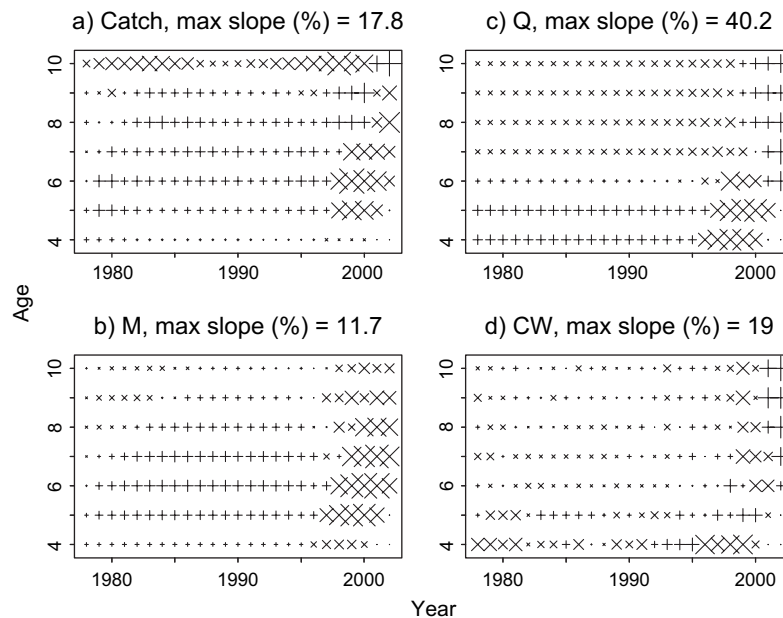


Figure 2. Local influence diagnostics for ρ . Each panel shows the elements of s_{\max} for a perturbation to: a) catches, b) M, c) cpue catchability Q, and d) case weights, CW. The size and type of plotting symbols are proportional to the absolute value and sign of the elements, respectively. At the top of each panel $\hat{\rho}_{\max}$ is shown in per cent of the ρ value in Figure 1.

perturbation. However, Cadigan and Farrell (2003) showed that the influence surface in the direction s_{\max} was somewhat nonlinear, and they found that a perturbation with $h = -5.6$ was not feasible because it resulted in some negative catches. The smallest value for h that could be used was $h = -2.5$. These perturbations are shown in Figure 3a; they are simply the results in Figure 2a scaled by -2.5 . The change in most catches is less than $\pm 25\%$, but a small number of catches are altered by more than $\pm 50\%$. Differences between the total observed and perturbed catches are shown in Figure 4. Perturbed catches are substantially larger in total after 1997. The retrospective pattern based on the perturbed catches is shown in Figure 5a. These catch perturbations resulted in a decrease of the ρ statistic from 1.3 to 0.47; however, a retrospective pattern still exists. The catch perturbations resulted in a poorer fit; the mean square error (MSE) increased from 0.22 to 0.29. These results suggest that fairly large changes to the catches are required to reduce the retrospective pattern. We conclude that small changes to the catches cannot remove the retrospective problem in the 4T fall herring SPA model formulation.

M perturbations

The M perturbations considered were of the form

$$M_{a,y}(\omega) = 0.2 \times \omega_{a,y},$$

where $\omega_o = 1$. The elements of s_{\max} are shown in Figure 2b. They are broadly similar to the catch perturbation

results in Figure 2a, although the influence of M and C at age 10 differs. These results suggest that an increase in M for the 1992–1996 cohorts and a decrease in M for most other cohorts will reduce the retrospective pattern.

If the influence surface is reasonably linear in the s_{\max} direction then a perturbation with h near -8.5 should reduce ρ to near zero. After a few trials, $h = -6.5$ was selected. The trials involved specifying a value for h near -8.5 , changing the values for M using $\{M_{a,y}\} = 0.2 \times (1 + h s_{\max})$, and then re-computing ρ to see if it was close to zero. The value of h selected was smaller in absolute value than the local approximation of -8.5 attributable to the nonlinearity in the influence curve. Figure 17c of Cadigan and Farrell (2003) showed that ρ decreased slightly faster than the linear approximation prediction, which is why $h = -6.5$ resulted in a smaller $|\rho|$ than h_{\max} . The perturbations based on $h = -6.5$ are shown in Figure 3b. The perturbations are large; that is, many of them are greater than 50% in absolute value, and some perturbations are greater than 150%. The perturbed retrospective patterns are shown Figure 5b. They are substantially reduced compared to those in Figure 5a, and to those in Figure 1, although the relative magnitude of the perturbations required to give these reductions is large (i.e. $|h| = 6.5$).

