Use of Quality Control Methods to Monitor the Status of Fish Stocks

James Scandol
NSW Department of Primary Industries, Cronulla Fisheries Centre, Cronulla, New South Wales, Australia

Abstract

Many fisheries that are data-limited are also of low economic value. Therefore, not only are the fisheries data-limited, but there are limited human resources available for undertaking stock assessment. Qualitative methods such as “eyeballing” the data are then often used to assess such systems. Quantitative methods need to be developed that are objective, but less demanding than dynamic stock assessment models. In particular, simple methods that can signal trends in empirical stock-status indicators need to be explored. One such approach is the use of quality control methods such as Shewhart, moving-average, and CUSUM (cumulative sum) control charts. Originally designed for industrial quality control, these methods can be parameterized to detect transient or persistent causes with specific false-positive and false-negative error rates. These signals can be interpreted within a managerial context as trigger reference points.

Results of a simulated study of yellowfin bream (Acanthopagrus australis) stocks from New South Wales (Australia) are presented. Empirical stock-status indicators including catch, catch per unit effort, mean age, mean length, recruitment fraction, total mortality, and fishery-independent surveys were processed using quality control methods. Performance of these indicators and algorithms were measured with receiver-operator-characteristic curves, which captured both false-positive and false-negative error rates. Biomass surveys performed best, followed by mean age and length, and recruitment fractions. Commercial catch rates and catch had the worst performance but were still acceptable. Age-based total mortality performed poorly unless very large numbers of samples were taken. Potential applications of these methods include a rapid diagnostic tool in data-limited situations, development of empirical reference points,
and empirical rule-based management systems. These methods are easily applied even when there is a short time-series of low-contrast data but a range of caveats must always be considered.

**Introduction**

There is a wide range of quantitative methods to assess the status of a fish stock (Quinn and Deriso 1999) and predict the effect of alternative management choices (Hilborn and Walters 1992). Researchers working in data-limited fisheries are, however, frequently frustrated when even simple methods such as biomass dynamic or delay-difference models cannot be satisfactorily calibrated to observations. These models generally fail because of insufficient contrast, or information content, within the available data sets.

Another issue that can dominate data-limited stock assessment is lack of analytical expertise. Fisheries of low economic value are frequently both data-limited and expertise-limited because of small budgets for research, monitoring, and assessment. This constrains what can practically be achieved for many stocks. For example, issues associated with low-contrast data can be overcome using Bayesian methods that “borrow” information from other fisheries (Punt and Hilborn 1997). Expertise to apply such methods is expensive and unlikely to be used for low-value fisheries.

When data are limited and available expertise is constrained, fisheries are often assessed by “eyeballing” graphs of empirical indicators (such as commercial catch per unit effort [CPUE]). Anybody who has worked in data-limited fisheries must have seen this happen, and given the efficiency of the human eye for identifying patterns, this is neither surprising nor undesirable. Such methods are, however, somewhat subjective. Given the importance of developing objective and transparent rules for managing fisheries this approach could be criticized.

This paper attempts to identify a middle ground between “eyeballing” data, and calibrating and applying sophisticated dynamic models. Fisheries management systems have developed that utilize results from stock assessment via indicators (such as fishing mortality $F$) and biological reference points (such as $F_{0.1}$). Relationships between indicators and biological reference points are then used to signal particular management actions. Stock assessment therefore involves coupling data analysis and modeling with a signaling system for management. Signaling systems are a core feature of statistical quality control (QC), where indicators are monitored to flag uncharacteristic values or an undesirable trend in a sequence of values. A common application is to ensure that a manufacturing facility is producing goods with consistent characteristics. QC algorithms are simple and numerically stable but cannot, of course, provide the same insight into a fishery that a dynamic model might achieve. This
paper examines the application of QC methods as a signaling system for fisheries management. Indicators derived directly from data would be analyzed (very simply) and signals from these indicators reported. Quality control methods may help fill the very large gap between “eyeballing” data and developing complex stock assessment systems.

This paper makes an important differentiation between empirical and estimated indicators of stock-status. An “empirical indicator” is one that is calculated directly from a specific set of raw data and, in the process of calculation, may use one or two parameters that can be easily defined. In contrast, an “estimated indicator” represents a fishery variable of interest that is derived from a range of data sets and is dependent upon additional parameters that may or may not be easily defined. For example, raw commercial CPUE and mean age are empirical indicators, while biomass and fishing mortality are estimated indicators. Empirical indicators would usually be calculated with a closed algorithm such as a database query, while estimated indicators frequently require iterative nonlinear algorithms. An exception to these generalities is the use of iterative general linear modeling algorithms to estimate an index-of-abundance.

Much fishery management science is oriented around the calculation and interpretation of estimated indicators. In data-limited situations where the calculation of estimated indicators is unreliable, the initial focus should be upon empirical indicators (also see Die and Caddy 1997, Caddy 1999). Several recent studies have measured the performance of empirical indicators with methods not dissimilar to those presented here. Punt et al. (2001) evaluated mean length, mean weight, and CPUE data collected in the east Australian broadbill swordfish fishery. Trenkel and Rochet (2003) completed a sophisticated power analysis on a range of stock-status indicators, including empirical indicators such as mean age and total mortality. Rochet and Trenkel (2003) extended these ideas to indicators of community structure. Scandol (2003) estimated the performance of landed catch as an indicator with CUSUM (cumulative sum, Page 1961) quality control methods. Previous research focused upon indicators, while this study examined the interpretation and performance of these indicators using methods from statistical quality control. Given the recent emphasis on the development of indicators for sustainable fisheries management (Garcia and Staples 2000), evaluation of statistical methods to interpret these indicators should be a research priority.

Materials and methods

Overview

This study used an operating model to generate observations that were transformed into nine stock-status indicators. Times-series of these indicators were then analyzed for transient and persistent causes (defined below) using three quality control methods: Shewhart control charts,
CUSUM control charts, and moving-average Shewhart control charts. Signals from these quality control (QC) methods were then compared to the known state of the simulated biomass, and the sensitivity and specificity of these algorithms determined for each of the nine indicators. All of these QC methods required one or two parameters to specify an out-of-control observation.

**The operating model**

The operating model is detailed in the Appendix. The model was an age-structured population model for yellowfin bream (*Acanthopagrus australis*), a common species harvested off the coast of eastern Australia (Yearsley et al. 1999). The model was parameterized using data from Gray et al. (2000). A small fishery was modeled that included the likely variation in catchability, applied effort, and recorded effort among 50 commercial fishers. Simulated indicators for CPUE ($U_{t}^{\text{cpue}}$) and commercial catch ($C_t$) were generated by the model, and an unbiased indicator that represented the surveyable stock ($U_{t}^{\text{survey}}$) was also included.

Each year a random sample of fish from the simulated commercial fishery was taken and six other indicators were calculated: mean age ($\mu_{t}^{\text{age}}$); mean length ($\mu_{t}^{\text{len}}$); total mortality estimated with catch-curve analysis from the same of ages ($Z_{t}^{\text{age}}$); total mortality using the Beverton and Holt method from length statistics and growth characteristics ($Z_{t}^{\text{len}}$); and recruitment fractions (the fraction less than a threshold length or age) from age ($\theta_{t}^{\text{age}}$) or length ($\theta_{t}^{\text{len}}$) frequency distributions. Equations for these indicators are provided in the Appendix. More sophisticated analyses were obviously possible with these simulated data, but the aim was to capture the types of simple calculations that could be completed in a data-limited and expertise-limited environment.

The temporal sequencing of the operating model was important and there were four distinct phases. First, the model was run without any variation (but with the simulated fishery operating) until the exploitable biomass converged to an equilibrium value ($B_{eq}$). Second, the model was run for a further 30 years with variation to stabilize any transient processes. Third, the model entered a ten-year phase where historical data were collected and indicators calculated. Fourth, the model entered a ten-year future phase where additional data were collected and indicators calculated. Simulated observations from the future phase were used to measure the performance of the QC algorithms, and observations from the historical phase were used to provide a reference for the QC algorithms.

During the future phase, or the historical and future phases, various deleterious impacts were randomly imposed upon the fishery. These impacts affected recruitment (decreased); fishing effort (increased); catchability (increased); or natural mortality (increased). The maximal rate of an impact was a 20% change in a parameter value per year. The exact
magnitude and type of an impact affecting a simulation were randomized, but these impacts were manipulated by two parameters: the probability of an impact occurring during the historical and future phases \( (P_{\text{hist}}) \); and the probability of an impact occurring during the future phase only \( (P_{\text{futr}}) \). This representation was required to generate scenarios where the fishery was historically unstable as well as stable.

**The quality control algorithms**

The quality control literature uses the term “transient cause” to describe a short-term influence that affects one or two consecutive observations in a time-series, and the term “persistent cause” to describe a long-term influence that distorts a process for a longer period and is responsible for a trend. Identification of an uncharacteristic value within a time-series is a relatively simple statistical problem and is the basis of the Shewhart control chart (Derman and Ross 1997). Processes are considered “in-control” if the observed mean of a consecutive set of observations is within a specified confidence interval. If the mean falls outside that interval, the process is considered “out-of-control.” Shewhart control charts are very effective at detecting transient causes but less effective for persistent causes because they have no memory of past events. Observations at one time-step are evaluated independently of data from the previous time-step. This deficiency can be circumvented by a series of runs rules (e.g., three out of four signals in a row indicates a persistent cause) or the use of control charts with a memory of past events such as the cumulative sum or moving-average control chart.

Cumulative sum control methods are effective at detecting persistent changes in observed processes (Page 1961). The underlying algorithm is a simple cumulative sum of the deviation of observations from the mean (Hawkins and Olwell 1997). An alternative to CUSUM charts for detecting persistent causes is the moving-average control chart.

