Report of the retrospective analysis working group

Alaska Groundfish Plan Teams: Bill Clark, Dana Hanselman, and Mike Sigler

September 2012

The Plan Teams tasked a group of three members to review the value of adding retrospective analyses to Alaska fisheries stock assessments. We were fortunate that a comprehensive review already was available (LeGault 2009). In our review, we summarize the comprehensive review, recommend an approach for Alaska stock assessments, and show some examples for Alaska stock assessments.

Summarize LeGault (2009)

A workshop on retrospective patterns in stock assessment estimates was held in Woods Hole in 2008 (Legault 2009). The group considered different measurements of retrospective differences, causes of retrospective patterns, feasibility of identifying the causes of retrospective patterns, and feasibility of fixing retrospective patterns. The following definition of a retrospective pattern and the group's conclusions and recommendations are taken from the report.

Definition

A retrospective pattern is a systematic inconsistency among a series of estimates of population size, or related assessment variables, based on increasing periods of data (Mohn 1999). There are two types of retrospective patterns: historical and within-model. The historical retrospective analysis is conducted by examining the results of each final assessment for a number of years in a row and determining whether there was a consistent pattern of overestimating or underestimating assessment values in successive years. This type of retrospective pattern can be caused by changes in the data, type of assessment model, or assessment model formulation and is most important to managers because it relates directly to the management choices made in the past based on the information available at the time. In contrast, the within-model retrospective analysis uses the same data, type of assessment model, and assessment model formulation and trims the most recent year's data in successive model runs. The within-model retrospective patterns are most useful for determining an internal inconsistency in the data because the only changes in the different runs are the number of years of data in the model.

Conclusions of the workshop

1. A retrospective pattern is an indication something is inconsistent (data and/or model).
2. Lack of a retrospective pattern does not mean all is well. Based on simulations, data or model inconsistency does not always produce a retrospective pattern. Retrospective patterning is just one diagnostic to be considered when conducting stock assessments.
3. Simulated retrospective patterns can be caused by time trending changes in biological characteristics, catch, survey catchability, or spatial concentration of the population. Multiple sources may occur in assessments.
4. The source(s) of the retrospective pattern can be anywhere in the time series. Some methods were presented to identify when the change took place (moving window, q surface, mean square residual local inference surface).
5. The true source(s) of a retrospective pattern have not been identified using current methods.

This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the National Marine Fisheries Service and should not be construed to represent any agency determination or policy.

Knowledge of events in the fishery or biological information may help identify probable sources.
6. Interventions (correlated errors) are more likely to cause retrospective patterns than random noise.
7. Splitting surveys, changing *M*, or changing catch may reduce the retrospective pattern, but do not necessarily produce an assessment closer to the truth, although the other diagnostics for the new assessment may be fine.
8. The retrospective statistic, rho, may be a useful measure of the amount of retrospective pattern. A strong retrospective pattern can be defined by the degree of overlap between confidence intervals from different terminal years.
9. Local influence surface analysis using rho is not useful for diagnosing the timing or source of retrospective patterns. In many stocks, strong retrospective patterns typically persist.

Recommendations of the workshop

1. Always check for the presence of a retrospective pattern.
2. If a model shows a retrospective pattern, then consider alternative models or model assumptions.
3. Develop objective and consistent criteria for the acceptance of assessments with retrospective patterns.
4. A strong retrospective pattern is grounds to reject the assessment model as an indication of stock status or the basis for management advice.
5. When a moderate retrospective pattern is encountered: (not an exhaustive list)
    a. Consider alternative states of nature approach to advice.
    b. Investigate the performance of alternative methods for retrospective adjustments through management strategy evaluations.
6. Use biological and fishery hypotheses and auxiliary information as a basis for adjustments for retrospective patterns.
7. Consider use of survey swept area numbers instead of mean catch per tow in assessment models.
8. The presence and implications of a retrospective pattern as a source of uncertainty in the assessment should be clearly communicated to managers.

Recommendation for stock assessment authors

There is currently not an accepted level of retrospective bias beyond which an assessment is deemed to exhibit a retrospective pattern (LeGault 2009). Simulation exercises specifically designed to mimic the level of uncertainty in the assessment data may reveal how often a pattern might be expected in real assessments could provide guidance. However, this approach would be labor intensive and be done for each model formulation of a specific assessment (LeGault 2009). Instead, the typical approach has been to focus on within-model effects, look at the plots and make a subjective decision based on the number of years which deviate from the full time series assessment in the same direction.