The MSE for the M-perturbed SPA dropped slightly from 0.22 to 0.2. This is another difference compared to the catch perturbations in the previous section. The M and catch perturbed SPAs give substantially different estimates of population size; for example, the maximum absolute

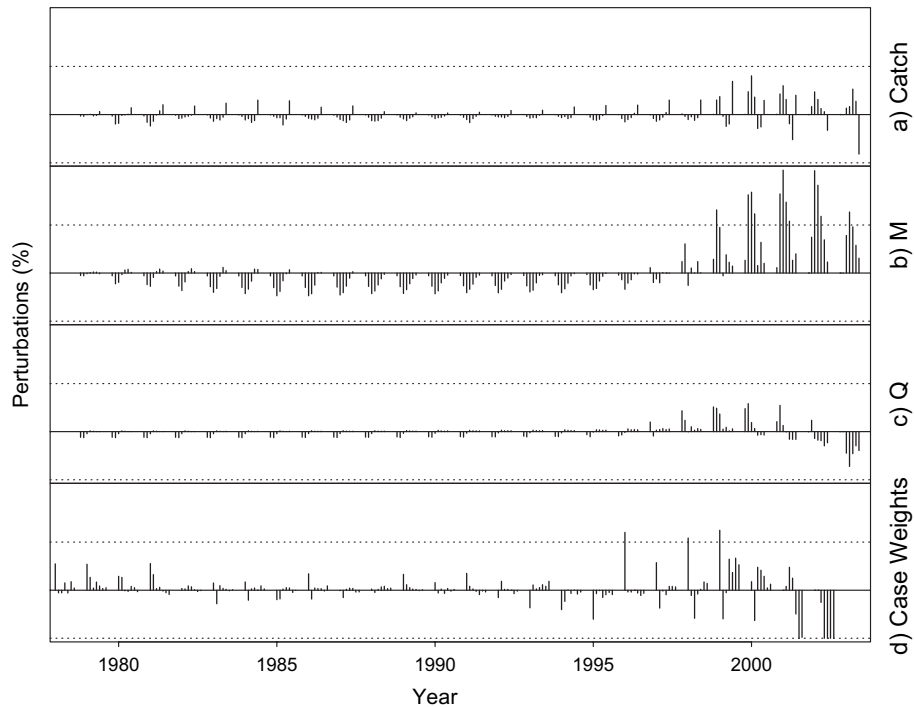


Figure 3. Perturbations to reduce the retrospective pattern in B_+ . Each vertical line shows the perturbation to a single component of: a) catch, b) M , c) cpue catchability Q , and d) case weights. Perturbations are clustered by year and shown sequentially for ages 4–10. The values are $-h \times s_{\max}$ in per cent, with a) $h = 2.5$, b) $h = 6.5$, c) $h = 2.0$, and d) $h = 5.0$. The horizontal dotted lines are $\pm 100\%$.

difference in total annual biomass for ages 4–10 from the catch and M -perturbed SPAs is 260 000 t in 2002, which is 54% of the unperturbed value.

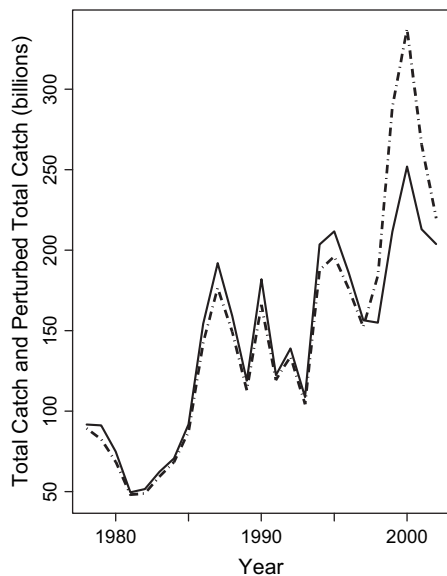


Figure 4. Total observed annual catch (solid line), and total perturbed annual catch (dashed line).