The three QC methods evaluated each required the indicators to be standardized \((z\text{-transformed})\) using a control mean and standard deviation. This control mean and standard deviation could be specified as a managerial goal but could also be estimated from historical observations of that indicator. After all nine indicators were calculated for the historical and future phases of the simulation, they were standardized using data from the historical phase, and the QC algorithm applied to each of the ten future years. For each future year the algorithm generated a signal if an out-of-control process was detected. This signal was compared with the actual system state of exploitable biomass at time \( t (B_{\text{ex}}^t) \) defined by the biological reference point \((\phi B_{eq})\). If the exploitable biomass was less than this reference point and the QC algorithm generated an out-of-control signal then a true-positive result was registered. Other combinations of \( (B_{\text{ex}}^t < \phi B_{eq}) \) and QC signal results registered true-negative, false-positive, and false-negative results. These results were incremented for the ten
years in the future phase of the simulation and the simulation replicated 1,000 times. This enabled an estimation of the probability of true-positive, true-negative, false-negative, and false-positive results of the QC algorithm for a particular set of parameter values. Due to dependencies among the future years these were not true probabilities. This could have been avoided by using different replicate simulations for each of the future years, but this would be inconsistent with the use of the scheme for a particular realization of a stochastic process (as would occur in any application of these methods).

**Implementation of the QC algorithms**

1. **Shewhart Control Chart.** This was the simplest to apply and required only a single control parameter, the decision interval \((h)\). A signal is raised at time \(t\) if \(|x_t|>h\), where \(x_t\) is value of the standardized indicator at time \(t\).

2. **Moving-Average Control Chart.** An incremental development to the Shewhart control chart is the calculation of a moving-average of the last \(w\) observations of the indicator, and then application of a Shewhart control chart to these smoothed values. If the absolute value of the smoothed indicator is greater than the decision interval then a signal is raised.

3. **CUSUM Control Chart.** The final QC method examined was the cumulative sum or CUSUM control chart. This method calculated an upper or lower CUSUM \((\pm \phi)\) with the following equations:

\[
\phi_t^+ = \max(0, \phi_{t-1}^+ + x_t - k) \\
\phi_t^- = \min(0, \phi_{t-1}^- + x_t + k),
\]

where \(x_t\) is value of the standardized indicator at time \(t\) and \(k\) is the chart tolerance (or variation that is ignored). These iterative equations were initialized using \(\phi_0^+ = 0\) and \(\phi_0^- = 0\) and when \(|\phi_t^{\pm}|>h\) an out-of-control signal was raised.

Both the moving-average chart and the CUSUM chart required additional parameters \((w\) and \(k\) respectively), but these algorithms were more effective at signaling persistent causes because they captured the memory within a sequence of indicator values.

**Measurement of performance and sensitivity analysis**

Evaluation of the sensitivity and specificity of any type of test can be used to assess the performance of that test. The more sensitive a test, the more reliable it is at signaling an anomalous situation. The more specific a test, the more reliable it is at not signaling a situation that is not anomalous. These are well-established concepts in health research (Sackett et al. 2000, Gigerenzer 2002) but do not appear to have been used in the fisheries literature. Sensitivity and specificity are conditional probabilities that are defined:

\[
sensitivity = P(T^+)/[P(T^+) + P(F^-)]
\]
specificity = \frac{P(T^-)}{P(F^+) + P(T^-)}

where \( P(T^+) \) is the probability of a true-positive result and equivalent notation applies to the other terms. Different values of the decision interval \( h \) will yield different values of sensitivity and specificity for a test and illustrate the trade-off between these characteristics. A plot of the complement of the specificity against the sensitivity over a range of decision intervals \( (h) \) creates a receiver-operator-characteristic (ROC) curve.

The ROC curve is a valuable diagnostic tool because it enables an objective comparison of various indicators and QC algorithms. In this study the overall performance of an indicator and a QC algorithm is measured as the area under the ROC curve (\( \Omega \)). This area was calculated using numerical integration. This performance measure had the additional advantage of integrating-out a nuisance parameter, the decision internal \( h \), from the analysis. The ROC curve was also used to determine the value of the decision interval that yielded a threshold value of sensitivity, and estimate the specificity at that point. These calculations were completed using linear interpolation of available information from the ROC curve. Sensitivity analyses were also completed upon: the number of fish sampled for age- and length-based indicators; the coefficient of variation of recruitment; and the steepness of the stock-recruitment relationship. Consideration was also given to the CUSUM tolerance parameter \( k \), the moving-average parameter \( w \), the reference point \( \phi \), and the probability of an historical impact in the fishery \( P_{hist} \). In all cases the sensitivity analyses are presented in terms of the overall performance (\( \Omega \)).

**Results**

**Illustrative results from a single replicate**

Figure 1 includes correlation plots of the \( \frac{B_t}{B_{eq}} \) versus the indicator values from a single replicate simulation of the operating model. Figure 1a, 1b, and 1c illustrate the age-based, length-based, and abundance indicators respectively. The replicate was a scenario with no historical impacts (\( P_{hist} = 0.0 \)) and surety of future impacts (\( P_{futur} = 1.0 \)). The cluster of points at \( \frac{B_t}{B_{eq}} \approx 1.0 \) illustrates the behavior of the indicators during the historical equilibrium phase and the points when \( \frac{B_t}{B_{eq}} < 1.0 \) show the response of the indicators as the biomass collapses. As expected, there was a decrease in mean length and age as well as an increase in total mortality (Fig. 1a and 1b). The recruitment fraction indicators increased as the stock collapses. Within the abundance indicators there was strong correlation with the survey indicator (\( U_{survey} \)), but commercial catch (\( C_t \)) and \( U_{cpue} \) did not indicate biomass in a robust manner (Fig. 1c).

Results for the indicator mean age (\( \mu_t^{age} \)) generated from this replicate were analyzed with the three QC algorithms examined in this paper. Figure 2a plots \( \mu_t^{age} \) (unstandardized and standardized using data from 0
Figure 1. Relationship between the empirical indicators and $B_t/B_{eq}$: (a) Age-based indicators $Z_t^{age}$ (symbol Z) and $\mu_t^{age}$ (symbol M) are plotted on the left axis scale; $\theta_t^{age}$ (symbol R) is plotted on the right axis scale. (b) Length-based indicators $Z_t^{len}$ (symbol Z), $\mu_t^{len}$ (symbol M), and $\theta_t^{len}$ (symbol R) are plotted in an analogous fashion to those on (a). (c) Abundance indicators $U_t^{survey}$ (symbol S), $U_t^{cpue}$ (symbol U), and $C_t$ (symbol C). Unit notes: $C_t$ (metric tons or t), $U_t^{cpue}$ (tons per effort unit or t/eu), and $U_t^{survey}$ (tons per survey unit or t/su).
Figure 2. Illustration of quality control charts. (a) Mean age (μ_{\text{age}}, solid line) and B/B_{eq} (solid line with symbols) plotted against time. Standardized mean age (dotted line) is also included. Note that B/B_{eq} \approx 0.5 at 15 years. (b) Control charts: Shewhart (dotted line); three year moving-average Shewhart (dashed line); CUSUM (k = 1.0, φ_{t} heavy dashed line, φ_{t} heavy solid line). Signals are generated each time a control value exceeds ±h, i.e., ±1. The CUSUM method generated the strongest signal by year 15.
to 10 years) along with \( B_t/B_{eq} \). The reference point for this analysis was \( B_t/B_{eq} = 0.5 \), and the biomass fell below this value after 15 years. Before this time signals were false-positive and signals on or after 15 years were true-positive. Figure 2b provides the control charts for the standardized \( \mu_t^{age} \) using a decision interval \( h \) of ±1. The Shewhart chart and moving-average (with a three year window) generated four false-positive signals before any impacts occurred (at ten years). The CUSUM chart (using \( k = 1.0 \)) generated a false-positive signal at 7 years and 14 years but gave the strongest signal thereafter.

**Comparison of QC algorithms**

Replicate simulations of the above algorithm enabled the estimation of an ROC curve for a particular indicator and QC algorithm. Figure 3 illustrates these ROC curves for \( \mu_t^{age} \), \( C_t \) and \( U_t^{survey} \) using a CUSUM control chart \( (k = 1.0) \). When the decision interval \( h \) is 0 then all curves have a sensitivity of 1 and a specificity of 0. As \( h \) increased, the sensitivity decreases and specificity increases. Indicators and control schemes that are both sensitive and specific will have the greatest area under the curve (Q). The ROC curve was also used to determine the specificity of an indicator at a certain value of sensitivity (or vice versa) as well as the value of the decision interval \( (h^*) \) that obtained these values. Table 1 summarizes the performance of indicators for Shewhart, CUSUM \( (k = 0.5, 1.0, 1.5) \) and moving-average control charts (with the average calculated over 2, 3, and 4 years). In general the choice of control chart did not have a large effect upon performance of an indicator, though the CUSUM charts always performed marginally better. This table also enables a comparison of indicators. As expected, \( U_t^{survey} \) had the best performance followed by \( \mu_t^{age} \), \( \mu_t^{len} \), \( Z_t^{len} \) (derived directly from \( \mu_t^{len} \), \( \Theta_t^{len} \), \( \Theta_t^{age} \) (all these age- and length-based indicators had very similar performance). Commercial catch rates \( (U_t^{cpue}) \) averaged lower performance across all QC methods than these previous indicators but were superior to \( Z_t^{age} \) and \( C_t \). The number of age samples taken (1,000) was not sufficient for a precise estimate of total mortality. Ironically, \( Z_t^{len} \) was biased, but was a superior indicator to its age-based counterpart within this framework.