For Alaska groundfish assessments with age-structured models (Tier 3 and higher),we recommend that a retrospective analysis be presented as part of the model evaluation. Specifically, stock assessment authors are requested to conduct the within-model approach and rerun the model, successively dropping data one year at a time. Specifically the analysis should include:

1. Running retrospectives back to 2002 (where 2002 would be a terminal year) for the base-case assessment in 2012 (i.e., drop 10 years of most recent data);
2. Plotting spawning biomass time series for each model run;
3. Plot of relative changes referenced to the terminal model run.

We envisage having this for all full assessments presented for the 2012 November-December Council review cycle. Example plots requested are (from Legault 2009):

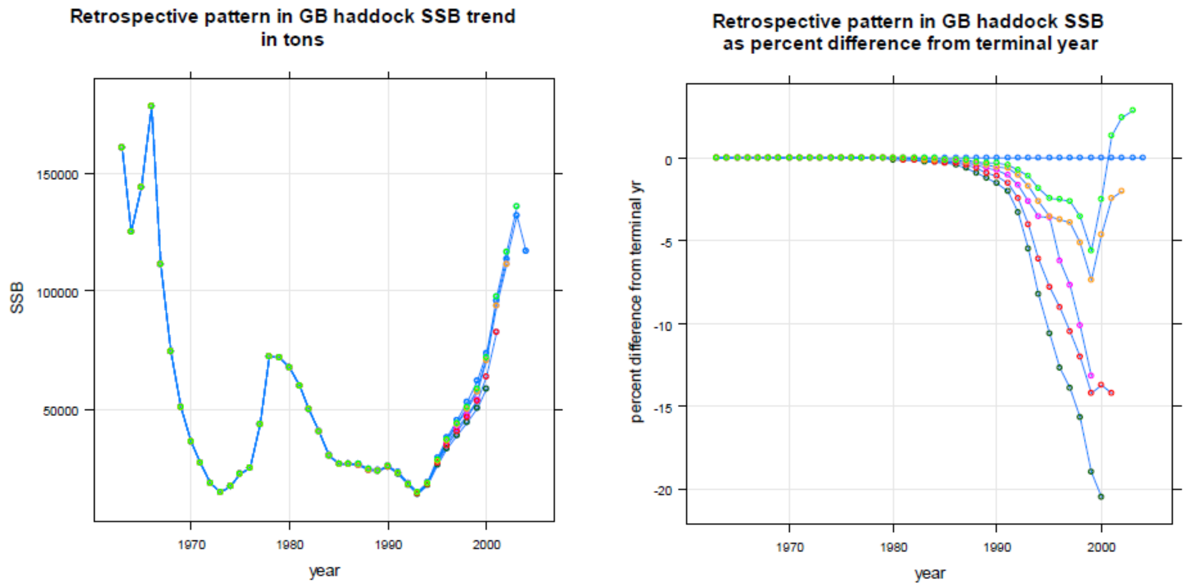

Figure 4. Retrospective plots for Georges Bank haddock, standard plot on left and relative plot on right.

Examples

We include several examples of the recommended analyses (Figs. 1-3). In these examples, there appear to be two distinct types of retrospective pattern. For Gulf of Alaska Pacific ocean perch and northern rockfish (Figs. 1 and 2), the retrospective patterns in the assessment results appear to be data-driven. The retrospective change in estimates occurs only when a new trawl survey is included (every two years). Both of these species have had relatively large, imprecise increases in trawl survey biomass estimates since 1999. However, the two species age structures are quite different. The bulk of the increase in survey biomass for the northern rockfish population appear to be driven by increases in the abundance in the older age-classes, while the bulk of the increase in Pacific ocean perch survey biomass is driven by increased abundance of the younger age classes. Increased survey abundance in the older age-classes over time is the likely cause of the changes in the historical estimates (1960-1990) of northern rockfish, because as each successive survey is added the model increases the historical abundance to explain the increased abundance in the older age classes that was observed in the survey. Alternatively, changes in historical abundance estimates is not as drastic in the Pacific ocean perch model, rather estimates of abundance change in recent years are likely due to observations of younger fish in the survey that influence estimates of recruitment in the model. These data-driven examples exhibit retrospective patterns that probably are of little concern.

