Atlantic States Marine Fisheries Commission

Healthy, self-sustaining populations for all Atlantic coast fish species or successful restoration well in progress by the year 2015

Development and Use of Reference Points

Written By:
Kim McKown
Steven Correia
Matt Cieri

Assessment Science Committee

December 2008

Table of Contents

I. Introduction 3
II. Terminology 3
III. Reference Point Regulations and Requirements 6
A. ASMFC 6
B. Federal 7
C. Management Oversight of ASMFC Species 7
IV. Development of Reference Points 7
A. Purpose 7
B. Methods 8
Supporting Analyses
1. Trend Analysis 10
2. Smoothing Techniques 10
3. Continuity 11
C. Indicators of Stock Health 11
V. Control Rules 13
A. Targets and Thresholds 13
B. Concept of control rules and overfishing/overfished definitions 14
C. Caddy's Traffic light approach 15
D. Rago's Non-Parametric multi-attribute assessment methodology 15
E. Precision and Bias 16
F. Uncertainty around reference points 16
G. Retrospective bias 16
VI. Summary 17
VII. References 18
Tables 21
Figures 24

I. Introduction

The purpose of this document is to provide a useful and concise guide to ASMFC technical committee members and other fisheries scientists conducting stock assessments in their efforts to determine reliable indicators of stock status. Several summary articles on reference points are available in the fisheries literature (Mace 1994, Rosenberg et al. 1994; Caddy and Mahon 1995; Caddy 2004). This document does not intend to repeat those efforts, but provides a useful guide for fisheries scientists on the options, methods, and assumptions related to the development of reference points, particularly for species undergoing first time assessments and for 'data poor' species.

Another aspect of reference point calculation pertinent to the ASMFC is that the Commission and the National Marine Fisheries Service manage several species jointly. In these cases, requirements for reference points are defined in the Magnuson-Stevens Act. For species managed by the ASMFC, these requirements do not exist. This permits greater flexibility in selecting reference points, but can also lead to confusion or even debate over reference point selection. In both cases (ASMFC managed species and jointly managed species), developing reference points that are pertinent to the stock, consistent with the assessment method, and clear to managers, is important.

This report begins by defining several terms related to reference points, and then examines requirements for reference points. The report concludes with a review on the calculation of reference points and control rules.

II. Terminology

When managing a fishery, managers usually set control rules which define predetermined management actions to be taken when a stock meets or exceeds a certain status indicator (e.g., reference points). Control rules allow for clear management actions to be made depending on the biological status of a stock as determined by comparison to reference points. Reference points are signposts indicating the desired state of the stock (targets) and marking the boundary of undesirable stock conditions (thresholds). Managers gauge the status of the stock relative to reference points that are commonly expressed as fishing mortality (rate of removal of stock by fishing operations) and stock size (reproductive capacity, often expressed as spawning stock biomass or abundance), then take action following the control rules, if necessary. Management action is typically required when fishing mortality exceeds a threshold reference point (end overfishing), or when biomass falls below a threshold reference point (rebuild). Target reference points define desirable fishing mortality rates or stock sizes managers aim to achieve for a particular stock.

Fisheries managers determine the amount of harvest to remove from a stock in order to achieve desired goals. One goal may be to optimize stock abundance for a variety of biological (avoid low spawning biomass, large fluctuations in biomass), social (maintain fishing communities), and economic reasons (avoid large fluctuations in harvest/revenues, maximize long-term revenues). In practice, this results in adjusting fishery removals, as other sources of mortality are difficult or impossible to control directly (i.e., natural mortality). In fisheries management, this is termed Optimum Yield or OY, the amount of removals for a particular stock that balances these often-competing needs (See also Section III.B).

In this guide, we distinguish between *model-based* reference points and *index-based* reference points used to define control rules for fish stocks. In practice, the difference between the two types of reference points may be ambiguous, as assessment scientists utilize a combination of available information (often both qualitative and quantitative results). For example, a model may provide estimates of recruitment, which in turn are used to derive SSB reference points using an empirical distribution of recruitments. An index-based method may apply a distribution of recruitments measured from a survey index to derive a biomass reference point in survey units. Both model-based and index-based methods have advantages and disadvantages for management, as well as varying amounts of uncertainty. Neither type of reference point is more valid for management purposes in all circumstances. Model-based reference points tend to be favored, because they are typically based on multiple sources of information integrated into a model framework, and uncertainty is often quantifiable.

Model-based reference points and their methods are derived from intrinsic and stock-specific biological characteristics such as growth, maturity, and stock-recruitment relationships. The reference points are estimated through population modeling and may include estimates of carrying capacity, population growth rates, spawning stock biomass per recruit, yield per recruit, and others. These reference points tend to be the more familiar F_{MSY} (fishing mortality rate resulting in maximum sustainable yield) and $\mathrm{B}_{\mathrm{MSY}}$ (biomass associated with producing maximum sustainable yield) produced by age-structured, surplus production, or other types of population models currently in use. Inherent in model-based reference points is the assumption that fishing has a direct impact on the reproductive potential or abundance of a stock, resulting in predictions of a population's trajectory and long-term average status under differing management regimes.

Model-based quantitative reference points have the advantage of quantitatively relating fishery removals to stock biomass. Catch, therefore, can be directly translated into stock abundance both in short-term projections and over the long term through stock-recruitment relationships and projections. In addition, the relationship between catch and other factors (stock biomass, recruitment, revenues) can be examined, providing managers with options when setting OY, and describing the implications of several harvest strategies. In general, model-based reference points are advantageous because they integrate a variety of information - catch or landings (by age or in total), fishery-independent or -dependent indices, and population growth and survival rates.

Model-based reference points have some disadvantages. Chief among these is the use of *long-term* biological values such as carrying capacity, intrinsic growth rates, spawner-recruit relations, and others. Most of these values, in reality, are constantly changing because of shifts in predator-prey relationships or the availability of suitable habitat, or large-scale climate changes. In most modeling frameworks, these values are based on an assumption of stationarity, that is, particular features (natural mortality (*M*), the intrinsic rate of increase (r)) vary randomly around a 'historical average' of the time series modeled. For example, an assumption about stationarity in natural mortality rates at age is the basis for the 'constant *M*' assumption in models. The annual *M* varies without trend around the long-term mean *M*. Inter-annual variability in population characteristics is a less important consideration for model-based reference points than moderate to long-term trends or non-stationarity in these characteristics. When these conditions change over the long term, model-based reference points, which reflect past environmental conditions, may no longer be applicable under current conditions. The reference points may not reflect recent changes in productivity, natural mortality, growth, or other important variables controlling the stock's absolute abundance or biomass.

In some cases, model results may seem incredible. For example, model-based reference points will describe stock conditions that have not been observed over the timeframe used to derive the reference points. This often happens when data are only available for the most recent time period of a fishery with a long history of exploitation. For example, the Gulf of Maine cod stock has landings data back to the 1940s, and survey (1960s), recreational catch (1980s), and discard data (1989) but a fishery has existed on the stock since the early 1600s. This poses a challenge for managers attempting to balance the social, economic, and biological needs of the fishery. Generally, this is less critical for fishing mortality reference points than biomass reference points. Biomass reference points lying outside of observed biomasses can be more problematic in cases where control rules mandate a defined rebuilding period as density-dependent effects may prevent reasonable attainment of management goals. In contrast, estimating reference points from a short time series for a heavily exploited stock is a form of recency bias that could result in long-term loss in yield and revenues. Finally, these reference points are often derived from complex modeling that may be difficult to understand without an extensive quantitative background. Effective communication on the scientific rationale behind the reference point (including uncertainties) is critical in promoting acceptance by the public and managers.

The methods used to develop model-based reference points range from fully integrated statistical catch at age models with reference points (F_{msy} and B_{msy}) estimated within the model; to assessment models with reference points estimated extrinsically to the model; to mixed models where a reference point such as $F_{40\%}$ serves as a proxy for F_{MSY} and SSB_{MSY} and is derived from empirical recruitments (e.g., mean recruitment) and SSB per recruit estimates.

Index-based reference points and their methods are typically derived using a summary statistic (mean, median, quantile) from the empirical distribution of historic observations of the population. The time series frequently used include historical catches, catch rates (CPUE), and fishery-independent indices or relative exploitation rates (catch divided by survey index). But these reference points are not derived from modeled population estimates. They are generally used in data-limited situations or in cases where models cannot provide robust reference points. Using index-based reference points to determine the catch that will achieve desired management results is usually difficult, because performing projections may be challenging, if not impossible.