The table also presents the value of \( h^* \) and the test specificity when the sensitivity is 0.8. QC algorithms are commonly applied by defining a fixed sensitivity (true-positive rate). The decision interval required for such sensitivity is then estimated along with the test specificity. In time-series models, where assumptions of statistical independence between observations are invalid, simulation can be used to estimate these threshold values (otherwise analytical methods are available, Hawkins and Olwell 1997). The \( h^* \) values in Table 1 provide some insight in the decision intervals that could be used for the three QC algorithms and how specific the signals would be. These numerical values are, however, only valid for this simulated fishery. Note that for the better performing
indicators such as $\mu_{ie}^{age}$, the $h^*$ values are larger and are associated with greater specificity than with the inferior indicators such as $C_i$.

**Reference points, effect size and effect timing**

The reference point used for this analysis (by default 0.5 $B_{eq}$ or $\phi = 0.5$) had a crucial role in this analysis and was analogous to an effect size in power analysis. Larger defined effects will be easier to accurately signal. Performance of indicators as a function of $\phi$ was estimated using the CUSUM scheme ($k = 1.0$) and the results plotted on Fig. 4a. As the effect size decreased (larger values of $\phi$) the performance of all indicators decreased except surveys that continued to improve until $\phi = 0.8$. Large effects degraded the performance of survey indicator because of poor specificity (false-positive rate). Performance of the indicator total
Table 1  Performance (Ω) of empirical indicators for Shewhart, CUSUM, and moving-average control charts. Values of the specificity* and decision interval (h*) at 80% sensitivity are also tabulated.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Metric</th>
<th>Shewhart</th>
<th>CUSUM</th>
<th>Moving-average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>k = 0.5</td>
<td>k = 1.0</td>
<td>k = 1.5</td>
</tr>
<tr>
<td>$\mu_{t}^{age}$</td>
<td>Ω (%)</td>
<td>93</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>5.8</td>
<td>11.6</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>$\theta_{t}^{age}$</td>
<td>Ω (%)</td>
<td>92</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>4.1</td>
<td>7.8</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>93</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>$Z_{t}^{age}$</td>
<td>Ω (%)</td>
<td>77</td>
<td>83</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>0.8</td>
<td>1.7</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>49</td>
<td>68</td>
<td>61</td>
</tr>
<tr>
<td>$C_{t}$</td>
<td>Ω (%)</td>
<td>77</td>
<td>85</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>1.4</td>
<td>3.5</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>54</td>
<td>71</td>
<td>68</td>
</tr>
<tr>
<td>$U_{t}^{cue}$</td>
<td>Ω (%)</td>
<td>88</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>2.4</td>
<td>4.2</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>84</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>$\mu_{t}^{len}$</td>
<td>Ω (%)</td>
<td>93</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>6.2</td>
<td>11.9</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>93</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td>$\theta_{t}^{len}$</td>
<td>Ω (%)</td>
<td>92</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>5.0</td>
<td>9.5</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>92</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>$Z_{t}^{len}$</td>
<td>Ω (%)</td>
<td>93</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>7.4</td>
<td>13.7</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>$U_{t}^{survey}$</td>
<td>Ω (%)</td>
<td>95</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>$h^*$</td>
<td>11.8</td>
<td>&gt; 30.0</td>
<td>&gt; 30.0</td>
</tr>
<tr>
<td></td>
<td>specificity* (%)</td>
<td>91</td>
<td>&lt; 93</td>
<td>&lt; 94</td>
</tr>
</tbody>
</table>
Figure 4. (a) Performance of indicators as a function of the reference point $\phi_{Beq}$. Most indicators degrade as the effect size decreases. (b) Performance of indicators as a function of the probability of a historical impact $p_{hist}$. When $p_{hist} \geq 0.4$ the performance of all indicators became destabilized. Length-based indicators performed similarly to $\mu_{t}^{age}$ and $\theta_{t}^{age}$ in both analyses. Symbols used: $U_{t}^{survey}$ (symbol S), $U_{t}^{cpue}$ (symbol U), $C_{t}$ (symbol C), $Z_{t}^{age}$ (symbol Z), $\mu_{t}^{age}$ (symbol M), and $\theta_{t}^{age}$ (symbol R).
mortality ($Z^{age}_{t}$) degraded badly when attempting to detect small changes to the underlying stock.

The controlling mean and standard deviation were calculated using observations from the historical years. Simulations to this point have assumed there were no impacts during this time ($P^{hist} = 0.0$). Figure 4b illustrates the effect on indicator performance when there have been changes to the stock during the historical years. Up to $P^{hist} \approx 0.4$ there was little effect on indicator performance, however; beyond this value performance changed in unpredictable ways. For example, the performance of $U^{survey}_{t}$ decreased but that of $U^{cpue}_{t}$ increased. Large amounts of historical variation will degrade the performance of all indicators using this scheme. These analyses on $\phi$ and $P^{hist}$ were only completed for the CUSUM scheme, but the magnitude of the effects was far greater than the differences between the three QC algorithms considered. Similar patterns would be expected for the Shewhart and moving-average methods.

**Sensitivity analysis**

Using the CUSUM ($k = 1.0$) algorithm, results for four parameters are presented (Table 2): the number of fish aged; the number of fish measured; recruitment variability; and the steepness of the stock-recruitment parameter. Increasing the numbers of fish aged and measured increased the precision and accuracy of all of the age- and length-based indicators respectively with the exception of $Z^{len}_{t}$ (which was always biased). There was, however, small marginal benefit from increasing samples from 1,000 to 10,000 fish. Most of increase in performance was obtained within 100–200 fish for all indicators except $Z^{age}_{t}$, which continued to increase in performance up to 10,000 fish. Actual fish populations would be expected to illustrate much greater variability in age and length structure than this model (which was not spatially structured and used an annual time-step). Nevertheless, these age- and length-based indicators appeared to perform very well with effective sample sizes in the hundreds rather than thousands of fish. The small (1%) performance variation in the indicators that were not age- or length-based reflected the random variation from different sequences of random numbers used in the simulations.

Table 2 also presents a sensitivity analysis on recruitment dynamics. Performance of indicators was extremely robust to changes in the steepness of the stock-recruitment parameter. Only very large values of the coefficient of variation of recruitment (50%) appeared to degrade the performance of these indicators. Other sensitivity results (not presented) were very predictable. For example, increasing the variability of surveys caused a marginal decrease in performance of $U^{survey}_{t}$. The results presented appeared to be robust to reasonable changes in the parameter values.
Discussion

The lack of contrast in data is a major stumbling block for stock assessment in data-limited fisheries. In such situations, it may not be practicable to estimate fishing mortality rates or stock biomass from an index-of-abundance, landings, and age- or length-composition data sets. Alternative methods for identifying and signaling important trends in empirical stock-status indicators require examination. This study tested the application of quality control algorithms to detect signals in such empirical indicators and showed that QC-based systems using indicators of average age and length perform exceptionally well, with the integrated area under an ROC curve of between 90 and 95% under a wide range of situations. Indicators based upon recruitment fractions had similar performance. Averages can be calculated precisely with hundreds (rather than thousands) of samples, so detection of a persistent shift in these statistics is feasible. Mean lengths and ages also dampen transient signals resulting from variation in recruitment. Empirical indicators such as CPUE and catch performed worse than age- and length-based indicators, though a theoretical CPUE indicator (the ratio of actual catch and actual effort, rather than recorded effort) performed similarly to surveys. Total mortality estimated from catch-curve analysis was not a robust indicator unless the estimates used a very large sample size (many thousands of fish). Note that in real-world studies, there will be the usual challenge of obtaining representative samples of fish age and length. The above

<table>
<thead>
<tr>
<th>Number of fish sampled for age- and length-based indicators</th>
<th>Coefficient of variation of recruitment</th>
<th>Steepness of stock-recruitment relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>10            100               1,000            10,000</td>
<td>0.01    0.1   0.2   0.5</td>
<td>0.55  0.75  0.95</td>
</tr>
<tr>
<td>$\mu_{t}^{age}$</td>
<td>78 92 94 95</td>
<td>94 94 92 86</td>
</tr>
<tr>
<td>$\theta_{t}^{age}$</td>
<td>75 90 93 94</td>
<td>94 93 90 82</td>
</tr>
<tr>
<td>$Z_{t}^{age}$</td>
<td>57 77 82 90</td>
<td>82 82 81 75</td>
</tr>
<tr>
<td>$C_{t}$</td>
<td>83 83 84 83</td>
<td>83 83 82 77</td>
</tr>
<tr>
<td>$U_{t}^{cpue}$</td>
<td>89 90 90 90</td>
<td>90 90 89 83</td>
</tr>
<tr>
<td>$\mu_{t}^{len}$</td>
<td>76 90 93 94</td>
<td>94 94 92 85</td>
</tr>
<tr>
<td>$\theta_{t}^{len}$</td>
<td>71 88 92 94</td>
<td>94 93 91 83</td>
</tr>
<tr>
<td>$Z_{t}^{len}$</td>
<td>78 90 93 94</td>
<td>94 94 92 86</td>
</tr>
<tr>
<td>$U_{t}^{survey}$</td>
<td>95 96 96 95</td>
<td>95 95 95 94</td>
</tr>
</tbody>
</table>
estimates of sample size have not considered spatial and within-year variation, both of which are likely to substantially increase the number of samples actually required.

There are three potential applications of the work presented here: (1) a simple signaling system for empirical indicators to prioritize research and management; (2) approximate calibration of observations to dynamic models and the use of these, along with a specification of required sensitivity/specificity to determine decision intervals and reference points; and (3) use of QC algorithms within decision-rules to ensure sustainable harvesting rates. These applications are discussed in turn.