Catchability perturbations

Another interesting perturbation scheme involves the survey catchabilities. A common assumption which is used in the 4T fall herring SPA is that the cpue indices are proportional to absolute stock numbers-at-age (Equation 2). This constant-catchability assumption may not be true; for example, it is possible that cpue catchability changes over time due to changes in fishing practices, etc. Mohn (1999) showed that violations of the constant-catchability assumption could cause retrospective patterns. To assess the potential for this we examined influence diagnostics for multiplicative q -perturbations,

$$q_{a,y}(\omega) = q_a \omega_{a,y}, \quad (7)$$

where $\omega_o = 1$. This involved perturbations to unknown model parameters. We estimated the q_a s in Equation (7) using the perturbed fit function,

$$\begin{aligned} l_\omega(\theta) &= \sum_{a,y} \{ \log(r_{a,y}) - \log[q_{a,y}(\omega)] - \log[N_{a,y}(t)] \}^2 \\ &= \sum_{a,y} \{ \log(r_{a,y}) - \log(\omega_{a,y}) \\ &\quad - \log(q_a) - \log[N_{a,y}(t)] \}^2. \end{aligned}$$

The perturbed catchabilities were estimated as $\hat{q}_{a,y}(\omega) = \hat{q}_a \omega_{a,y}$. However, the $\omega_{a,y}$ s were fixed and obtained from the local influence diagnostics.

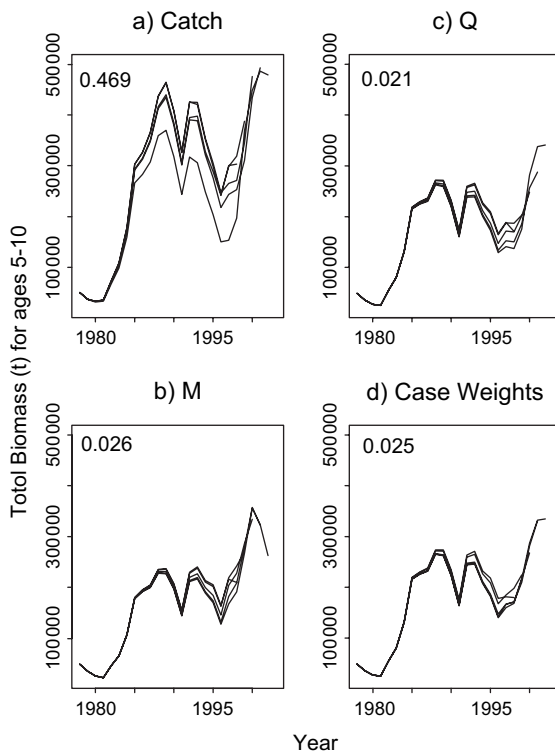


Figure 5. Retrospective estimates based on s_{\max} perturbations for: a) catches and $h = 2.5$, b) M and $h = 6.5$, c) cpue catchability and $h = 2.0$, and d) case weights and $h = 5.0$. The value of ρ is shown in the top left-hand corner.

The elements of s_{\max} are shown in Figure 2c. They suggest that a trend in the cpue catchability in 1996–2002 can reduce the retrospective pattern. The s_{\max} perturbations with $h = -2$ are shown Figure 3c. On a relative scale these are much smaller perturbations than those in Figure 3b. The estimates of q_a in Equation (7) were almost identical to the unperturbed estimates; for example, the unperturbed \hat{q}_{10} was 2.562×10^{-3} , while the perturbed estimate was 2.517×10^{-3} . Differences for other ages were of similar or smaller magnitude. The perturbed retrospective patterns are presented in Figure 5c. These perturbations greatly reduced the retrospective pattern, and improved the fit as well because the MSE decreased from 0.22 to 0.20.

Age specific retrospective patterns in estimates of population numbers-at-age based on the q -perturbations are shown in Figure 6. They are substantially improved when compared to the unperturbed results (see Figure 9 in Cadigan and Farrell, 2003) for ages 6 and 7 and marginally improved for ages 8–10. For age 5 the retrospective patterns are about the same magnitude but reversed in sign. The q -perturbed retrospective patterns are worse for age 4.

Case weight perturbations

Many methods for assessing influence involve the perturbation of case weights. A case weight is an extrinsic weight

of the squared residual in the fit function, which we perturb as

$$l_{\omega}(\theta) = \sum_{a,y} \omega_{a,y} \{ \log(r_{a,y}) - \log(q_a) - \log[N_{a,y}(t)] \}^2.$$

This perturbation scheme can be used to assess, for example, the impact of deleting cpue indices for a particular year. In this section we assess whether changes in case weights can reduce retrospective patterns.