Index-based reference points tend to be more flexible than model-based reference points and are often easier to understand. Managers have the ability to set reference points that reflect more recent stock conditions, or recent conditions in the fishery, to better balance the competing biological, social, and economic needs. Because they are set in close consultation with managers, index-based reference points may be more readily accepted by managers and the public in comparison to model-based reference points. In general, index-based reference points and their calculations are more easily understood by the layperson.

Index-based reference points have disadvantages as well. They serve as rough proxies for true/model-based reference points, but the relationship to true/ model-based reference points is unknown. Particularly, they suffer from an inability to link stock removals to future abundance. However, they provide landmarks useful for developing management advice in an uncertain environment. For example, managers may know current biomass is low compared to reference points, but scientific advisors may not be able to determine the harvest reduction needed to rebuild the stock or to stop overfishing. However, advice to move toward lower exploitation may be robust even though the magnitude of the exploitation needed to rebuild is unknown.

A second disadvantage is that index-based reference points tend to be based on limited information or data. If the desired data had been available, scientists would have calculated model-based reference points. Index-based reference points lack the more rigorous quantitative backing of model-based reference points. This can make decisions politically difficult if reductions are needed, and social and economic sectors might call into question the validity of perceived ‘arbitrary’ measures of stock status and reference points. Managers may be tempted to use the uncertainty in reference points as rationale for selecting lower biomass targets and higher F targets. Finally, these proxy reference points may be re-evaluated in another modeling framework as more information becomes available. Changes to reference points may be seen as “moving the goal post” in one direction or another.

Regardless of the types of reference points and methods available, managers should be aware of the variability, accuracy, and precision of these measures (see below under Precision and Bias). Managers need to take a precautionary approach when setting F_{target}, TAC (total allowable catch) or OY, and when initiating control rules, particularly when rigorous analyses are not possible due to data constraints. This approach will improve the likelihood of not exceeding the F reference point or going below the $B_{\text{threshold}}$. The amount of precaution should be based on the uncertainty associated with the type of reference points or methods chosen.

III. Reference Point Regulations and Requirements

A. ASMFC

Goal 1 of ASMFC is to rebuild and restore depleted Atlantic coastal fisheries, and maintain and fairly allocate recovered fisheries through cooperative regulatory planning. One strategy to achieve this goal is to monitor the progress of cooperative fisheries conservation and management programs of the states relative to stock status.

The ASMFC Guidance Document for ASMFC Technical Support Groups includes a section on general guidance for addressing uncertainty. The guidance recommends technical committees include uncertainty in all assessment documentation. One of the specific recommendations is the use of target and threshold quantitative reference points to evaluate the status of stocks. The guidance does not specifically recommend types of quantitative reference points (unlike fish stocks managed jointly by the ASMFC and the National Marine Fisheries Service). This provides for more flexibility in the choice of reference points for ASMFC managed species. It also allows for the development of index-based reference points in lieu of traditional model-based reference points (fishing mortality, reproductive capacity) since the estimation of maximum sustainable yield and related reference points are not required for ASMFC managed species.

B. Federal

The Magnuson-Stevens Act is the federal law defining the standards and requirements for federal fishery management plans. The Act calls for all federal management measures to prevent overfishing and achieve the optimum yield for each fishery. Optimum yield is established by determining the maximum sustainable yield of a stock that can be harvested, and then reducing this level of harvest to account for biological, economic, and social considerations. Federal fisheries must "assess and specify the present and probable future condition of, and the maximum sustainable yield and optimum yield from, the fishery". Therefore, federal fisheries assessments must calculate reference points (F_{MSY} and B_{MSY} or proxies thereof) allowing scientists to compare current status to these values. F_{MSY} is the rate of fishing and B_{MSY} is the stock size producing the maximum sustainable yield under equilibrium conditions. The law requires that plans have measures to prevent overfishing and rebuild depleted or overfished stocks.

Amendments made into law in 2007 require the regional Federal Management Councils to prepare and implement a plan to end overfishing within two years of an overfishing determination. This mandate takes effect 30 months after enactment. The maximum rebuilding time of 10 years (if able to rebuild with F=0 in less than 10 yrs; otherwise sum of generation time plus rebuilding time at F=0) for overfished stocks remains in place, except for summer flounder. Exceptions to the 10 year maximum rebuilding period exist for cases where the biology of the fish, other environmental conditions, or international management agreements dictate otherwise.

C. Management Oversight of ASMFC Species

Table 1 identifies the management authority and types of reference points for each species managed by ASMFC.

IV. Development of Reference Points

A. Purpose

Reference points are typically used to describe the stock status relative to the fishing mortality rate and some measure of stock size (number (abundance), total weight (biomass), spawning stock biomass, or exploitable biomass). Mortality rate based reference points define when 'overfishing' is occurring on a stock. Abundance/biomass based reference points determine when a stock is 'overfished' or stock size is unacceptably low. The primary purpose of an absolute minimum biomass threshold is to provide complete protection to a seed stock (i.e., reproductive capacity) to allow recovery in the event a stock becomes exceedingly low in spite of other management actions (Rosenberg et al. 1994). Control rules can be designed to prevent biomass from falling below the threshold, thus reducing the risk of low stock productivity over long periods. Three categories of reference point levels are typically employed in fisheries management:

Limit Reference Point (LRP):

1. Indicates an undesirable state of a fishery and/or resource which management action should be taken to avoid (Caddy and Mahon, 1995).
2. Sets the safe upper limit of fishery exploitation (Caddy and McGarvey, 1996).
3. The perceived maximum degree of safe exploitation for a stock. A LRP should rarely be exceeded (Prager et al., 2003).

Thresholds:

1. A line in the sand; a good possibility the stock will decrease if we go beyond this level (e.g., F_{collapse}).
2. A 'red area' where continuity of resource production is in danger and immediate action is needed. A threshold reference point indicates when such a danger area is about to be entered.
3. Selection of the appropriate threshold should depend on how quickly the stock can be expected to recover if the threshold is crossed (Rosenberg, 1994).

Target Reference Point (TRP):

1. Indicates a desirable state of a fishery and/or resource; management action, whether during development or stock rebuilding, should strive to reach this state (Caddy and Mahon, 1995).
2. Guiding a fishery to a desirable management target (Caddy and McGarvey, 1996).
3. A TRP uses the same metric as the corresponding LRP and defines the degree of exploitation strived for under management. TRP and LRP constitute margins of safety preventing frequent occurrences of exploitation beyond the LRP and thus promoting sustainability (Prager et al., 2003).

Limit, threshold, and target levels of reference points need to be distinctly different to provide clear guidance for management decisions. In the U.S., some limits and thresholds are synonymous (e.g., the F_{MSY} threshold is a limit reference point).

B. Methods

Model-based methods used to develop reference points are more data intensive than index-based methods. To model any population, required data include fishery-independent and dependent indices, catch (by age and/or in aggregate), measured or assumed estimates of natural mortality rates, and population growth rates. Although many of these values can be estimated from fishery-independent or dependent data, direct measurements increase confidence in the model, and in resulting management advice. Most models require extensive fishery-independent and dependent sampling, time to develop species- or stock-specific variables, and a high degree of expertise.

For example, many models estimate maximum sustainable yield (MSY). Managers often use this value and set an OY below MSY because of relevant biological factors such as predation, or because of variability or uncertainly in the population model or data.

Such models function best with longer time series of data, broader geographic coverage of fishery-independent or dependent data, and when landings and or fishery-independent indices show contrast (periods of both high and low abundance). In general, the contrast should have multiple episodes of increasing and decreasing population/exploitation rates. The classic 'one-way trip' is unlikely to be informative relative to model-based reference points, particularly with respect to sustainable exploitation rates. Greater contrast and reliable input data will result in more robust estimates of model-based reference points.

Note that population models are only as good as the data and assumptions used as inputs. Models and resulting population status relative to model-based reference points sometimes give false confidence to managers with respect to reliability because they are "more quantitative". With any model, sufficient diagnostics and uncertainty estimation should be expressed to managers whenever possible, in addition to standard model outputs.