Empirical indicators lie at the base of any type of stock assessment. In most cases (even within data-limited situations) there will be sufficient information on catch, CPUE, or length-composition to estimate a controlling mean and standard deviation. Estimation of these statistics, even from 2-3 years of data, will enable the standardization of new observations. QC algorithms could then be applied as an objective procedure to signal if the most recent observation is uncharacteristic and action should be taken. This approach is at the heart of any monitoring system. In most situations there will be ongoing uncertainty about the value of the decision interval $h$ and the appropriateness of the controlling mean and standard deviation. If the recent historical state of a fishery is severely depleted, then there will be risks associated with using such information to estimate the controlling mean and standard deviation. In such cases, the decision interval could be asymmetrical, so that a signal is generated if there is any decrease in an indicator such as mean length or CPUE.

Results presented in Table 1 provide some guidelines on the decision interval values (for a sensitivity of 80%) but these vary greatly according to the indicator and the desired sensitivity. Rather than provide numerically specific recommendations about the application of these methods, the author recommends that interested readers compile relevant examples from their fisheries in a simple electronic spreadsheet to understand the trade-offs that occur when using these algorithms. Specification of the decision interval will depend upon the managerial context (what are the actions resulting from a signal), attitudes toward (and consequences of) false-positive and false-negative signals, and the historical context of the fishery (see above). In many data-limited cases, it is likely that the decision interval will need to be negotiated among the stakeholders of the fishery. Such an approach will certainly not be the basis of a rigorous stock assessment but will enable rapid identification of stocks where things are starting, or continuing, to go astray. Inferences will also be much stronger if there are two or more independent empirical indicators available (also see Caddy 1999, Scandol 2003). The simplicity of these schemes would enable them to be integrated into the reporting algorithms of database management systems that would (or should) be used to store the data required to compile empirical indicators.
The second, and much more technically sophisticated, application will require a meaningful calibration of a dynamic model to a particular fishery. This was the general approach taken by Punt et al. (2001). Once an estimated relationship between an empirical and estimated indicator is available, then the standard limit/target biological reference points (BRPs) such as $B/B_0 = 0.2$ or $F_{0.1}$ can be converted to values of an empirical indicator or indicators. The degree of precaution or risk-aversion to be applied with the management system then could be expressed via the sensitivity of the test and the corresponding value of the decision interval $h$. This signal system could then replace a BRP target/limit system. The main advantage that such a scheme could have over standard BRP management systems is improved transparency.

The third potential application of QC methods is within the managerial decision-rules that are now advocated by some fishery scientists (e.g., Starr et al. 1997, Hilborn 2002). Starr et al. (1997) used rules based upon moving-averages of CPUE to adjust the total allowable catch in a fishery to support sustainability. Similarly, the values from a CUSUM chart could be used to adjust fishing mortality using either input or output controls. Schemes that use control loops based upon indicators derived from age- or length-composition data should be particularly valuable as they will be more robust to changes in catchability. CUSUM charts are likely to be very effective because they deliver very strong signals if a persistent impact occurs (see Fig. 2b). Systems could be designed that are robust to uncertainties in most parameters.

Many fisheries management agencies that are dealing with data-limited fisheries have limited human resources to develop sophisticated stock assessments. There will, however, often be statutory requirements for performance reporting of fisheries. Simple graphical charts of empirical indicators will therefore be a mandatory and achievable goal. Charts are also a transparent and rapid method of communication. The research presented here introduces some simple ideas from the statistics of quality control to assist in the quantitative interpretation of these charts of empirical indicators. Such methods can be applied in a very simple manner but more sophisticated extensions are also feasible.

**Acknowledgments**

The author would like to thank Geoff Liggins and two anonymous reviewers for their valuable comments on this manuscript. Financial support from the Australian Centre for International Agricultural Research (ACIAR) enabled the author to present this paper in Anchorage, Alaska.
References


Appendix: The operating model

Simulations were completed using the following age-structured model. Parameter values were, where possible, estimated from observations of yellowfin bream (Acanthopagrus australis) from Gray et al. (2000) or given assumed values. Default values of these parameters are given below. Average individual fish length (cm) at age $a$ was represented with:

$$T_a = 29.0 \left[ 1 - \exp\left[ -0.36(a+1) \right] \right]$$

Lengths were converted to weights (in kg) with $w_a = 2.48 \times 10^{-5} l_a^3.0$. All selectivity and maturity schedules were based upon the logistic function:

$$\Lambda(x, x_{50}, x_{95}) = \left[ 1 + \exp\left( -\ln(19)(x - x_{50}) / (x_{95} - x_{50}) \right) \right]^{-1}$$

Vulnerability at length was $v_{l,\text{len}} = \Lambda(l, 24.4 \text{ cm}, 29.4 \text{ cm})$; vulnerability at age was $v_{a,\text{age}} = \Lambda(a, 4.7 \text{ year}, 8.8 \text{ year})$; maturity at age was $m_{a,\text{age}} = \Lambda(a, 3.7 \text{ year}, 5.3 \text{ year})$; and for surveys the vulnerability at age was $v_{a,\text{surv}} = \Lambda(a, 0.5 \text{ year}, 1.0 \text{ year})$.

Exploitable and spawning biomass were respectively represented with:

$$B_{i,\text{ex}} = \sum_{a=0}^{A} N_{a,i} a_{a,\text{age}} v_{a,\text{age}} w_a$$
$$B_{i,\text{sp}} = \sum_{a=0}^{A} N_{a,i} a_{a,\text{age}} m_{a,\text{age}} w_a$$

The maximum number of age classes ($A$) was 20 years and no “plus group” was included in this model. Standard deviation of length at age was $\sigma_a = 1.9 + 0.1a$ (cm) and the elements within the age-length key were calculated with

$$P_{l,j} = \int_{l_j}^{l_j + \frac{1}{2}} \frac{1}{\sqrt{2\pi} \sigma_j} \exp\left[ -\frac{(x - T_j)^2}{2(\sigma_j)^2} \right] dx$$

This integral was calculated over each of the 50 $\times$ 1 cm length classes. Distribution of fish at length $l$ at time $t$ was calculated using

$$N_{l,t}^\text{len} = \sum_{i=0}^{A} P_{l,j} N_{a,i} a_{a,\text{age}}$$

The model was initialized with $N_{a,0} a_{a,\text{age}} = R_0 \exp(-M \times a)$ where $R_0 = 1,000 \times 10^3$ fish and $M = 0.2$ year$^{-1}$. Initial spawning numbers were calculated

$$N_{0}^\text{sp} = \sum_{i=0}^{A} m_{a,i} a_{a,\text{age}} N_{a,0} a_{a,\text{age}}$$

which enabled parameterization of the stock-recruitment relationship using

$$A^\text{sr} = \frac{R_0}{N_0^\text{sp}} \left[ 1 - \frac{z - 0.2}{0.8} \right] \quad \text{and} \quad B^\text{sr} = \frac{z - 0.2}{0.8zN_0^\text{sp}}.$$
The steepness parameter $z$ had a default value of 0.75. Mean recruitment into the $t+1$ year was given by $\bar{R}_{t+1} = N_{t+1}^{sp}/(A^{sp} + B^{sp} N_{t}^{sp})$.

The age-structured population was updated with

$$N_{t}^{age} = N_{t-1}^{age} \exp[-(M + F_{a,t})]$$

(the fishing model for $F$ is described below). Catch was calculated using

$$C_{t}^{age} = \frac{F_{a,t}}{M + F_{a,t}} \times N_{t}^{age} \times \{1 - \exp[-(M + F_{a,t})]\}.$$

Numbers of fish in the zero age class were specified with

$$N_{0,t} = \bar{R}_{t} \exp(R_{\nu, \epsilon} - R_{\nu, \epsilon}^{2}/2),$$

where $\epsilon$ is a normally distributed random variable and $R_{\nu}$ has a default value of 10% (sensitivity analysis considered larger values of this parameter). Total fishing mortality $F$ was the sum of individual activity from 50 fishers with individual effort ($e_{k,t}$) distributed as a random exponential variable (mean 0.4 units of effort). Individual catchability was lognormally distributed (mean 0.1, cv 50%). This gave an average fishing mortality ($F$) of 0.2 year$^{-1}$. Fishing mortality for an age class was assumed to be proportional to the vulnerability of that age class.

Impacts on a fishery were modeled by the generation of a random uniform variate ($u$) for each impact variable: natural mortality; catchability; effort; and recruitment (future impacts only). The geometrically imposed rate of impact $r$ per year was

$$\text{if } (u \geq 1 - P) \text{ then } r = 0.2(u - P - 1)P \text{ else } r = 0$$

where $P$ was the probability of an impact (either $P_{\text{hist}}$ or $P_{\text{futr}}$).

Indicators were calculated using the following equations:

$$C_{t} = \sum_{a=0}^{A} C_{t}^{age} w_{a}$$

$$U_{t}^{\text{cpe}} = C_{t} / \sum_{k} \text{int}(e_{k,t} + 1) \text{ (tons per effort unit or t/eu; note the corruption of effort data)}$$

$$U_{t}^{\text{survey}} = 0.001 \left( \sum_{a=0}^{A} \nu_{a}^{\text{survey}} N_{a,t}^{age} w_{a} \right) [\exp(0.05 \epsilon - 0.05 \epsilon^{2} / 2)]$$

(tons per survey unit, or t/su).

Age- and length-based indicators were derived from random samples of fish (1,000 age samples, 3,000 length samples) drawn from the catch. $Z_{t}^{age}$ was calculated using catch-curve analysis (with a threshold for recruitment of 4 years); $\nu_{t}^{age}$ and $\nu_{t}^{len}$ were calculated using the usual arithmetic equations; $\theta_{a,t}^{age}$ and $\theta_{a,t}^{len}$ were the fraction of the samples less than or equal to 3 years and 22 cm respectively. Finally $Z_{t}^{len} = 0.36(29.0 - \nu_{t}^{len})/\nu_{t}^{len} - 19.4)$. 