The results in Figure 2d suggest that, compared to the catchability perturbations, larger changes to case weights are required to reduce the retrospective problem. This is because the values of $\hat{\rho}_{\max}$ are larger; however, smaller changes to case weights are required compared with C and M perturbations. Figures 3 and 5 demonstrate that case weight perturbations with $h = -5$ are sufficient to remove the retrospective pattern. Some individual perturbations are large; for example, the perturbed weights for the cpue indices at ages 9 and 10 in 2001 were zero, and the weights for ages 7–10 in 2002 were zero. Reducing the estimation weights for recent indices will obviously reduce the retrospective pattern; for example, if all indices in the retrospective “block” are zero-weighted then the retrospective pattern would be removed completely. This is not an interesting

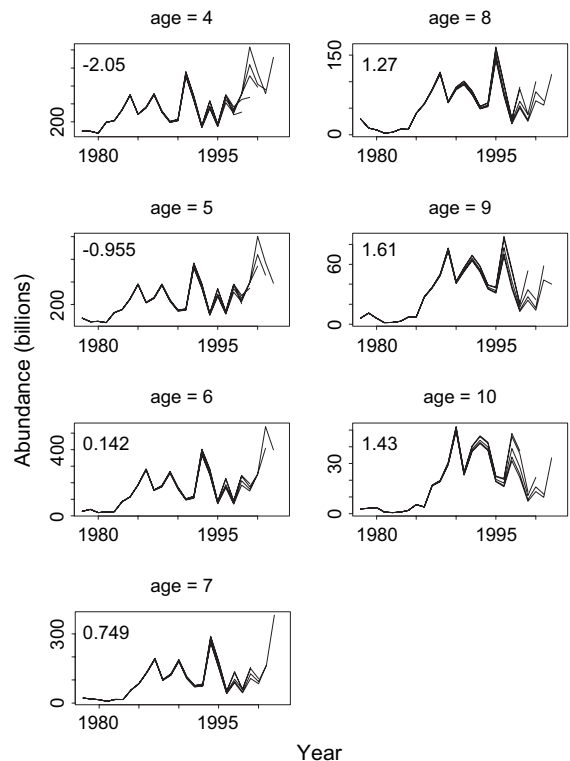


Figure 6. Retrospective estimates of 4T herring SPA stock abundance-at-age estimates based on s_{\max} perturbed cpue qs with $h = -2$. Each panel shows the results for an age. The retrospective ρ statistic is shown in the top left-hand corner of each panel.

perturbation, and may only mask the mis-specification problem.

Discussion

We have presented a practical methodology based on local influence diagnostics to assess the potential magnitude of changes in SPA inputs required to reduce or remove retrospective patterns. In the context of a data example, we illustrated how these methods can be employed to identify relatively small perturbations to SPA component inputs that result in greatly reduced retrospective patterns. The diagnostics can be used to assist in identifying the more likely sources of the retrospective patterns, or at least in eliminating components that are not the underlying source of the patterns. That is, if the smallest perturbation in a component that removes the retrospective pattern is unrealistic then we can reasonably conclude that the pattern is not caused by the component. Although this is obviously not as desirable as identifying the components that are at the root of the retrospective pattern, the determination of which components that are unlikely to have caused the problem is nevertheless useful.

We applied the proposed methods to the 4T fall herring SPA, and studied the influence of commercial catches, natural mortality, survey catchability, and estimation case weights on retrospective patterns. We concluded that relatively small perturbations to assumptions about the catchability of cpue indices could remove the retrospective patterns. Larger changes to case weights or the assumptions about M were required to remove the patterns, while our local catch perturbations were not able to eliminate them completely. This suggests that changes in cpue catchability are the most likely source of the retrospective patterns. Implicit in this conclusion is an assumption that the perturbation schemes are comparable. If the perturbation schemes are not comparable then the local slopes (e.g. $\hat{\rho}_{\max}$) are also not comparable and are not useful for determining which perturbations are more likely. For example, if uncertainty about M is much larger than uncertainty about whether cpue catchability is constant, then one might reasonably conclude that the M perturbations are more realistic and that incorrect assumptions about M are a more likely source of the retrospective pattern. We suggest *prima facie* that our perturbation schemes are comparable, but the plausibility of the perturbations, and the SPA perturbed stock estimates, are best assessed by 4T herring experts who are knowledgeable about the fishery and other scientific information for this stock.