Another type of retrospective pattern is illustrated by Alaska sablefish (Fig. 3). This retrospective pattern is unlikely to be considered severe, but at issue is the “one-way” pattern. As data are added, the recent estimates of spawning biomass decrease slightly for each of the 10-year retrospectives. This contrasts with the rockfish examples because the patterns of fits relative to the survey indices varied (both increasing and decreasing). This is likely because the Alaska sablefish model integrates a larger number of datasets (e.g., the model has three abundance indices compared to one for the rockfish examples). Also, in the last several years the magnitude of the retrospective pattern seems to have dissipated. It is difficult to isolate the cause of this pattern but several possibilities exist. For example, hypotheses could include environmental changes in catchability, time-varying natural mortality, or changes in selectivity of the fishery or survey.

While the patterns in these examples may be acceptable, it would be worthwhile for assessment authors to evaluate which parameters or model configurations that might best be altered to remove the pattern. Although, as advised in Legault (2009), isolating a parameter that fixes the retrospective pattern does not necessarily warrant changing that parameter in the model or that the fixed model is any closer to the 'truth', but is a good exercise in model exploration and sensitivity.

If an approach and the final model specification fail to correct a strong one-way retrospective pattern, then we recommend that this be highlighted as a rationale for potentially setting the recommended ABC below the maximum permissible value if the model-driven pattern is biased high. Conversely, a strong retrospective pattern that is consistently biased-low could be used as evidence to set the ABC at maximum permissible despite other evidence of low stock size. The Plan Teams will need to review these retrospective analyses across Alaska stocks to determine what constitutes a "strong" one-way pattern.

Acknowledgement

We thank Pete Hulson for providing the rockfish retrospective runs and figures.

References

Legault, C.M., Chair. 2009. Report of the Retrospective Working Group, January 14-16, 2008, Woods Hole, Massachusetts. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 09-01; 30 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/

Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56: 473-488.