Per-Recruit Simulation as a Rapid Assessment Tool for a Multispecies Small-Scale Fishery in Lake Malombe, Malawi, Africa

Olaf L.F. Weyl and Anthony J. Booth
Rhodes University, Department of Ichthyology and Fisheries Science, Grahamstown, South Africa

Kissa Mwakinyongo¹ and Moses M. Banda
Department of Fisheries, Fisheries Research Unit, Monkey Bay, Malawi

Abstract

The 390 km² Lake Malombe supports a small-scale fishery that is dominated by the nkacha net, a locally developed purse seine. The nkacha fishery contributes in excess of 95% to the haplochromine cichlid catch from Lake Malombe. Since 1990 the annual haplochromine cichlid yield declined from 9,500 t to less than 4,000 t. In Lake Malombe, as in most African fisheries, the use of assessment methods that allow for dynamic simulation analysis incorporating the response of the stock to changes in management strategy was negated by the lack of directed age-based catch data and the lack of information pertaining to the biology of the target species. In this study, a rapid assessment framework for the assessment of biological and fishery input-parameters needed for the application of multispecies yield-(YPR) and spawner-biomass per recruit (SBR) analysis is presented. During a one-year assessment period, species selectivity by the nkacha fishery was determined. Length-frequency analysis allowed for a first estimate of growth rate for five major target species. These growth rate data were then used to estimate age-selectivity into the fishery, mortality rates, and age-specific maturity for each of the target species. Subsequently, YPR and SBR analyses were used to investigate commonly used “target reference points” as management targets for five target species in the fishery. The results are discussed with reference to the application

¹Deceased.
of per-recruit analysis for the rapid derivation of management advice for multispecies fisheries in data limited situations.

Introduction

Lake Malombe (14°30'–14°45'S to 35°12'–35°20'E) is situated in the Mangochi District of Malawi (Fig. 1). The lake is a shallow, 390 km² widening of the Upper Shire River about 15 km from its outflow from Lake Malawi, and provides a livelihood for about 2,000 small-scale fishers (Weyl et al. 2001b). While catch and effort has been monitored by the Malawi Department of Fisheries since 1976 (Tweddle et al. 1995), the monitoring system aggregates more than 60 species into 13 commercial categories (Tweddle et al. 1995). Despite this aggregation, major changes in the fishery have been detected. The fishery for tilapiine cichlids, which contributed more than 6,000 tons to the total catch in 1982, collapsed in 1991 and was replaced by a fishery for a variety of small haplochromine cichlid species (Tweddle et al. 1995). This haplochromine fishery yielded 9,500 tons in 1990, but by the mid-1990s had declined to less than 4,000 tons annually (Weyl et al. 2001a). Since a major objective of the Malawi fisheries policy is “the maximization of harvests within safe sustainable yield levels” (Government of Malawi 1999), the decline in this fishery necessitated a stock assessment.

Fishing gear used to harvest haplochromine cichlids on Lake Malombe includes gillnets, beach seines, purse seines, hooks, and traps (Weyl et al. 2004). Despite this diversity in gear use, 147 nkacha nets (purse-seine type gears that are operated offshore) contributed more than 95% to the total haplochromine catch (Weyl et al. 2001a). Nkacha nets utilize mesh sizes ≤ 19 mm and 57 species have been recorded in their catches (Weyl et al. 2004). While the species composition in the catch is highly diverse, five species, *Copadichromis chrysonotus*, *Copadichromis virginalis*, *Lethrinops turneri*, *Otopharynx argyrosoma “ssp. red”* (informal subspecies name after Turner [1996]), and *Otopharynx tetrastigma*, contributed more than 60% to the catch (Weyl et al. 2004).

The presence of an existing fishery on the lake requires the use of assessment methods that allow for dynamic simulation analysis incorporating the response of the multispecies stock to changes in management strategy (e.g., a change in selectivity or effort). However, the choice of suitable assessment methods was constrained by the lack of directed catch and effort data, length- or age-based catch data, and other biological data pertinent to the application of age-based models. Further complications arose from the high cost associated with the collection of these data and the immediate need for management recommendations. For this reason, a cost-effective stock-assessment technique that is not data intensive is required.
As a result of similar constraints, fisheries managers in developing countries have focused on the application of the yield-per-recruit (YPR) model, which is an abbreviation of the full dynamic-pool model (Beverton and Holt 1956, 1957), in the management of lacustrine fisheries (Amarasinghe and De Silva 1992, Manyala et al. 1995, Thompson and Allison 1997). The application of these YPR models for the assessment of predefined fishing mortality targets (target reference points—TRPs) has become common practice in fisheries management (Clarke 1991, Punt 1993, Punt and Butterworth 1993, Caddy and Mahon 1995). The YPR approach allows for the determination of at least two commonly used TRPs: first, the fishing mortality which corresponds to the maximum of

Figure 1. Map of Lake Malombe showing Mwalija and Chapola fish landing sites where nkacha net catches were sampled between April 2000 and March 2001.
the yield-per-recruit curve ($F_{\text{MAX}}$) and second, the marginal yield or $F_{0.1}$ strategy (Gulland and Boerema 1973, Deriso 1987), which is the rate of fishing mortality at which the slope of the yield-per-recruit curve falls to 10% of its value at the origin. However, the assumption that recruitment is constant and independent of spawner-biomass has led to some criticism of YPR-based TRPs and scientists concerned with the management of marine species have tended to base their TRP recommendations on the results of spawner biomass-per-recruit (SBR) models (Butterworth et al. 1989, Smale and Punt 1991).

In the absence of information on the surplus production function or the spawner biomass-recruitment relationship, SBR-based TRPs are currently considered the most robust, allowing for the determination of a fishing mortality rate that will provide relatively high yields at lower risks (Clarke 1991, Punt 1993). The definition of a SBR-based TRP ($F_{SB(x)}$)

### Table 1. Input parameters used for the application of yield- and spawner-biomass per recruit models on *Copadichromis chrysonotus* (CC), *Copadichromis virginalis* (CV), *Lethrinops turneri* (LP), *Otopharynx argyrosoma* “ssp. red,” (OA) and *Otopharynx tetrastigma* (OT) in Lake Malombe, Malawi.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>CC</th>
<th>CV</th>
<th>LT</th>
<th>OA</th>
<th>OT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_\infty$</td>
<td>140</td>
<td>118</td>
<td>144</td>
<td>149</td>
<td>139</td>
</tr>
<tr>
<td>$K$</td>
<td>1.0</td>
<td>0.61</td>
<td>0.38</td>
<td>0.46</td>
<td>0.44</td>
</tr>
<tr>
<td>$t_0$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.000083</td>
<td>0.00021</td>
<td>0.000043</td>
<td>0.000195</td>
<td>0.0000103</td>
</tr>
<tr>
<td>$\beta$</td>
<td>2.58</td>
<td>2.33</td>
<td>2.70</td>
<td>2.33</td>
<td>3.03</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.73</td>
<td>1.02</td>
<td>2.54</td>
<td>1.09</td>
<td>1.65</td>
</tr>
<tr>
<td>$\sigma_m$</td>
<td>0.12</td>
<td>0.20</td>
<td>1.23</td>
<td>0.25</td>
<td>0.27</td>
</tr>
<tr>
<td>$t_r$</td>
<td>0.57</td>
<td>1.07</td>
<td>1.03</td>
<td>0.84</td>
<td>1.11</td>
</tr>
<tr>
<td>$\sigma_r$</td>
<td>0.07</td>
<td>0.13</td>
<td>0.07</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>$M$</td>
<td>1.31</td>
<td>0.7</td>
<td>0.3</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>$F_{\text{CUR}}$</td>
<td>1.36</td>
<td>1.09</td>
<td>1.65</td>
<td>1.02</td>
<td>0.19</td>
</tr>
<tr>
<td>$F_{\text{CUR}}^{\text{max}}$</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>$q$</td>
<td>0.009251701</td>
<td>0.007415</td>
<td>0.011224</td>
<td>0.006939</td>
<td>0.001293</td>
</tr>
<tr>
<td>$R_i$</td>
<td>1.52E+08</td>
<td>1.94E+08</td>
<td>3.93E+08</td>
<td>1.38E+08</td>
<td>2.14E+08</td>
</tr>
</tbody>
</table>

*Parameter estimates derived from Weyl et al. (2005). $L_\infty$ = asymptotic length (mm TL), $K$ = Brody growth coefficient (yr$^{-1}$), $t_0$ = age at zero length, $\alpha$ and $\beta$ = length weight parameters, $\phi$ = age-at-50%-maturity (years), $\sigma_m$ = variance of logistic maturity ogive, $t_r$ = age-at-50%-selection (years), $\sigma_r$ = variance of logistic selection ogive (years), $M$ = natural mortality rate (yr$^{-1}$), $F_{\text{CUR}}$ = current fishing mortality rate (yr$^{-1}$), $F_{\text{CUR}}^{\text{max}}$ = maximum age of fish (years), $q$ = catchability coefficient, $R_i$ = estimated recruitment.*
involves setting the fishing mortality to a level at which spawner biomass-per-recruit is reduced to \( x\% \) of its pristine level. Although there is no conventional \( F_{SB(x)} \) TRP, the maintenance of SBR at between 25% and 50% of unexploited levels has been recommended (Deriso 1987, Sissenwine and Shepherd 1987, Butterworth et al. 1989, Punt 1993, Booth 2004). Because haplochromine cichlids are mouth-brooders (Turner 1996), strong density dependence between spawner-biomass and recruitment is implied and the maintenance of SBR at levels where the stock can replace itself is likely to be important for the sustainability of haplochromine-based fisheries.