Although our local influence analyses have produced a large amount of output to examine, the overall sensitivity of ρ can be summarized by the values of $\hat{\rho}_{\max}$. These slopes of ρ are 18%, 12%, 40%, and 19% for the C, M, q, and case weight perturbations, respectively. They suggest that smaller changes to the cpue catchability model are required to reduce the retrospective pattern than for the other SPA

model components. This is because the q slope is largest, which suggests that ρ can be reduced more for a fixed size q perturbation than for the other perturbation schemes. We repeated our analyses after excluding the 2002 data and basing ρ on retrospective stock sizes for 1997–2001. The slopes for C, M, q, and case weight perturbations are 32%, 18%, 60%, and 37% of $\rho = 0.98$, respectively. The per cent slopes are larger for the 1978–2001 series compared with the 1978–2002 series. This is partly because the value of ρ for the shorter series is smaller than for the full series (i.e. $\rho = 1.3$) but also because biomass estimates for the shorter series are more sensitive to perturbations. Nonetheless, the rank of the slopes is the same for the 1978–2001 series compared with the full series and our conclusion, that smaller changes to the catchability model can reduce ρ , is the same.

Retrospective-corrected stock scenarios are shown in Figure 7. Catch and M perturbations lead to somewhat different estimated stock trajectories, since they both play a very similar role in SPA by accounting for population deaths; however, an important difference in their roles involves the constraints on fishing mortalities at the oldest age A that is used in the 4T herring SPA. These constraints are commonly used and often make an SPA more sensitive to errors in catches at age A than similar sized errors in M at that age. This is more so when the SPA is not well “converged” (see Pope, 1972), which is the case for 4T herring. Also, the cpue, q, and case weight perturbations resulted in stock trajectories that are very similar to the unperturbed result. The conclusion we draw from this is that a retrospective-corrected SPA may not produce lower estimates of stock size. This is contrary to the common perception that a retrospective problem implies that the SPA will overestimate current stock size.

We focused our assessment of influence on the direction of maximum slope, s_{\max} , and the slope in this direction. We can also use the local slopes, given by $S(d)$ in the section on the Local influence approach, to examine the effects of perturbations in other directions. For example, if there was some indication that M doubled around 1995, then we can use the local influence diagnostics to predict the effect on ρ of this change simply by summing up the individual slopes for 1995–2002. The sum is -0.62 , and the local influence prediction of ρ based on a doubling of M in 1995–2002 is $1.30 - 0.62 = 0.68$, which is almost a 50% reduction in the ρ value. The re-estimated change in ρ is 0.64, and the influence-based prediction is close to this value.

Cadigan and Farrell (2003) found that the retrospective diagnostics were fairly similar when ρ was based on different stock quantities, although some differences were observed in the effect of the perturbations on retrospective patterns. One would expect the diagnostics based on total abundance and B_+ to be similar. Perhaps a better set of diagnostics would involve the numbers of young and old fish. In many applications spawner biomass would be an appropriate way to quantify the numbers of old fish. Another factor to consider is the number of years used to measure the

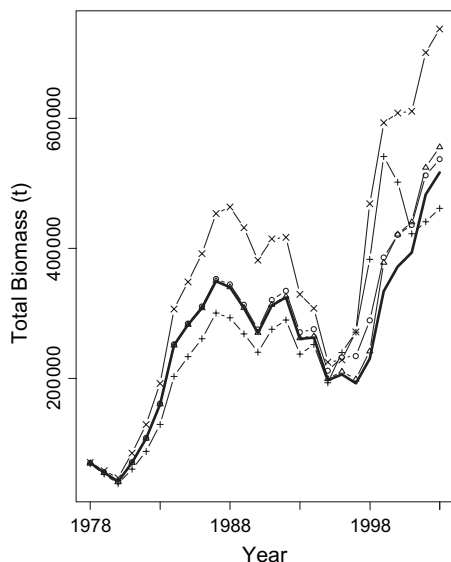


Figure 7. Comparison of estimated total biomass at ages 4–10: unperturbed (heavy solid line), M-perturbed (plus), catch perturbed (cross), q-perturbed (open triangle), and case weight perturbed (open circle).

retrospective pattern. Also, the retrospective metric proposed by Mohn (1999) and used in this paper (Equation 3) is a relative average which may not be appropriate in some cases, especially when some $S_{y,y}$'s are large.