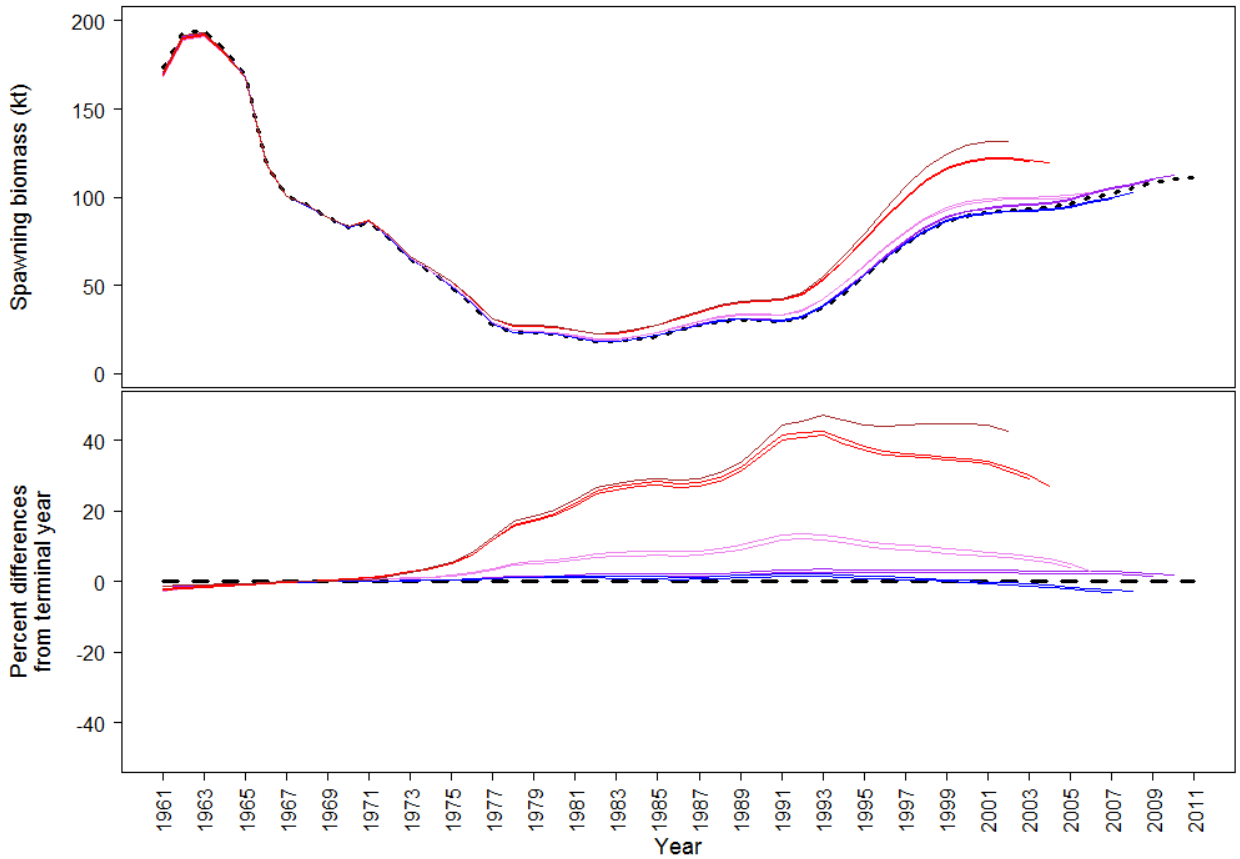

Figure 1. Within – model retrospective plots for Gulf of Alaska Pacific ocean perch. Top panel is absolute change in female spawning biomass. Bottom panel is the relative difference in each year to the terminal year estimates. Black dashed line is the terminal year estimates, while pairs of the same color are based on the same survey results due to biennial surveys.