Index-based methods used to develop reference points require fewer data and are often used in data limited situations. In general, they are likely going to be based on summary statistics (mean, median, quantile) of data such as average catch, 75^{th} quantile of the survey index, or median exploitation ratio (catch divided by survey index) from an empirical distribution (or time series) of data. Even in the absence of fishing mortality estimates or landings, management may use relative biomass reference points for deriving management advice. Like model-based reference points, index-based reference points function best with multiple data sources and longer time-series of data. Implementing control rules and basing decisions on a single survey that exhibits large short-term fluctuations may be difficult, but adding information on age- or size-structure, recruitment, changes in catch and effort, or additional surveys will improve decision making if multiple indicators show similar trends. As with model-based reference points, index-based targets and thresholds need to be distinctly different to provide meaningful information about the status of the stock and fishery.

For example, if landings and a fishery-independent index are available, managers may set a target biomass reference point based on the median of that index across the time series. Managers may also set a target fishing mortality reference point using a proxy, such as the median of catch divided by biomass (a relative exploitation rate).

Fishery dynamics and the population dynamics of a stock are inherently included in the calculation of model-based reference points. Fishing mortality and stock size reference points synthesize much information into a few numbers (e.g., $F_{MSY,}$ B_{MSY}), while index-based reference points are generally set for groups of individual data sets. Thus, the utility of index-based reference points is improved with the addition of more data sources. One method uses a 'traffic light approach' (see Caddy's Traffic Light Approach in Section III). In this type of management régime, abundance, recruitment, age structure, and other fishery-independent or -dependent information is used to assess the current status of the stock. In utilizing multiple sources of information, indicators of stock status may be more informative when compared with reference points or methods based on a single data source.

Supporting Analyses

1. Trend Analysis

The primary interest in trend analysis is to determine whether most of the evidence from fishery-independent surveys supports a hypothesis of population change. Population decline is of particular concern. Trend analysis results can be combined using multivariate techniques. Multivariate statistical techniques such as Principal Component Analysis (PCA) can be used in finding commonalities, trends, and differences among multiple independent surveys.

Standardizing survey indices allows for comparing trends among surveys using the same units. A number of methods exist for standardizing indices. Multiple surveys can be transformed to similar scales using simple ratios of annual indices divided by the long-term mean, or by using z-transformations. These are useful in making direct comparisons among survey indices when time series are the same length, but comparisons can be less useful and even misleading when various time series are of different length.

Annual survey index values can be influenced by various factors (e.g., T°, depth) unrelated to changes in abundance. Statistical approaches are available for standardizing indices to account for variation in these factors. Common approaches are generalized linear models (GLM) (McCullagh and Nelder, 1989), generalized additive models (GAM) (Hastie and Tibshirani, 1990), and mixed models (Pinheiro and Bates, 2002). These techniques can be used to produce indices that are standardized relative to a number of factors (year, gear, time of year, bottom type) or covariates (depth, temperature, latitude-longitude) that influence catchability.

2. Smoothing Techniques

Smoothing techniques are useful for visualizing underlying trends in a time series despite the noise produced by annual variability in environmental and other factors. A number of methods exist for smoothing data. Running averages is a simple approach, where each annual data point is the average of several years (such as 3, 4, 5, 10 year intervals). The number of years to average depends on the periodicity of the trend you are looking at, the biology of the species, and the implications of truncation at one or the other tails of the time series. A longer time frame might be used when looking for trends related to environmental change, while short-term trends in abundance due to recruitment and management measures would utilize a shorter time frame.

Several statistical approaches are available for smoothing time series to detect trends (e.g., loess, splines, and GLM (such as delta lognormal) methods (Lo et al, 1992)). Time series methods such as spectral analysis, ARIMA, ARMA, and seasonal loess time series decomposition can be used to smooth time series and separate seasonal components from long-term trends. Smoothing can also be used for evaluating relationships between two or more variables. An example of using loess smoothing is presented in Figure 1 and compared with a Beverton-Holt stock-recruitment function. Smoothing can be an exploratory data analysis technique useful in evaluating the relationship between two variables. For example, loess smoothing is used to assess the shape of the relationship between stock and recruit of Massachusetts DMF windowpane flounder survey indices (Figure 1). In this case, the form of the stock recruit relationship indicated by the loess fit is consistent with a Beverton-Holt stock-recruitment function. The use of localized smoothing on stock and recruit data, combined with knowledge of a species' biology, can be a helpful guide to selecting candidate stock and recruit models used in deriving reference points.

3. Continuity

For assessment models where reference points are internally estimated for each model update, reference points need to be updated. For models where reference points are externally estimated, decisions must be made on whether to revise reference points with each assessment update. The reference points must be re-examined when a new assessment model is used (e.g., striped bass in ASMFC's 2007 assessment), or when reference points are calculated using a new approach (lobster in 2005). Decisions must be made on whether to revise reference points when modeling and reference point approaches are unchanged, and only new data are added to the assessment. Using the most up-to-date and accurate reference point value from the recent assessment is an intuitive approach. However, frequently changing reference point values creates moving targets managers may find difficult to achieve, potentially resulting in diminished public buy-in. In general, a reference point should be considered a point estimate (with an associated coefficient of variation) representing long-term conditions. Minor reference point adjustment shouldn't be considered in response to small annual variations in selectivity, weights at age, or the addition of one or two year classes to the stock-recruitment history. However, reference points may need updating when large changes in inputs such as fishery selectivity or long-term changes in productivity (faster or slower growth) have a large effect on the value of reference points.

C. Indicators of Stock Health

Identifying Useful Indicators

A number of stock indicators can be examined to interpret stock status relative to reference points, in addition to fishing mortality and population biomass parameter estimates. These stock indicators are used to corroborate or refute model results and provide additional information about the overall health of each stock.

Multiple indicators bring together a variety of monitoring results, traditional stock assessment model results, and fishery indices to form management advice. This offers transparency in methodology and purpose to all stakeholders. The advantage of using a multiple indicator approach is that any single indicator is associated with uncertainty as to what change means relative to stock status. But, additional indicators, especially when obtained from independent observations, and when considered collectively, will tend to reflect the true state of the stock (Caddy, 2004; Koeller et al., 2000). One disadvantage of indicators is they only represent the state of being over the time series of data, which may not reflect the stock's entire range of productivity. Another disadvantage is the same trend in an indicator may have multiple and differing interpretations. For example, the concentration of biomass in a single year class could indicate (1) a population with truncated age distribution caused by high mortality, or 2) the influence of an exceptional year class in a healthy stock with a broad age distribution.

The following are population and fishing indicators useful in assessing the health of a stock:

1) Abundance
   a) Total abundance
   b) Spawning stock abundance (this is a proxy typically used for reproductive capacity)
   c) Recruitment
   d) Geographic distribution
   e) Structure of abundance
      i) Size or age composition
      ii) Sex ratio
      iii) Maturity at age

2) Mortality
   a) Total mortality (Z)
   b) Fishing mortality (F)
   c) Natural mortality (M)

3) Productivity
   a) Trends in growth
      i) Length- and weight-at-age

4) Fishery
   a) Effort
      i) Temporal trends
      ii) Spatial trends

   b) Catch
      i) Temporal trends
      ii) Spatial trends
      iii) Structure of catch

5) Environmental
   a) Quantitative effects of environmental parameters (i.e., temperature and dissolved oxygen) on population parameters
      a. Relationships between bottom temperature, or North Atlantic Oscillation (NAO), and productivity

Criteria must be developed to determine the adequacy of data used to develop indicators. Suggested criteria:

- Data should be representative of the whole stock area
- Time series should be collected consistently over time

- Data used in estimating population and fishery parameters should be consistent with data used for the assessment model
    - the same surveys should be used in all types of analyses
    - recruit and exploited-size indices should be developed in the same way
- Consistent types of data and empirical indicators should be available for all stock units (e.g., Gulf of Maine vs. Southern New England)

Critical values must also be developed to determine safe and unsafe levels of indicators (see the traffic light approach in Section V).

Considerations for Management

Indicators of stock health should be selected based on their ability to clearly exhibit changes associated with management actions. Those more likely to be influenced by other anthropogenic or environmental factors should be avoided. Managers should understand not all variables are under their control. They should account for this by using harvest strategies and control rules that apply a precautionary approach to scientific uncertainty.