It is possible that retrospective patterns are caused by mis-specifications of two or more model components. For example, the retrospective patterns could be caused by mis-reported catches and incorrect assumptions about cpue catchabilities. Local influence diagnostics based on perturbations to multiple components can be constructed directly from the diagnostics for individual component perturbations. For example, if $\dot{g}_{1,0}$ and $\dot{g}_{2,0}$ are the slope vectors for two perturbations ω_1 and ω_2 , then the direction of maximum slope for the combined perturbation scheme $\omega = [\omega_1, \omega_2]$ can be obtained using Equation (5) with $\dot{g}_0 = [\dot{g}_{1,0}, \dot{g}_{2,0}]$.

The methods we have proposed are related to minimizing $\|\omega - \omega_0\|$ with respect to ω , subject to $\rho(\omega) = 0$, where $\|x\|^2 = \sum_i x_i^2$. Note that $\rho(\omega)$ is based on SPA parameter estimates that minimize $MSE(\theta; \omega)$. In addition, we considered whether the mean square error (MSE) was reduced along with ρ when assessing perturbation schemes.

An alternative approach could be based on minimizing $MSE(\theta; \omega_0) + \lambda \|\omega - \omega_0\|$ with respect to both θ and ω , subject to $\rho(\omega) = 0$, for some specified penalty weight λ . With this approach, one searches for perturbations close to ω_0 that remove the retrospective pattern (as measured by ρ) and improve the fit simultaneously. Yet another approach could involve an unconstrained minimization with a second penalty term, $\rho^2(\omega)$, added to the objective function. Diagnostics based on improving fit as well as removing

the retrospective pattern may have increased power for detecting the correct source of mis-specification, if it exists. If λ was chosen to be very large then the first alternate approach would be equivalent to minimizing $MSE(\theta) = MSE(\theta; \omega_0)$ with respect to θ , subject to $\rho(\theta) = 0$. This would provide the basis for a statistical test of the hypothesis $\rho(\theta) = 0$.

The local influence diagnostics can be improved by utilizing more realistic perturbation schemes. For example, potential errors in catches may not be entirely multiplicative in nature because the magnitude of errors in large reported landings could be larger or smaller than the errors in small reported landings. The size of multiplicative perturbations to catch may not be directly comparable with the size of multiplicative perturbations to M because our uncertainties about catch may be more, or less, than our uncertainties about M . In practise it will be useful to have external information about the accuracy of SPA inputs and to use this information to assess the plausibility of perturbations. Also, the perturbation schemes we used may be too general in some situations, and possibly too sensitive to noise in the inputs as well as potential mis-specification signals. If the mis-specifications are believed to be smooth or more structured than have been modelled, then smooth or more structured perturbations may be more appropriate. Such modifications are relatively straight forward to implement within the diagnostic framework we have presented. The efficacy of the retrospective diagnostics also requires simulation testing.

Acknowledgements

The research work of Patrick Farrell is funded by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

- Cadigan, N. G., and Farrell, P. J. 2002. Generalized local influence with applications to fish stock cohort analysis. *Applied Statistics*, 51: 469–483.
- Cadigan, N. G., and Farrell, P. J. 2003. Local influence diagnostics for the retrospective problem in sequential population analysis. DFO Canadian Science Advisory Secretariat Research Document 2003/057. http://www.dfo-mpo.gc.ca/csas/Csas/DocREC/2003/RES2003_057_E.pdf.
- Cook, R. D. 1986. Assessment of local influence (with discussion). *Journal of the Royal Statistical Society, Series B*, 48: 133–169.
- Evans, G. T. 1996. Using the elementary operations of sequential population analysis to display problems in catch or survey data. *Canadian Journal of Fisheries and Aquatic Sciences*, 53: 239–243.
- LeBlanc, C. H., MacDougall, C., and Poirier, G. A. 2002. Assessment of the NAFO 4T southern Gulf of St. Lawrence herring stocks in 2001. DFO Canadian Science Advisory Secretariat Research Document 2002/053.

- LeBlanc, C. H., Poirier, G. A., Chouinard, G., and MacDougall, C. 2003. Assessment of the 4T southern Gulf of St. Lawrence herring stocks in 2002. DFO Canadian Science Advisory Secretariat Research Document 2003/040.
- Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. *ICES Journal of Marine Science*, 56: 473–488.
- Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. *ICNAF Research Bulletin*, 9: 65–74.
- Sinclair, A., Gascon, D., O'Boyle, R., Rivard, D., and Gavaris, S. 1991. Consistency of some Northwest Atlantic groundfish stock assessments. *NAFO Scientific Council Studies*, 16: 59–77.