This study utilizes data derived from a one-year survey of the nkacha net fishery to derive input parameters for, and apply YPR and SBR models to assess, four commonly used TRPs for the derivation of management advice for \( C. \) chrysonotus, \( C. \) virginalis, \( L. \) turneri, \( O. \) argyrosoma “ssp. red,” and \( O. \) tetrastigma in the Lake Malombe nkacha net fishery.

**Methods**

The application of per-recruit analysis requires age-based estimates for mortality, maturity, selectivity, and the length-weight relationship. *Copadichromis chrysonotus* has been assessed using per-recruit analysis and all input parameters are available (Weyl et al. 2005). For the other four species, \( C. \) virginalis, \( L. \) turneri, \( O. \) argyrosoma “ssp. red,” and \( O. \) tetrastigma, the length-weight relationship and length-based selectivity and maturity parameters were available in the literature (Banda 1995, Weyl et al. 2004).

**Growth estimation**

Due to the use of small mesh sizes (≤ 19 mm) by the nkacha net fishery resulting in negligible differences in size-selectivity (Weyl et al. 2004), the length-frequency of the target species in this gear showed discernable modal progression over time and was therefore considered suitable for the estimation of length-at-age (\( l_t \)) using length-frequency analysis (Pauly and David 1981, Wetherall 1986, Shepherd 1987).

Monthly length-frequency samples were collected from the nkacha net fishery at two major landing sites on the western shore of the lake, for three consecutive days a month from March 2000 to April 2001. The methods used during this survey are detailed in Weyl et al. (2004, 2005).

Length-frequencies of \( C. \) virginalis, \( L. \) turneri, \( O. \) argyrosoma “ssp. red,” and \( O. \) tetrastigma in sampled catches were regrouped into 5 mm size classes and raised to represent the length-frequency of the estimated catch of the species during each sampling period using the methods outlined by Weyl et al. (2005) for \( C. \) chrysonotus.
To estimate length-at-age \( (l_t) \), length-frequency analysis (Pauly and David 1981, Wetherall 1986, Shepherd 1987) was used to derive estimates for asymptotic length \( (L_\infty) \) and the growth coefficient \( (K) \) for the von Bertalanffy growth model: 
\[
l_t = L_\infty (1 - \exp(-K(t-t_0)))
\]
(Ricker 1975). A preliminary estimate of the growth parameter \( L_\infty \) was obtained from pooled monthly length-frequency samples using the Powell-Wetherall method (Wetherall 1986). Electronic length frequency analysis (ELEFAN) (Pauly and David 1981) was used to derive a first estimate of \( K \). This analysis was performed using the FAO-ICLARM Fisheries Assessment Tools (FiSAT) software (Gayanilo et al. 1997). There was insufficient resolution in the data to calculate the age-at-zero length \( (t_0) \) and this parameter was, therefore, taken to be zero.

**Mortality**

A first approximation of the instantaneous rate of total mortality \( (Z) \) was estimated by catch-curve analysis (Ricker 1975). Annualized length-frequency data presented in Weyl et al. (2004) were analyzed by means of a linearized length-converted catch curve (Pauly 1983, 1984a,b). This method uses the von Bertalanffy growth parameters to plot \( \ln(f/dt) \) against \( t \), where \( f \) is the frequency of individuals in each length class and \( t \) is the relative age of the fish. The value \( dt \) is the time taken for the fish to grow through a particular length class and allows for decreased growth with increased age. The negative of the slope of the resultant linear regression line through the descending data points gives a first approximation of \( Z \).

Natural mortality \( (M) \) was estimated using the Pauly (1980) empirical equation; 
\[
M = -0.0152 - 0.279 \ln L_\infty + 0.6543 \ln K + 0.463 \ln T;
\]
where \( L_\infty \) (cm) and \( K \) are the von Bertalanffy growth parameters and \( T \) is the mean annual water surface temperature (26.2°C, calculated from monthly temperature readings presented in Jambo [1997]). The estimates for \( M \) derived using the Pauly (1980) equation were similar to independent estimates for closely related species in Lake Malawi (Tweddle and Turner 1977) and were accepted as a first estimate for the five species.

The fishing mortality rate was derived by subtraction \( (F = Z - M) \). Because the fishery targets all five species, the coefficients of proportionality between fishing effort and fishing mortality (i.e., the catchability coefficients) will vary between species due to differences in their availability and vulnerability to the gear (Murawski 1984). Therefore, at any given level of effort, the \( F \) for each species in a multispecies fishery will be different. Catchability coefficients were estimated using the linear relationship, 
\[
F_i = q_i \times f,
\]
where \( q_i \) is the catchability coefficient of species \( i \), and \( f \) is the total effort by all gears in the fishery during the assessment year (i.e., 147 nkacha nets enumerated in 2001).
**Maturity and selectivity**

The proportion of mature fish-at-age $\psi_t$ was estimated by age-converting the Banda (1995) length-based logistic maturity ogives as

$$t = t_0 - \frac{1}{K} \ln \left( 1 - \frac{L_t}{L_w} \right).$$

The age-converted ogive width parameter $\sigma_t$ was calculated from the length-based equivalents as

$$\sigma_t = t_0 - \frac{1}{K} \ln \left( 1 - \frac{\sigma_t \ln 3}{L_w - \phi_t} \right)$$

(Booth and Weyl 2004).

Age conversion of the length-based selectivity for the four species provided in Weyl et al. (2004) was calculated using the maturity methods described above.

**Per-recruit analysis**

An assumption of per-recruit analysis is that the parameters for recruitment, growth, and natural mortality are constant from one year to the next and, therefore, the stock is in a steady state. Under these assumptions, yield-per-recruit [$YPR(f)$] and spawner biomass-per-recruit [$SBR(f)$] of any species $i$ as functions of fishing effort $f$, were calculated as:

$$YPR(f)_i = \max_{t_r} \int_{t_r}^{\max} N_{t,i,j} W_{t,i,j} S_t q_t f dt$$

$$SBR(f)_i = \max_{t_r} \int_{t_r}^{\max} N_{t,i,j} W_{t,i,j} \psi_{t,i,j} dt$$

In all YPR and SBR models, for any species $i$, $W_t$ is the mass-at-age $t$ (derived from the relationship $W_t = \alpha(l_t)^\beta$ where $l_t$ is length-at-age, $\alpha$ and $\beta$ are parameters describing the length-weight relationship), $S_t$ is the selectivity-at-age, $t_r$ is the age of fish first recruiting into the fishery, $\psi_t$ is the maturity-at-age, $q_t$ is the catchability coefficient, $f$ is the fishing effort (number of nkacha nets) and $\text{max}$ the maximum recorded age. Because no direct estimate of age through hard part analysis was possible, the parameter $\text{max}$ was set at 5 years for all species as this was the maximum number of modes seen in the length frequency distributions (Fig. 2).
Figure 2. Monthly length-frequency (bars) and fitted von Bertalanffy growth curves (lines) for *Copadichromis virginalis* ($L_{\infty} = 118$ mm TL, $K = 0.61$, $t_0 = 0$), *Lethrinops turneri* ($L_{\infty} = 144$ mm TL, $K = 0.38$, $t_0 = 0$), *Otopharynx argyrosoma* “ssp. red” ($L_{\infty} = 149$ mm TL, $K = 0.46$, $t_0 = 0$), and *O. tetrastigma* ($L_{\infty} = 139$ mm TL, $K = 0.44$, $t_0 = 0$) sampled from the Lake Malombe, Malawi, nkacha net fishery between April 2000 and April 2001.
Numbers at age \( t \) \((N_t)\), were calculated as \( N_{t,i} = e^{-(t_r,i)M} - (a-t_r,i)(S_{i,q}f+M) \) where \( M \) is the age-independent rate of natural mortality. The YPR and SBR integrals were solved numerically using Simpson’s rule with 50 steps.

To calculate total YPR for all five species at any given level of \( f \), \( TYPR_f \), the \( YPR(f) \_i \) of each species \( i \) was corrected according to the estimated recruitment \( R_i \) and \( TYPR_f \) was then additive, such that:

\[
TYPR_f = \sum_{i=1}^{5} YPR(f) \_i
\]

Recruitment for species \( i \), \( R_i \) was obtained by dividing the total catch of each species (\( Y \)) during the 2001 calendar year by the total yield-per-recruit (\( YPR_i \)) from the fishery determined at “base case” fishing and natural mortality scenarios.

Target reference points

Four target reference points (TRPs) were investigated for the YPR and SBR curves: \( F_{\text{MAX}} \) (that fishing mortality corresponding to the maximum of the YPR curve); \( F_{0.1} \) (where the slope of the YPR curve is 10% of that at the origin); \( F_{SB50} \) and \( F_{SB35} \) (the fishing mortalities that correspond to a reduction in the SBR curve to 50% and 35% of its unexploited equilibrium level).

Due to the inherent difficulty in the estimation of the instantaneous rate of natural mortality (\( M \)), the sensitivity of the YPR and SBR models to changes in the rate \( M \) was assessed using a Monte Carlo estimation procedure (Manly 1998) described in Weyl et al. (2004). In this procedure, 1,000 \((U_{1000})\) random mortality samples \((M_U, U=1,2,...,U_{1000})\) were generated, with a normally distributed error structure around the “base case” \( M \) of each species. A coefficient of variation (CV) of 25% around the \( M \)-estimate was utilized as this represented the CV of five \( M \) estimates derived for similar cichlid species in Lake Malawi (Tweddle and Turner 1977). Subsequently, a corresponding set of \( \hat{F}_1, \hat{F}_2, ..., \hat{F}_{U_{1000}} \) TRPs was computed for each \( M \sim N[M,(0.25 \times M)^2] \) and the mean, CV, and 95% confidence intervals derived. The percentile method was used to estimate 95% confidence intervals, where the 2.5% and 97.5% quartiles from the sorted \( \hat{F} \) vector were chosen to represent the upper and lower 95% confidence intervals respectively (Buckland 1984).