Indicators of stock status generally improve with the length of the time series. Indicators should be reported annually, but year-to-year variations usually are not indicative of management success or failure (McBride and Houde 2006). We are more interested in responding to patterns and trends in indicators than annual variation (see smoothing approaches in Section IV-B).

V. Control Rules

A. Targets and Thresholds

Stock size (reproductive capacity) and fishing mortality thresholds establish the criteria for determining when a stock is fished at unacceptably high fishing mortality rate or when the stock size (reproductive capacity) is unacceptably low. This may be due to fishing mortality, natural mortality, or recruitment failure. If overfishing caused the decline, then the stock may be labeled 'overfished'. If the decline occurred due to other factors, then the stock could be described as 'depleted' or 'at low abundance'. When overfishing occurs, fishing mortality needs to be reduced. An overfished stock triggers the need for a rebuilding plan, although rebuilding may be undertaken for stocks at low levels due to causes other than fishing. If a stock size threshold is crossed, fishing mortality should be reduced to allow for rebuilding.

In general, fishing mortality and stock size thresholds are set to prevent a stock from reaching conditions where productivity (reproduction) or sustainability may become compromised and serious harm to the stock may occur. Fishing mortality targets are generally set far enough from fishing mortality thresholds to account for uncertainty in fishing mortality estimates. Stock size thresholds are established to assure the sustainability of the stock and fishery, by providing an adequate buffer to prevent the population from reaching a stock size where productivity may be compromised. Control rules and overfishing definitions should accommodate the information available and uncertainty in the data. Incorporating uncertainty in making recommendations about specific values for targets and limit reference points is particularly important. The more uncertainty in the estimate of fishing mortality rate, the greater the buffer zone should be between F_{target} and $F_{\text{threshold}}$ (NAFO SCS Doc. 4/12, 2004). The biomass threshold should be set to avoid a stock size where long-term productivity is compromised and management actions should be implemented to avoid crossing the threshold.

If stock status is between the thresholds and targets, additional advice on stock status can be generated using either the 'traffic light' approach of Caddy, or the 'Consumer Report' approach of Rago (personal communication, Figure 3). To the extent practicable, these indicators should have a quantitative basis. Multiple indicators should be considered. The use of only a single indicator can be misleading if there is an anomaly in a data set. For example, one survey tow with an extremely large catch can cause significant change in a survey index's trend. Advice for management action should be based on the evaluation of several indicators in combination.

B. Concept of control rules and overfishing/overfished definitions

The following recommendations are from Rosenberg (1994).

- An overfishing definition can be comprehensively expressed as a threshold harvest control law relating target and threshold fishing mortality rates to stock biomass or abundance.
- The overfishing definition should be explicitly linked to management actions and rebuilding programs agreed upon ahead of time while the stock is healthy. This avoids delays in implementing remedial action should the stock become overfished.
- The process for updating overfishing definitions when new information becomes available should be explicit. The updating can take two forms: 1) an update of the parameter value within the same model or 2) revision of the reference points using a new model framework.

Rosenberg et al. (1994) suggest defining overfishing using a combination of maximum fishing mortality rate, a precautionary stock size target below which the allowable fishing mortality rate is reduced, and an absolute minimum stock size threshold.

A form of this control rule is used in the New England groundfish management plan. The fishing mortality threshold is set at F_{MSY} or proxy. When overfishing is determined, action is required to end overfishing. A biomass target (B_{target}) is set to ensure stock biomass remains at a sustainable biomass. This is often set at B_{MSY} or proxy. A minimum biomass threshold is set at some fraction of B_{target} such as ½ B_{target} or $(1\text{-}M)B_{MSY}$ (Restrepo et al. 1998) to prevent the population from falling to a biomass where long-term productivity is diminished. A requirement for a formal rebuilding program is triggered when biomass drops below the biomass threshold. Once a stock is declared overfished, the rebuilding program must specify a time frame to rebuild the stock to B_{target}.

If the stock is at the lower end of the biomass range, then fishing mortality should be low. Managers may decide to end directed harvest and only allow a bycatch fishery when biomass reaches a specified threshold. Figure 2 presents an example of how control rules can be used for status determination and management action.

The use of these four criteria provides a broad view of stock status. However, other factors such as age distribution, geographic distribution, fish condition, recruitment, age distribution of the spawning population, mean weights-at-age, etc., also provide information on stock status. In some cases, these other factors provide important information on stock status augmenting the broader view. For example, spawning biomass may temporarily rise above $B_{threshold}$ based on the influence of a strong single cohort that may pass quickly through the fishery.

C. Caddy's Traffic light approach

A more qualitative approach is the traffic light system of Caddy (1998, 2004). Caddy's approach (Caddy 2004) provides a way to use multiple indicators and their critical reference points for managing populations. Multiple indicators measuring quantifiable life history characteristics would be scored as 'green' if they fall into the safe zone, 'red' if they fall over the reference point representing a dangerous condition for the stock, or 'yellow' if the indicators represent an uncertain situation as transitional between the green and red zones. The problem occurs in not just how to define the reference points (the color boundaries) but, when using multiple indicators, how to weigh them before combining within a 'characteristic' (i.e., abundance or fishing mortality). Combining multiple indicators within a formal management rule is complex. The key objective is that it should be integrated into a precautionary harvest law or management information system. An example of the traffic light approach applied to the American lobster stock (ASMFC 2005) is presented in Table 2.

Halliday et al. (2001, in Caddy 2004) proposed that decision rules be based on the integrated score of indicator values measuring at least three characteristics (abundance, production, and fishing mortality). A gradation of responses is likely to occur, since individual indicators of characteristics may not be triggered simultaneously. This may provide some redundancy and smoothing if the proportion of indicators triggered determines the degree of management response.

Prager (1994) and Prager et al. (2003) (also in Caddy, 2004) argued the advantages of normalizing time series, or expressing indicators in dimensionless form such as the ratio of current value to the 'optimal' value. This puts all the indicators on the same scale, though it might be advisable to use standard deviation units (Z transformed) to also eliminate differences in variability.

D. Rago's Non-Parametric multi-attribute assessment methodology

Rago (personal communication) has developed a nonparametric multi-attribute assessment methodology using psychometric measures ('consumer reports approach'). It was developed as a graphical method to summarize the results of stock assessments. This approach determines percentiles (quintiles) of population attributes. Colors are assigned to each quintile (i.e., red, pink, white, gray, and black from worst to best). These color-coded quintiles are presented in tabular form. Areas with many red quintiles means there is concern about the status of the stock, while areas with many black quintiles means the stock seems healthy. An example of this for mean weights-at-age for Georges Bank yellowtail flounder and Georges Bank haddock is shown in Figure 3 (from Figure 3.6 of GARM II report).

E. Precision and Bias

To effectively present risk to managers, conveying the precision of an estimate is important. Presenting risk (i.e., uncertainty) estimates, like confidence intervals, for current levels of fishing mortality or stock size allows managers to understand the probability of exceeding $F_{\text{threshold}}$ or falling below the $B_{\text{threshold}}$ and provides a more realistic view of stock status.

F. Uncertainty around reference points

Reference points calculated from models have uncertainty associated with them that is derived from uncertainty in data inputs, among other sources, and they are calculated with error. This uncertainty can be expressed as a distribution around the point estimate, similar to the distribution of an estimate of fishing mortality produced via a bootstrap. Monte Carlo methods can be employed to estimate the uncertainty around a reference point by drawing model inputs from a distribution. Helser et al. (2002) demonstrate this approach in developing a reference point of the replacement fishing mortality rate as an overfishing threshold for a blue crab fishery.

G. Retrospective bias

The uncertainty of fishing mortality is highest in the last or current year of several types of models (e.g., ADAPT, MARK). As younger cohorts are repeatedly measured, more accurate ‘retrospective’ estimates of fishing mortality and stock size can be obtained. The differences between fishing mortality in the current year and later retrospective values are a measure of the reliability of the current year estimates, and can be displayed as a histogram of ‘uncertainty’ in current year *F* (Caddy and McGarvey, 1996).