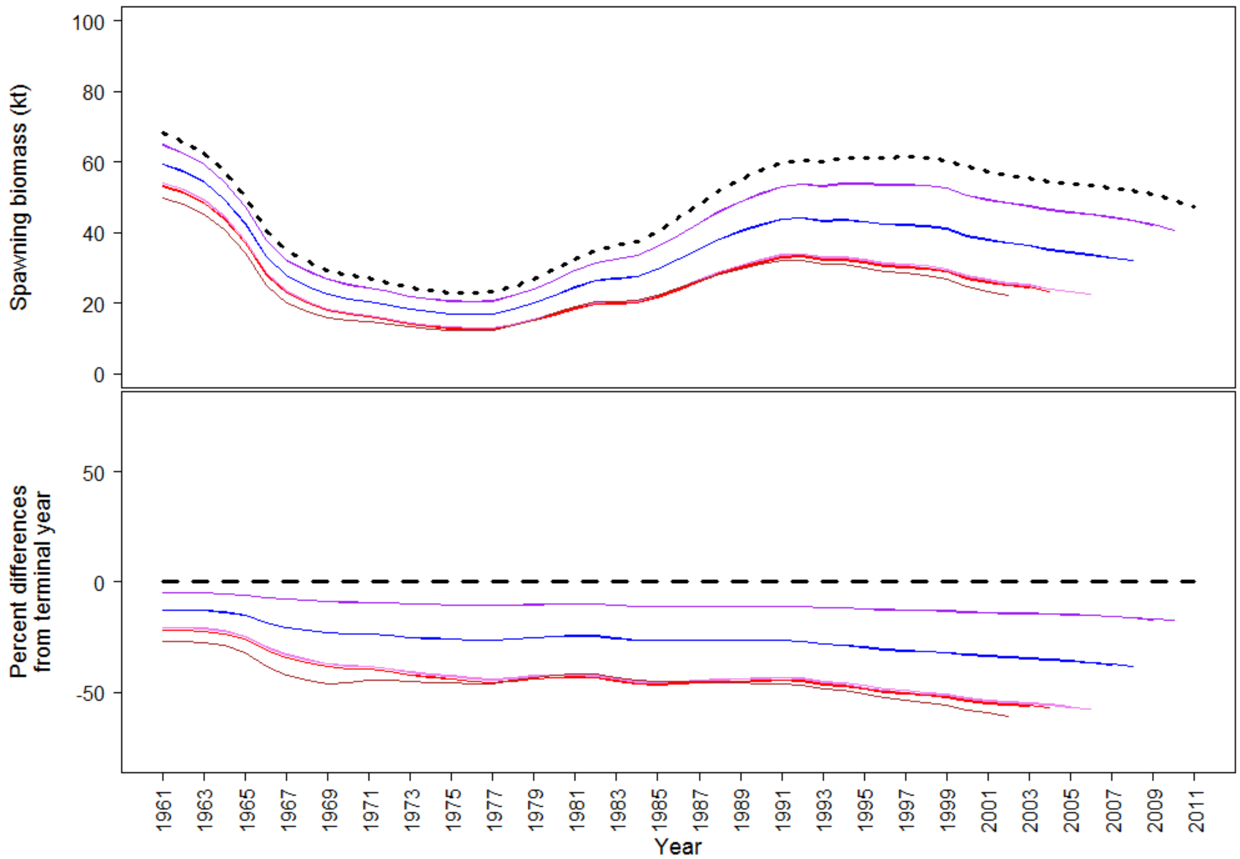

Figure 2. Within – model retrospective plots for Gulf of Alaska Pacific northern rockfish. Top panel is absolute change in female spawning biomass. Bottom panel is the relative difference in each year to the terminal year estimates. Black dashed line is the terminal year estimates, while pairs of the same color are based on the same survey results due to biennial surveys.

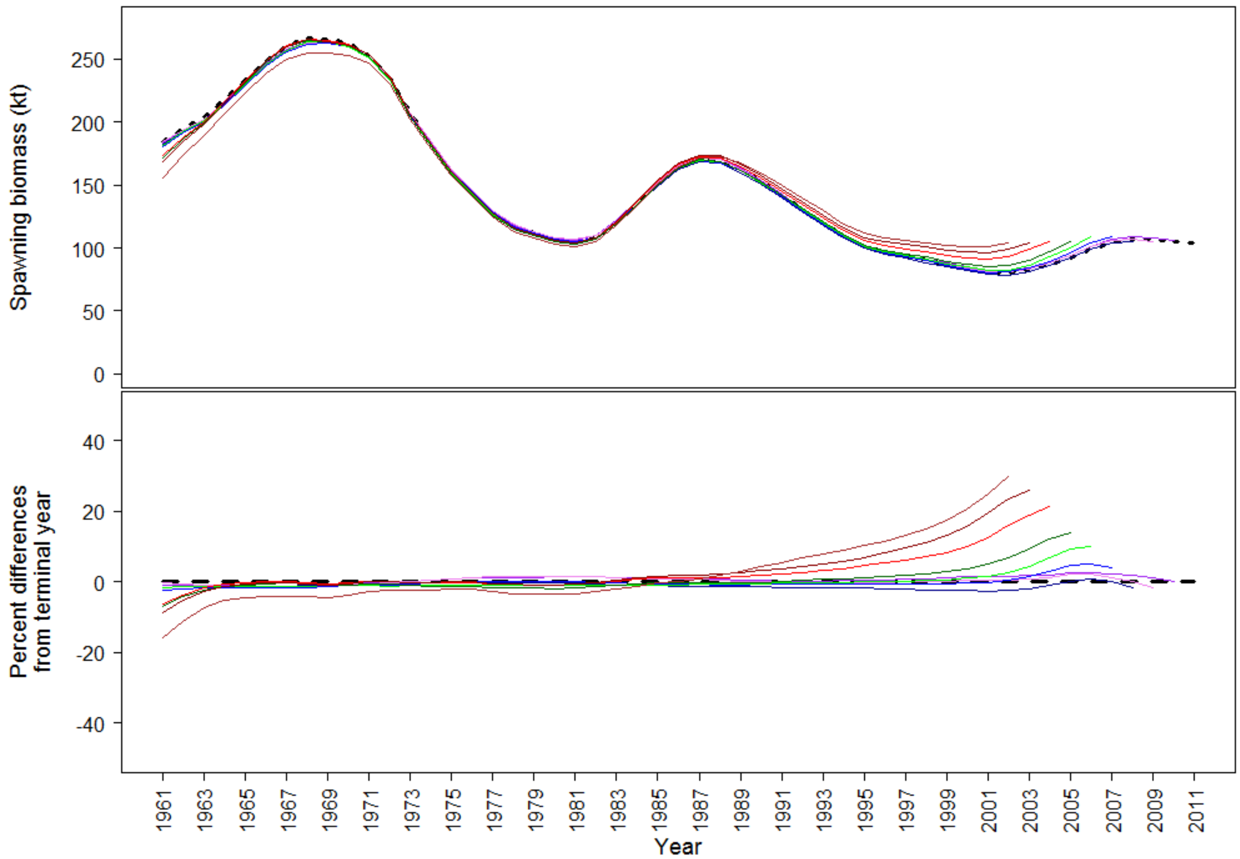

Figure 3. Within – model retrospective plots for Alaska sablefish. Top panel is absolute change in female spawning biomass. Bottom panel is the relative difference in each year to the terminal year estimates. Black dashed line is the terminal year estimates.