Results

A total of 3,989 \( C. \text{virginalis} \), 6,335 \( L. \text{turneri} \), 3,512 \( O. \text{argyrosoma “ssp. red,”} \) and 3,406 \( O. \text{tetrastigma} \) were measured during the assessment period, and the von Bertalanffy growth curves fitted to monthly length frequency distributions are shown in Fig. 2. The five species were relatively small with estimated \( L_\infty \) ranging between 118 and 140 mm TL (Table 1).
Length-age converted catch curves and resultant Z estimates are shown in Fig. 3. First estimates of $M$, $F$, and $q$ for *C. chrysonotus*, *C. virginalis*, *L. turneri*, *O. argyrosoma* “ssp. red,” and *O. tetrastigma*, and age-converted selectivity and maturity parameters are summarized in Table 1.

For all species, the $F$ required to attain the $F_{MAX}$ TRP was considerably higher than that required to attain $F_{0.1}$ and $F_{SB35}$ (Fig. 4, Table 2). In all species, an $F_{MAX}$ strategy resulted in severe SBR depletion and the $F_{0.1}$ TRP approximated the $F_{SB35}$ TRP (Fig. 4, Table 2).

The response of YPR and SBR indicate differences in resilience to fishing effort by the different species. From a YPR perspective, both *C. chrysonotus* and *C. virginalis* are harvested at close to optimum levels with current fishing mortalities ($F_{CUR}$) between the $F_{0.1}$ and the $F_{MAX}$ TRPs. From an SBR perspective, $F_{CUR}$ reduces the SBR of the two species to about

---

**Table 2. Summary statistics for four target reference points (TRP) ($F_{MAX}$, $F_{0.1}$, $F_{SB50}$, and $F_{SB35}$) derived from a Monte Carlo estimation procedure using 1,000 iterations for five target species in the Lake Malombe nkacha net fishery.**

<table>
<thead>
<tr>
<th>Species</th>
<th>TRP</th>
<th>$F$ (yr$^{-1}$)</th>
<th>Lower-upper 95% CIs</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. chrysonotus</em></td>
<td>$F_{MAX}$</td>
<td>2.41</td>
<td>1.11-4.41</td>
<td>34.1</td>
</tr>
<tr>
<td></td>
<td>$F_{0.1}$</td>
<td>1.14</td>
<td>0.64-1.75</td>
<td>24.5</td>
</tr>
<tr>
<td></td>
<td>$F_{SB50}$</td>
<td>0.68</td>
<td>0.43-0.97</td>
<td>20.1</td>
</tr>
<tr>
<td></td>
<td>$F_{SB35}$</td>
<td>1.13</td>
<td>0.70-1.64</td>
<td>20.6</td>
</tr>
<tr>
<td><em>C. virginalis</em></td>
<td>$F_{MAX}$</td>
<td>1.69</td>
<td>0.86-2.93</td>
<td>57.6</td>
</tr>
<tr>
<td></td>
<td>$F_{0.1}$</td>
<td>0.76</td>
<td>0.53-1.03</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td>$F_{SB50}$</td>
<td>0.54</td>
<td>0.39-0.71</td>
<td>15.4</td>
</tr>
<tr>
<td></td>
<td>$F_{SB35}$</td>
<td>0.94</td>
<td>0.65-1.29</td>
<td>17.9</td>
</tr>
<tr>
<td><em>L. turneri</em></td>
<td>$F_{MAX}$</td>
<td>0.62</td>
<td>0.53-0.71</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>$F_{0.1}$</td>
<td>0.42</td>
<td>0.38-0.47</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>$F_{SB50}$</td>
<td>0.28</td>
<td>0.27-0.30</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>$F_{SB35}$</td>
<td>0.45</td>
<td>0.42-0.48</td>
<td>3.8</td>
</tr>
<tr>
<td><em>O. argyrosoma</em> “ssp. red”</td>
<td>$F_{MAX}$</td>
<td>0.72</td>
<td>0.57-0.92</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>$F_{0.1}$</td>
<td>0.46</td>
<td>0.39-0.55</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>$F_{SB50}$</td>
<td>0.33</td>
<td>0.30-0.37</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>$F_{SB35}$</td>
<td>0.53</td>
<td>0.47-0.61</td>
<td>6.7</td>
</tr>
<tr>
<td><em>O. tetrastigma</em></td>
<td>$F_{MAX}$</td>
<td>0.68</td>
<td>0.57-0.83</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>$F_{0.1}$</td>
<td>0.46</td>
<td>0.40-0.53</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>$F_{SB50}$</td>
<td>0.32</td>
<td>0.29-0.35</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>$F_{SB35}$</td>
<td>0.50</td>
<td>0.46-0.55</td>
<td>4.6</td>
</tr>
</tbody>
</table>

*C. chrysonotus* data are from Weyl et al. 2005.
Figure 3. Length-age converted catch curves for *Copadichromis virginalis*, *Lethrinops turneri*, *Otopharynx argyrosoma* “ssp. red,” and *O. tetrastigma* sampled from the nkacha net fishery in Lake Malombe, March 2000 to April 2001. Total mortality ($Z$) estimates were obtained through regression analysis of the descending limb (closed circles) of the linearized catch curve (closed circles). The ascending limb of the catch curve and outliers (open circles) were excluded from analysis.
Figure 4. The response of *Copadichromis chrysonotus*, *C. virginalis*, *Lethrinops turneri*, *Otopharynx argyrosoma “ssp. red,”* and *O. tetrastigma* yield-per-recruit (YPR) and spawner biomass-per-recruit (SBR) to fishing mortality ($F$) in Lake Malombe, Malawi. Point estimates for the current fishing mortality rate, $F_{\text{CUR}}$, and four target reference points, $F_{\text{MAX}}$, $F_{0.1}$, $F_{\text{SB50}}$, and $F_{\text{SB35}}$ are illustrated. (*C. chrysonotus* data are from Weyl et al. 2005.)
30% which exceeds the $F_{SB35}$ TRP. *L. turneri* is severely over-exploited from both an YPR and SBR perspective with $F_{CUR}$ being more than twice as high as $F_{MAX}$ and SBR being depleted to about 6% of pristine levels. In *O. argyrosoma* “ssp. red,” $F_{CUR}$ also exceeds $F_{MAX}$ and depletes SBR to about 17% of unexploited levels. With $F_{CUR}$ well below that required to attain $F_{0.1}$ and an SBR of about 66% of pristine levels, the *O. tetrastigma* stock appears healthy.

From a TYPR perspective, $F_{MAX}$ was attained at $f$ levels of about 130 gears and $F_{0.1}$ was attained at an $f$ of 68 gears (Fig. 5). Effort levels of 130 gears, while lower than current effort, would not lead to a significant rebuilding of the SBR in the four over-exploited species. The $F_{0.1}$ strategy would maintain the SBR of *C. chrysonotus, C. virginalis,* and *O. tetrastigma* at levels above 50% of pristine, and reduce *L. turneri* and *O. argyrosoma* “ssp. red” SBR to 19% and 39% of pristine SBR, respectively.

**Discussion**

It has long been recognized that when a common gear harvests a number of species, it is impossible to manage each species at its optimum level (Beverton and Holt 1957, Anderson 1975, Pope 1979, Pikitch 1987). This is due to differences in the life histories of the different species which infers different population responses to exploitation rates as well as behavioral differences that affect the availability of each species to the gear (Murawski 1984). For instance, slow growing species with low natural mortality rates, long life-spans, late maturity, and precocial breeding habits are more likely to have relatively stable population sizes, and therefore relatively stable catch levels (Adams 1980). However, once they are overfished, it would require a relatively long period (depending on the extent of overfishing) for the stock to rebuild. Conversely, fast growing species with high rates of natural mortality, early maturity, and altricial breeding habits would support more productive fisheries, where fish can be harvested at younger ages from the population. However, these fisheries are likely to be of a “boom and bust” nature (Adams 1980), characterized by high initial stock sizes, and the potential for both growth and recruitment overfishing.

These effects are evident in the target species of the nkacha net fishery. While all five species are maternal mouth-brooders (Turner 1996), *L. turneri* had the latest maturity and lowest natural mortality rate (Table 1). Subsequently, this species appeared to be most vulnerable to the fishery with relatively low effort levels leading to asymptotic YPR and severe SBR depletion. At current effort levels, this species is over-exploited from both an YPR and SBR perspective. This over-exploitation is also evident from catch estimates for this species which show a decline in *L. turneri* catch from about 2,200 tons in 1991 to less than 900 tons in 2001 (Weyl et al. 2004).
Figure 5. The response of total yield-per-recruit (TYPR) and spawner biomass-per-recruit of *Copadichromis chrysonotus*, *C. virginalis*, *Lethrinops turneri*, *Otopharynx argyrosoma* “ssp. red,” and *O. tetrastigma* as a function of effort (expressed as the number of nkacha nets) in Lake Malombe, Malawi, Africa. Note that for TYPR the yield-per-recruit for each species considered is additive and expressed as a percentage of the maximum.
Conversely, the two *Copadichromis* species had the fastest growth rates and highest natural mortality rates. Consequently, the same effort levels that over-exploit *L. turneri* resulted in less severe YPR and SBR reduction in these two species. Despite similar catchability coefficients and SBR reductions to about 30% of pristine levels, *C. chrysonotus* catch increased from about 120 tons in 1991 to about 510 tons in 2000, while *C. virginalis* landings declined by more than 50% over the same period (Weyl et al. 2004). While this implies a higher vulnerability of *C. virginalis* to the fishery, further investigation of this effect was not possible.