In some assessments, a pattern of consistently overestimating or underestimating population parameters such as stock size or fishing mortality occurs as more years are added to the model. This retrospective bias can seriously erode the quality and robustness of management advice. The presence and magnitude of a retrospective pattern should be reported. Large retrospective bias indicates a serious problem with the assessment model. Large bias occurs when the converged estimate lies outside some pre-specified percentiles of a bootstrap estimate of the terminal year of a previous assessment. Retrospective biases can be detected using spaghetti plots (Figure 4), plotting historical amounts as deviations from the most recent assessment, or the calculation of Mohn's Rho statistic (summed proportion difference relative to the full-dataset-based estimates) (Legault 2008).

An example of a retrospective bias for Gulf of Maine winter flounder is shown in Figure 5 (from Groundfish Assessment Review Meeting II). This stock was selected for illustrative purposes because the retrospective pattern is severe. This particular pattern is troubling as fishing mortality is consistently underestimated, spawning stock biomass is overestimated, and the bias is large. The bootstrap distributions of retrospective terminal year estimates of spawning stock biomass for years 2000-2003 are shown in Figure 5. The panel for SSB in 2002 (upper left panel) provides a clear example of the retrospective problem. In terminal year 2002, the box plot indicates a very high probability that the true SSB lies between 3000 to 5000 mt. With the addition of a single year of data (terminal year 2003), the bootstrap distribution of 2002 SSB estimates shifts: nearly 75% of the bootstrap estimates are under 3000 metric tons. The addition of another year of data (terminal year 2004) causes a further shift in the distribution, and the bootstrap distribution of 2002 SSB does not overlap the distributions generated in terminal years 2003 and 2002. Clearly a conflict exists as the estimates from the converged portion of the VPA lie outside the bootstrap terminal year estimates. Managers should consider using a precautionary approach to management when retrospective bias is present in an assessment.

VI. Summary

This guide provides fisheries scientists with a variety of options and methods to use in developing reference points and determining reliable indicators of stock status. Guidance is also provided to managers who interpret stock status with respect to reference points, and subsequently implement management strategies according to control rules. It is important to develop reference points that are appropriate for the stock, internally consistent with assessment methodology, and provide clear meaning to managers.

To assist assessment scientists and managers, several terms related to reference points were defined, and existing regulations and requirements for reference points were examined. The calculation of reference points (both model-based and index-based) and their relationships to control rules were also reviewed. We distinguished between the reference points derived from statistical models and the index-based reference points based directly on field observations. Both model-based and index-based methods have advantages and disadvantages for management, as well as varying amounts of uncertainty. Neither type of reference point is more valid than the other for management purposes. However, model-based reference points tend to be favored because they are typically derived from multiple sources of information integrated into a model framework, and uncertainty is often quantifiable.

Three standard levels of reference points are typically used in fisheries management - limits, thresholds, and targets. Levels of these reference points need to be distinctly different to provide clear guidance for management decisions. Mortality rate based reference points define when 'overfishing' is occurring on a stock. Abundance/biomass based reference points determine when a stock is 'overfished' or stock size is unacceptably low. Control rules can be designed to prevent biomass from falling below the threshold, thus reducing the risk of low stock productivity over long periods.

In addition to fishing mortality and population biomass parameter estimates, several stock indicators can be examined to interpret stock status relative to reference points. These stock indicators are used to corroborate or refute model or empirical results and provide additional information about the overall health of each stock. Multiple indicators bring together a variety of monitoring observations, traditional stock assessment model results, and fishery indices to form management advice. The use of multiple indicators will tend to reflect the true state of the stock. Indicators of stock health should be selected based on their ability to clearly exhibit changes associated with management actions.

Regardless of the types of reference points and methods utilized, managers should be aware of the variability, accuracy, and precision of these measures. Managers need to take a precautionary approach when setting rates and TAC, and when taking action as defined in control rules, particularly when rigorous analyses are not possible due to data constraints. This approach will improve the likelihood of consistently reaching targets and not exceeding thresholds. The amount of precaution exercised should be based on the uncertainty associated with the types of reference points or methods chosen.

VII. References

Atlantic States Marine Fisheries Commission (ASMFC). 2002. Guidance Document for ASMFC Technical Support Groups. September, 2002. Washington, DC. 37 pp.

ASMFC. 2006. MSA Reauthorization Provisions of Interest to Atlantic States. December, 2006. Washington, DC. 2 pp.

Caddy J. F. 2004. Current usage of fisheries indicators and reference points, and their potential application to management of fisheries for marine invertebrates. Canadian Journal of Fisheries and Aquatic Science 61:1307-1324.

Caddy J. F., and R. Mahon. 1995. Reference points for fisheries management. FAO Fisheries Technical paper # 347.

Caddy J. F., and R. McGarvey. 1996. Targets or limits for management of fisheries? North American Journal of Fisheries Management 16(3):479-487.

Cleveland, W. S. 1993. Visualizing Data. Hobart Press, Summit, New Jersey, U.S.A.

Gerard, P.D., D.R. Smith, G. Weerakkody. 1998. Limits of retrospective power analysis. Journal of Wildlife Management 62(2):801-807.

Gerrodette, T. 1991. Models for power of detecting trends – a reply to Link and Hatfield. Ecology 72(5):1889-1892.

Hastie, T.J. and Tibshirani, R. J. 1990. Generalized Additive Models. First edition, Chapman and Hall/CRC, Boca Raton, FL 335 pp.

Helser, T. E., A. Sharov and D. M. Kahn. 2002. A stochastic decision-based approach to assessing the status of the Delaware Bay blue crab stock. pp. 63-82 *in* Berkson, J. M., L.

L. Kline and D. J. Orth, Editors. Incorporating Uncertainty into Fishery Models. American Fisheries Society Symposium 27, Bethesda MD.

Excerpts from http://ipl.unm.edu/cwl/fedbook/magfish.html

Koeller, P., Savard, L., Parsons, D.G., and Fu C. 2000. A precautionary approach to assessment and management of shrimp stocks in the Northwest Atlantic. Journal of Norwest Atlantic Fisheries Science 27:235-246.

Krebs, C. J. 1999. Ecological Methodology Second Edition, Benjamin/Cummings, Menlo Park CA 620 pp.

Legault, C.M., Chair. 2009. Report of the Retrospective Working Group, January 14-16, 2008, Woods Hole, Massachusetts. U.S. Department of Commerce, Northeast Fisheries Science Center Ref Doc. 09-01; 30 pp. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/

Lo, N.C.H., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences 49: 1515-1526.

Mace, P.M. 1994. Relationships between common biological reference points used as thresholds and targets of fisheries management strategies. Canadian Journal of Fisheries and Aquatic Sciences 51:110-122.

Magnuson-Stevens Fishery Conservation and Management Act 16 U.S.C. 1801-1882, April 13, 1976, as amended 1978-1980, 1982-1984, 1986-1990, 1992-1994 and 1996.

Manly, B.F.J. 2001. Statistics for environmental science and management. Chapman and Hall. Pp. 123-125.

McBride, M. and Houde, E.D. 2006. Fisheries Ecosystem Planning for Chesapeake Bay. NOAA Chesapeake Bay Fisheries Ecosystem Advisory Panel. American Fisheries Society, 410 pp.

McCullagh, P., and J. A. Nelder. 1989. Generalized Linear Models. Second edition. Chapman and Hall/CRC, Boca Raton, Florida. 532 pp.

National Marine Fisheries Service (NMFS). 2007. http://www.nmfs.noaa.gov/msa2007/. March 6, 2007.

Pinheiro, J. C. and Bates, D. M. 2002. Mixed Effects Models in S and S-Plus 3rd edition, Springer, 528 pp.

Prager M. H., C. E. Porch, K. W. Shertzer, J. F. Caddy. 2003. Targets and limits for management of fisheries: a simple probability-based approach. North American Journal of Fisheries Management 23:349-361.

Restrepo, V. R., G. G. Thompson, P. M. Mace, W. L. Gabriel, L. L. Low, A. D. MacCall, R. D. Methot, J. E. Powers, B. L. Taylor, P. R. Wade, J. F. Witzig. 1998. Technical Guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. U.S. Dep. Commer., NOAA Tech Memo., NOAA-TM-NMFSF/SPO-31.

Rosenberg A., P. Mace, G. Thompson, G. Drach, W. Clark, J. Collie, W. Gabriel, A. MacCall, R. Methot, J. Powers, V. Restrepo, T. Wainwright, L. Botsford, J. Hoenig, K. Stokes. 1994. Scientific review of definitions of overfishing in U.S. fishery management plans. NOAA Technical Memorandum. NMFS-F/SPO-17.

Thomas, L. 1997. Retrospective power analysis. Conservation Biology 11:276-280.

Table 1. Reference points for species managed by the Atlantic States Marine Fisheries Commission.

Species	**Reference Point**	**Management Authority**	**Assessment Type/ Review Date**
American Eel	None	ASMFC	Trends – 2005
American Lobster (GOM, GB, SNE)	Abundance or *F* /Trends	ASMFC	CSA - 2005
American Shad	Abundance or *F*	ASMFC	Trends/*Z* estimates - 2007
Atlantic Croaker (Mid-At/S.Atl)	Biomass and *F*	ASMFC-SAFMC	Age-structured Production Model - 2004
Atlantic Herring	Biomass and *F*	ASMFC-NEFMC	SCA - 2006
Atlantic Menhaden	Biomass, *F*, and Fecundity	ASMFC	SCA - 2003
Atlantic Sturgeon	None	ASMFC	Index - 1998
Black Sea Bass	$B_{40\%}$, $F_{40\%}$	ASMFC-MAFMC	SCALE - 2008
Bluefish	Biomass and *F*	ASMFC-MAFMC	ASAP - 2005
Horseshoe Crab	None	ASMFC	Trend Analysis - 1999
Large Coastal Sharks	Relative Abundance and Relative *F*	ASMFC-NMFS	Production Model 2006
Northern Shrimp	Biomass and *F*	ASMFC	CSA/Production - 2007
Red Drum	%SPR (% Spawning Potential Ratio)	ASMFC-SAFMC	VPA - 2000
River Herring	None	ASMFC	None
Scup	$B_{40\%}$, $F_{40\%}$	ASMFC-MAFMC	ASAP - 2008
Small Coastal Sharks	Relative Abundance and Relative *F*	ASMFC-NMFS	Production Model 2007
Spanish Mackerel	Biomass and *F*	ASMFC-NMFS	VPA - 2003
Spiny Dogfish	Relative Biomass or *F* /Trends	ASMFC-NEFMC-MAFMC	Index based - 2006
Spot	None	ASMFC-SAFMC	None
Spotted Seatrout	%SPR (% Spawning Potential Ratio)	ASMFC-SAFMC	None
Striped Bass	Biomass and *F*	ASMFC	SCA - 2007
Summer Flounder	$B_{40\%}$, $F_{40\%}$	ASMFC-MAFMC	ASAP - 2007
Tautog	Biomass or *F*	ASMFC	VPA - 2005
Weakfish	Biomass and *F*	ASMFC	VPA/Production - 2006
Winter Flounder (GOM / SNE)	Biomass and *F*	ASMFC-NEFMC	VPA - 2008

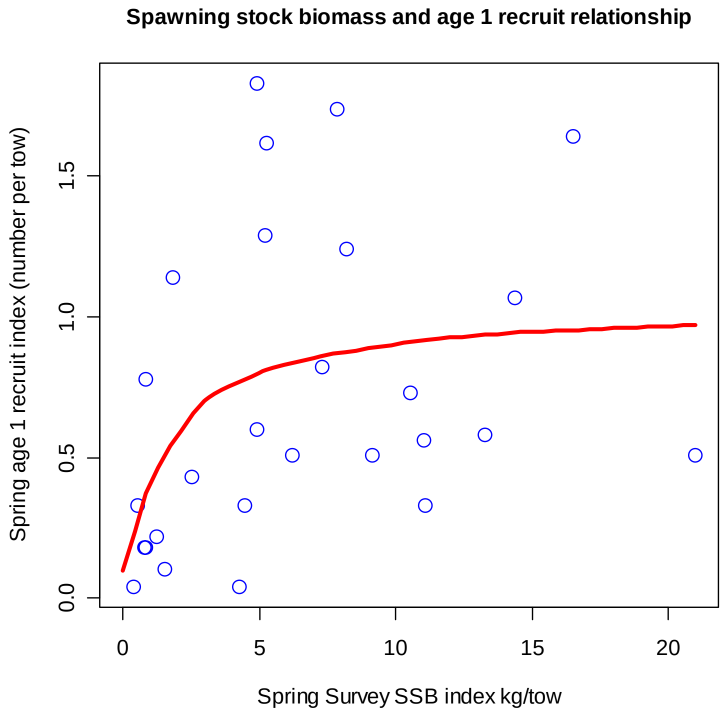
Table 2. Types of reference points, their advantages and disadvantages, and the modeling approaches and data inputs used to derive them.

Model	**Data Inputs**	**Reference Points**	**Advantages**	**Disadvantages**
Production	Historical catch and standard effort data	F_{MSY}, B_{MSY}	Minimum data requirements	- Need contrast in catch history - assumes equilibrium growth rate and carrying capacity
Yield/recruit	Age or length structure, average individual growth, M, vulnerability to fishery	F_{max}, $F_{0.1}$	Useful when several fleet components exploit different age groups and when gear regulations affect age/size of first capture	Does not include information on reproductive potential; assumes constant recruitment
SSB(egg)/recruit	Extension of yield/recruit which incorporates a time series of recruitments, S/R curve or age/size at maturity	$F_{X\%}$, SSB_{MSY} or proxy	Incorporates the magnitude of spawning potential	May utilize S/R curve which may not be well defined

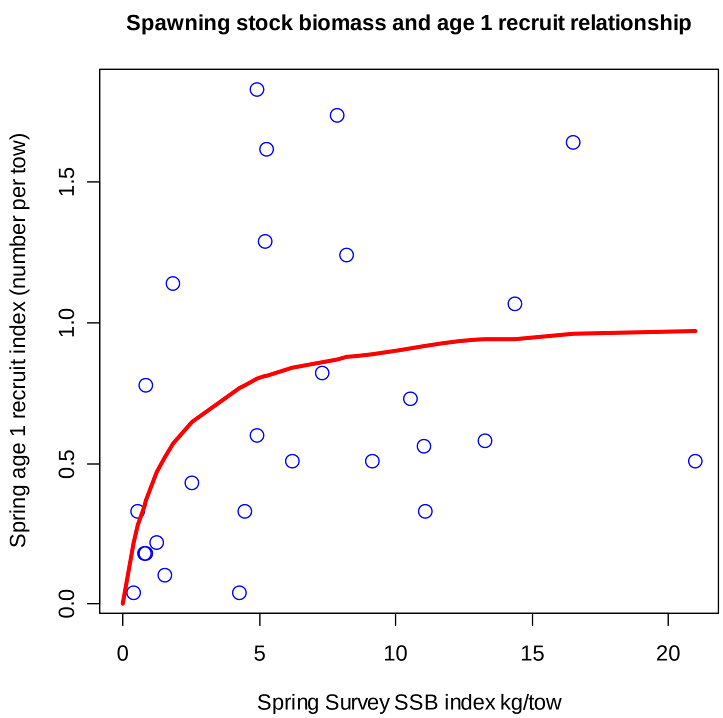
Table 3. An example of Caddy's (2004) traffic light approach using Gulf of Maine lobster.

Stock Indicators	*Mortality Indicators*				*Abundance Indicators*				*Fishery Performance Indicators*			
	1. Exploitation Rate	2. Z (model)	3. Mean Length	4. Proportion of the Exploitable Stock Comprised of Recruits	5. Spawning Stock Abundance Index	6. Recruit Abundance (sexes combined Model)	7. Full Recruit Abundance	8. Settlement Index	9. Effort (traps)	10. Landings (mt)	11. Mean Length (Landings)	12. Gross CPUE
1982	0.57	1.08	98.8	0.71	2.43	32.95	13.16		2390415	14,669	89.35	13.53
1983	0.57	1.08	92.9	0.72	3.83	32.90	13.09		2599642	15,069	89.35	12.78
1984	0.46	0.84	96.7	0.71	5.42	37.85	15.58		2450165	13,797	89.12	12.41
1985	0.55	1.05	96.3	0.51	9.35	23.99	23.17		2079758	14,558	89.02	15.43
1986	0.51	0.93	96.8	0.67	6.97	34.01	16.63		1926713	13,816	89.06	15.81
1987	0.48	0.86	92.1	0.61	5.20	30.68	19.99		2265169	13,952	88.83	13.58
1988	0.61	1.19	93.1	0.50	5.02	21.02	21.43		2409689	14,696	88.45	13.45
1989	0.55	1.03	96.5	0.75	4.21	39.48	13.01	1.64	2396941	16,708	88.68	15.37
1990	0.53	1.00	92.6	0.69	7.28	42.74	18.83	0.77	2545777	19,244	88.92	16.67
1991	0.57	1.09	92.4	0.63	7.24	39.39	22.85	1.54	2444711	20,215	88.97	18.23
1992	0.57	1.10	97.4	0.62	6.05	33.79	20.91	1.30	2434537	17,738	88.89	16.06
1993	0.53	0.97	97.0	0.69	7.35	40.96	18.29	0.45	2222578	18,802	89.19	18.65
1994	0.54	1.01	92.8	0.67	6.26	46.44	22.49	1.61	2821359	23,869	89.45	18.65
1995	0.43	0.77	97.4	0.74	7.39	69.76	25.01	0.66	3025934	23,001	89.51	16.76
1996	0.46	0.82	95.8	0.45	13.21	35.69	43.76	0.47	2908361	22,155	89.17	16.79
1997	0.45	0.80	103.1	0.64	12.06	61.92	34.92	0.46	3036822	26,726	89.28	19.40
1998	0.48	0.87	95.7	0.54	15.67	51.36	43.47	0.14	3258231	25,836	89.49	17.48
1999	0.44	0.80	100.2	0.62	13.33	66.01	39.87	0.65	3461777	30,038	89.78	19.13
2000	0.49	0.89	95.1	0.58	20.45	67.48	47.93	0.13	3202571	31,845	90.05	21.92
2001	0.41	0.74	101.4	0.57	18.06	63.71	47.95	2.08	3388671	26,517	89.74	17.25
2002	0.42	0.75	102.9	0.57	24.32	72.11	53.56	1.38	3515509	33,806	89.78	21.20
2003	0.33	0.58	107.3	0.58	30.61	82.73	59.90	1.75	3623066	29,198	89.9	17.77
2001-03 Avg.	0.39	0.69	103.8	0.57	24.33	72.85	53.80	1.74	3509082	29,840	90	19
Median	*0.49*	*0.57*	*93.6*	*0.62*	*7.35*	*40.96*	*22.85*	*1.03*	*2599641.9*	*20,215*	*89*	*17*
25th	*0.45*	*0.81*	*96.6*	*0.58*	*5.57*	*33.84*	*18.42*	*0.47*	*2400127.9*	*14,789*	*89*	*15*
75th	*0.55*	*1.04*	*98.5*	*0.69*	*13.30*	*63.26*	*42.57*	*1.57*	*3161133.6*	*26,347*	*90*	*19*

Figure 1. A comparison of loess smoothing and a Beverton-Holt function to examine relationships between spawning stock biomass and age 1 recruits.



loess smoothed



Beverton Holt

Figure 2. Elements of a control rule. **Bold** indicates thresholds and targets. Underlined is stock status. *Italics* describe management actions agreed upon by managers that are triggered by changes in stock status.

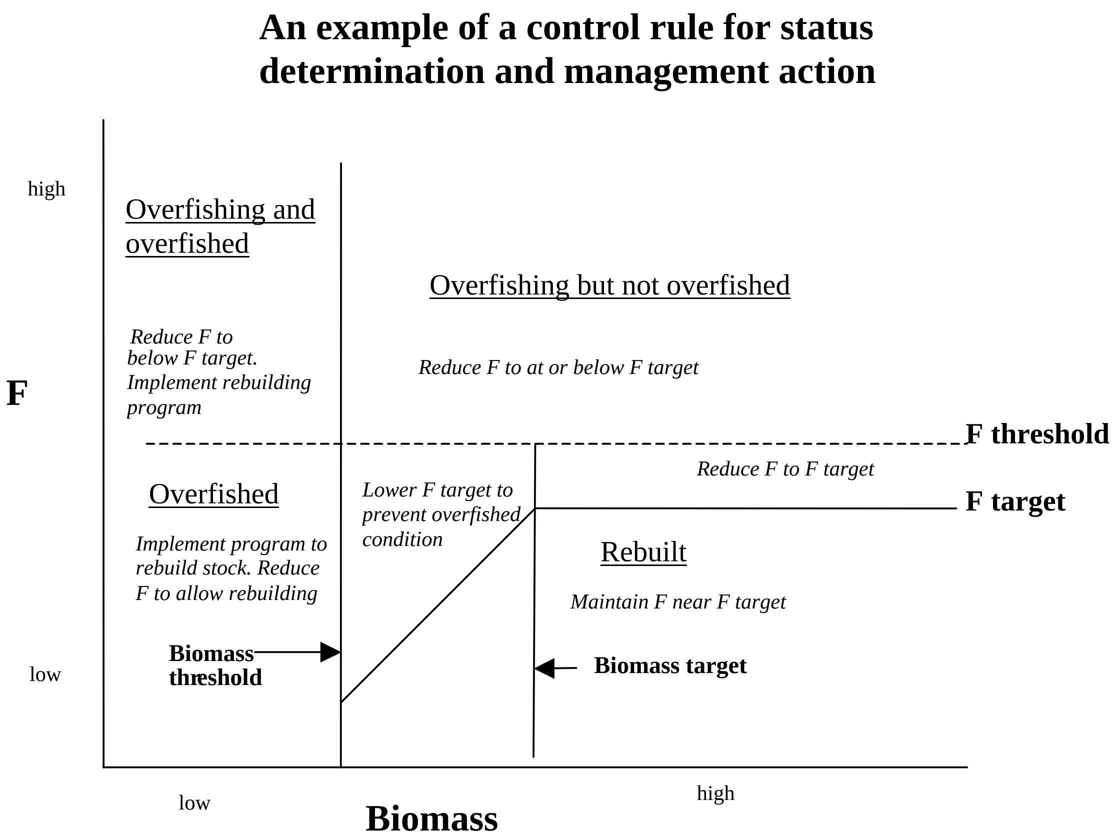


Figure 3. Example of Rago's consumer reports method for relative changes in average stock weights-at-age for Georges Bank cod (GBCOD), haddock (GBHAD), and yellowtail flounder (GBYT) [Figure from GARM II (NEFSC 2004)].

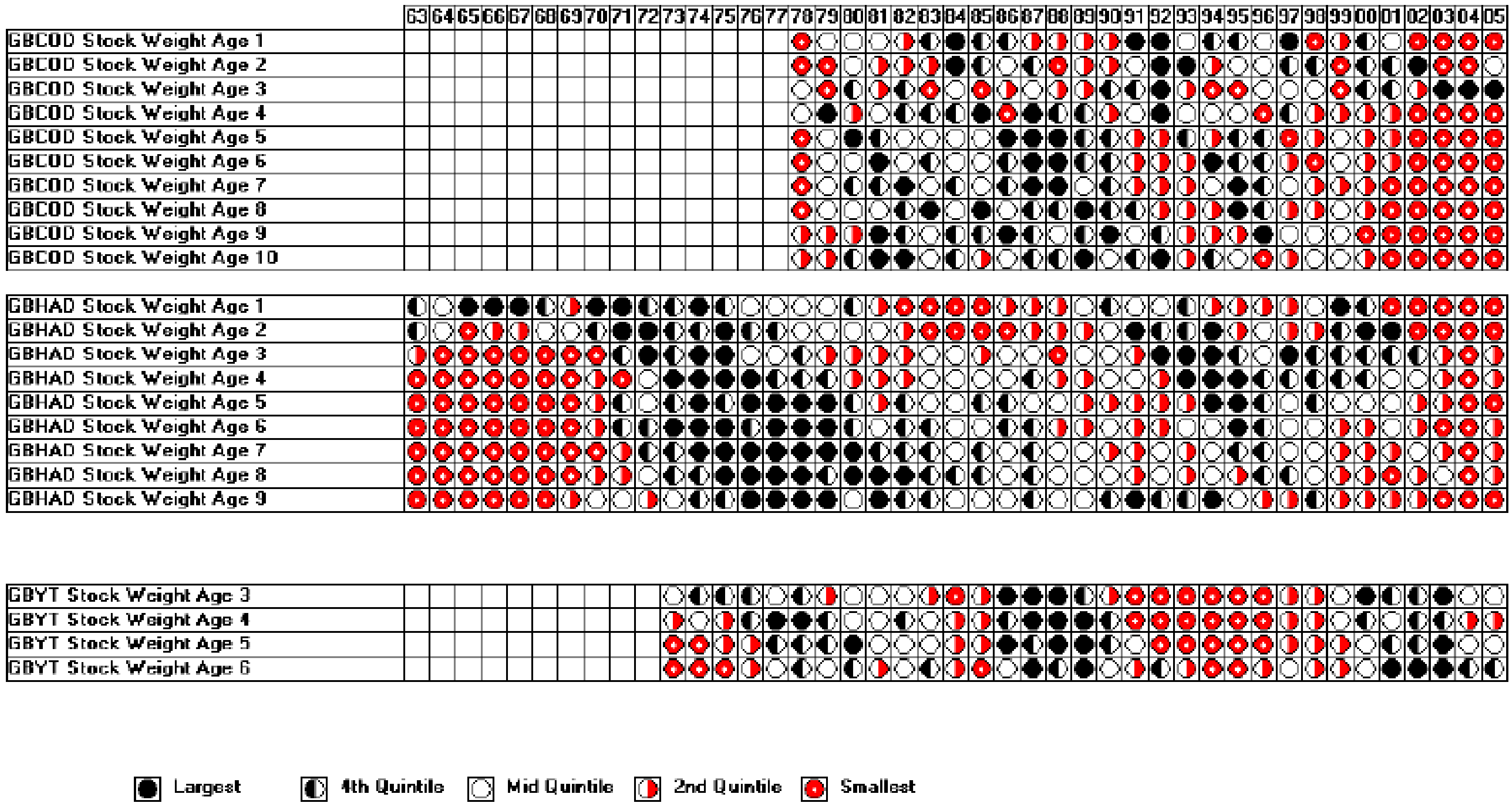
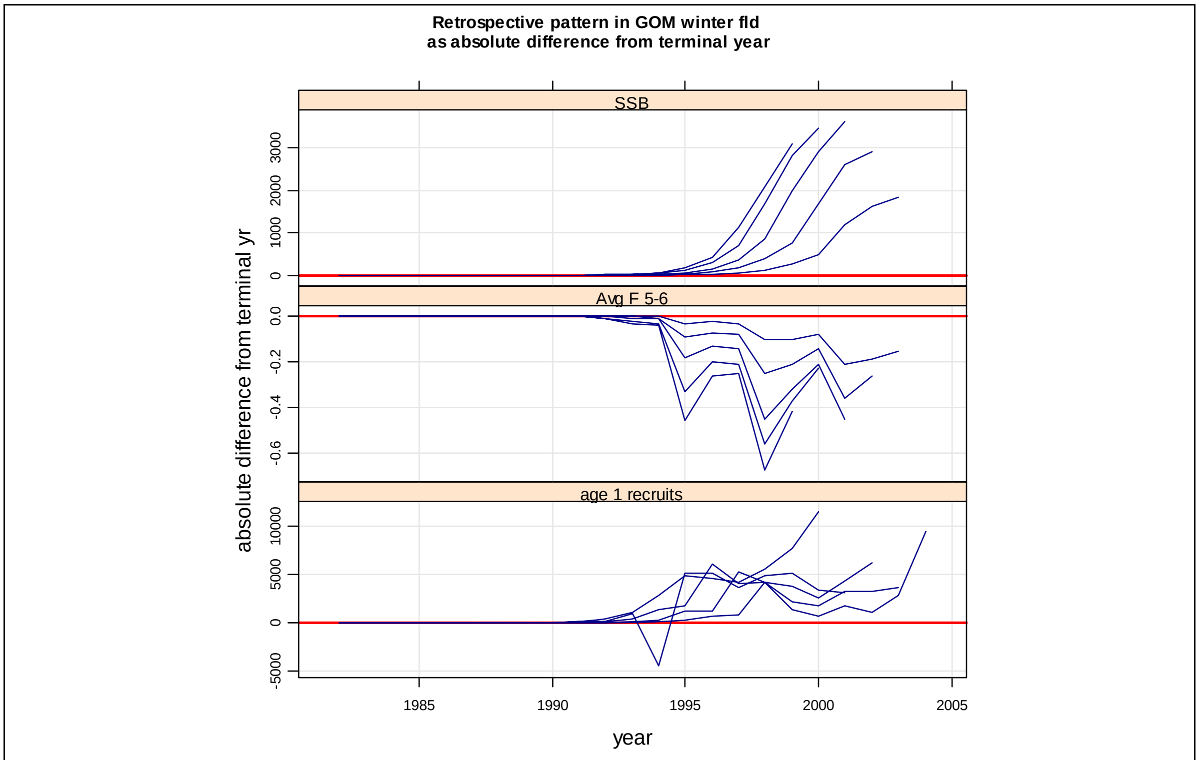


Figure 4. Two displays of retrospective bias for Gulf of Maine winter flounder (GARM II, NEFSC 2004). Top panel: retrospective is measured as difference between retrospective years and terminal year 2004 values (zero line is terminal year). Bottom panel: retrospective showing values for each retrospective terminal year. SSB units are metric tons, age 1 recruits are in thousands of fish.



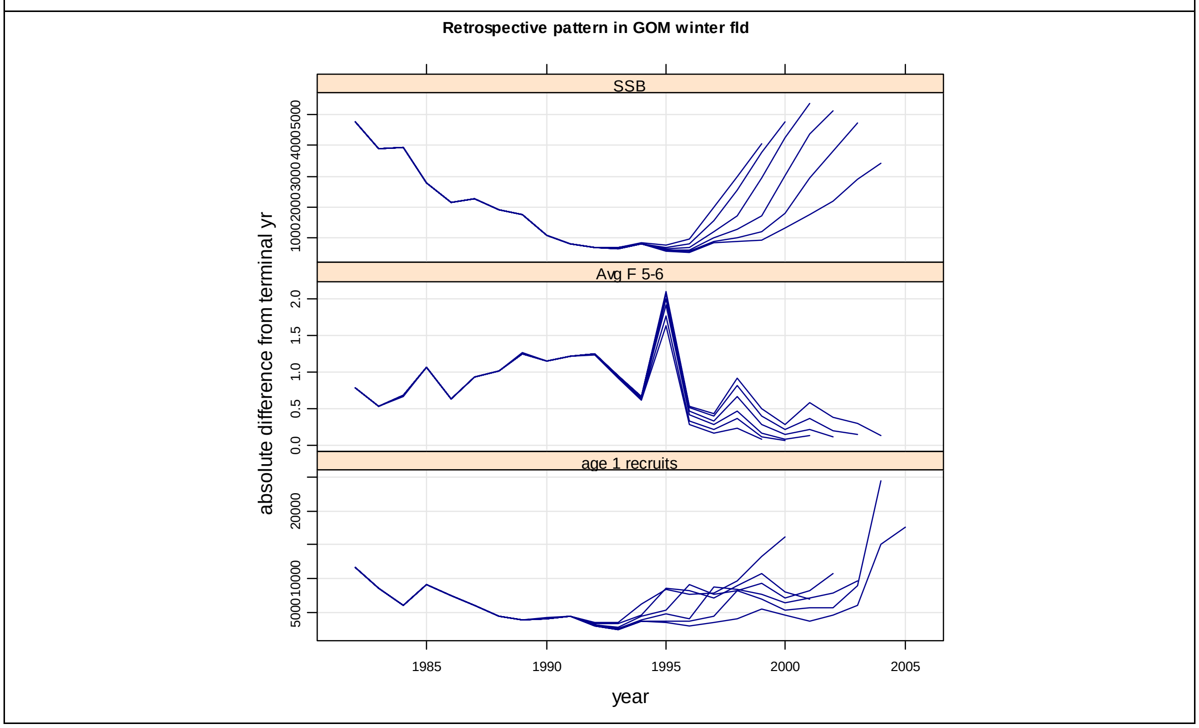


Figure 5. Distribution of retrospective bootstrap estimates of terminal year spawning stock biomass for Gulf of Maine winter flounder. Each panel shows the bootstrap distribution of terminal year estimates for SSB for a particular year. Y-axis is the terminal year in the retrospective analysis.