The two *Otopharynx* species had similar growth and natural mortality rates, but responded differently to the fishery. Current effort levels exceed the $F_{MAX}$ TRP for *O. argyrosoma* “ssp. red” and deplete its SBR to about 17% of unexploited levels (Fig. 4). *O. tetrastigma*, on the other hand, appears far less vulnerable to the gear and the species is underexploited on a YPR basis and its SBR is estimated to exceed 66% of unexploited levels (Fig. 4). Such effects imply different vulnerabilities of the five species to the fishing gear, which may be a consequence of differences in spatial distribution or behavior. Unfortunately, little is known about the distribution of the five species within the lake. What is known is that in the adjacent Lake Malawi, *O. tetrastigma* favors well-vegetated areas, while *O. argyrosoma* “ssp. red” inhabits open areas (Konings 1995). It is therefore possible that a part of the *O. tetrastigma* population is unavailable to nkacha nets, which operate in offshore areas. It must also be considered that, due to the shallow depth (maximum depth of 5 meters) and uniform sandy/muddy bottom, most of the lake is fishable with nkacha nets and spatial exclusion would be valid only for species favoring extremely shallow areas. The lack of knowledge on distribution and behavior of the five species therefore forms a major bottleneck in understanding the dynamics of the system. But all species have the same commercial value and it is unlikely that the fishery would target any one of the five species preferentially.

The Malawi fisheries policy (Government of Malawi 1999) stresses the maximization of harvests within safe sustainable yield levels. Peak yield-per-recruit is attained if an infinite fishing mortality is applied when the biomass of a cohort is at its maximum (Pereiro 1992). However, it is evident that, at current recruitment levels, higher yields are unlikely to be attained from this fishery (Fig. 5). In mouth-brooding cichlids recruitment may be dependent on spawner-biomass and the maintenance of the spawner stock at levels where it can replace itself is vital. Taking SBR recommendations for other species (Gabriel et al. 1989, Quinn et al. 1990, Clarke 1991, Punt 1993, Mace 1994, Caddy and Mahon 1995, Booth 2004) into account, the maintenance of SBR at levels between 35% and 50% of pristine was taken as an initial management target and, ideally, the SBR of all five major target species should be maintained at levels exceeding 35% of unexploited biomass.
A visual representation of the possible management choices, presented in Fig. 6, shows the variability inherent to the $F_{SB35}$ management targets. Taking into account the variability of the $F_{SB35}$ estimates, current effort levels are likely to maintain the SBR of *C. chrysonotus*, *C. virginalis*, and *O. tetrastigma* within acceptable levels. Effort would have to be reduced to at most 88 nkacha nets if the $F_{SB35}$ TRP is to be attained for at least four species. If SBR is to be maintained at a $F_{SB35}$ TRP for all five species, effort would have to be reduced to about 40 gears. However, as the nkacha net fishery cannot target any species in isolation and the market price for all five species is identical, a strategy maximizing total yield from the fishery, without regard for individual SBR, might be favored. But in the absence of information on interspecies effects and the SBR-recruitment relationship, such a strategy cannot be recommended.

While per-recruit models are relatively data-intensive and some data from previous studies (Banda 1995, Weyl et al. 2004, Weyl et al. 2005.) were used in the analysis, all input parameters for the models could have been estimated during a one-year assessment period. But it must be realized that the per-recruit models assume that age-specific maturity, age-specific selectivity, the growth equation, natural mortality and the current rate of fishing mortality are constant, that there is no recruitment variability, and that it is possible to impose a specified fishing mortality (Punt 1993). Despite these constraints, management is an immediate rather than a future concern in African fisheries and a per-recruit approach enables the rapid provision of a number of management options for a fishery given constraints in available historic data. It must also be noted that the age estimates derived from length-frequency analysis in this study only provide a first-estimate of growth for the five species until alternative aging methods can be applied. In this regard, sectioned otoliths are currently considered the most suitable hard tissue for age and growth determination in tropical and subtropical areas (Campana 2001). For this reason, it is vital that the per-recruit approach presented in this paper is seen as a first-assessment and that efforts be made to collect data for the future application of more quantitative methodologies.

**Acknowledgments**

The senior author would like to thank Rhodes University for the award of a Rhodes University postdoctoral fellowship. The fieldwork for this study was financially and logistically supported by the Gesellshaft Für Technische Zusammenarbeit (GTZ GmbH, Eschborn), supported by the National Aquatic Resource Management Programme (NARMAP) of the Malawi Department of Fisheries. We would like to express our sincere thanks to Uwe Scholz and Peter Jarchau for their moral support during the fieldwork phase of this study. We also express our sincere thanks to the fishermen of Lake Malombe whose cooperation facilitated this study.
Figure 6. Mean and 95% confidence intervals (CIs) for the $F_{SB35}$ target reference point (TRP) expressed as a function of fishing effort for *C. padichromis chrysonotus* (CC), *C. virginalis* (CV), *Lethrinops turneri* (LT), *Otopharynx argyrosoma* “ssp. red” (OA), and *O. tetrastigma* (OT) in the nkacha net fishery of Lake Malombe, Malawi, Africa. The 95% CIs were derived from a Monte Carlo estimation procedure using 1,000 iterations for five target species in the Lake Malombe nkacha net fishery. Current fishing effort $f_{CUR}$ is illustrated.

The attendance of the senior author at the 21st Lowell Wakefield fisheries symposium was facilitated through financial assistance from Rhodes University and the Alaska Sea Grant College Program.

References


Pauly, D., and N. David, N. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequencies data. Meeresforschung 28(4):205-211.


Do Commercial Fishery CPUE Data Reflect Stock Dynamics of the Baltic Herring?

Tiit Raid, Ahto Jäärvik, Leili Järv, and Olavi Kaljuste
Estonian Marine Institute, University of Tartu, Tallinn, Estonia

Abstract
Herring are important commercial fisheries of the Baltic Sea. The structure and dynamics of Baltic herring stocks have been monitored thoroughly since international management started in the mid-1970s but management has not always been successful, particularly during periods of rapid change. Shortcomings in certain assessment data are a concern and the difficulty of obtaining appropriate biological samples from the huge and extremely heterogeneous herring stocks is a problem. Another problem is the lack of good tuning data for age-structured analyses. Presently the results of an annual International Acoustic Survey are used for tuning in extended survivors analysis (XSA). A critical shortcoming is that the acoustic survey does not cover several key areas of herring distribution. Therefore, there is a need for alternative or supplementary data to minimize effects of biased acoustic results. CPUE (catch per unit effort) in commercial trawl fishery may be a potential source of additional data. In theory, the CPUE should be proportional to the stock size so changes in CPUE, adjusted for gear efficiency, could provide, at low cost, supplementary information on stock trends. The CPUE data collected from the Estonian pelagic trawl and pound-net fishery (a small part of the larger Baltic herring fishery) show good accordance with the results of VPA estimates of separate stock units in 1990-2000s. Further exploration of commercial CPUE as a stock status index would be reasonable.

Introduction
Pelagic stocks provide the vast majority of commercial fish production of the Baltic Sea. The total landings of Baltic herring (Clupea harengus membras) and the Baltic Sprat (Sprattus sprattus balticus) exceeded
600,000 t, sometimes reaching 700,000 t in the last decade. Like herring in other parts of the world, Baltic herring form local stocks or populations (Stephenson 1991). Probably there are nine or ten different Baltic herring populations, each with different biological characteristics and stock dynamics (Ojaveer 1981, Stephenson 1991). Three herring populations, the Gulf of Riga herring, the Gulf of Finland herring, and the open-sea herring of the northeastern Baltic, inhabit the northeastern part of the Baltic Sea. The gulf herring stocks spend all year within gulf waters, while the open-sea herring occur mostly in the Baltic proper, performing spawning migrations to the spawning grounds located in the gulfs or archipelago area.

The dynamics of Baltic herring populations (stocks) have been assessed and managed internationally since the mid-1970s. In early years, the separate assessments were performed for several local stocks, mostly defined on geographical basis (ICES 2001b). Since 1990, all local herring populations in the main basin and in the Gulf of Riga and the Gulf of Finland were combined and assessed as one stock (central Baltic herring in subdivisions 25-29 and 32). At present, the Baltic herring is assessed in five assessment units (Fig. 1):

- Herring in subdivisions 25-29 and 32 excluding the Gulf of Riga.
- The Gulf of Riga herring.
- Herring in subdivision 30.
- Herring in subdivision 31.

The international management of those stocks has not always been successful, particularly in case of the central Baltic herring where biomass has decreased continuously throughout the approximately 30 year period of observations (Ojaveer 2002). Therefore the complex stock structure of the Baltic herring could be one of the reasons for difficulties in understanding and managing these stocks.

There may be other reasons; for example difficulties in data collection also can have a substantial role (Raid 2002). In general, there are only few problems with collecting of general biological data from the commercial catches, discards, and landings. The respective data sets cover all Baltic herring fisheries, areas, and seasons for past 30 years reasonably well (ICES 2003). However, the tuning data for analytical assessments are fragmentary and this has been one of the shortcomings of herring assessment for years. At present, the results of the annual International Acoustic Survey (IAS) are used in order to tune the VPA in the XSA (extended survivors analysis) in the routine herring assessments by the ICES Baltic Fisheries Assessment Working Group. The IAS, initially directed on sprat
Figure 1. ICES subdivisions in the Baltic Sea. The light-shaded area shows the coverage by the International Acoustic Survey. The overlapping areas between different national surveys are dark-shaded.
stock, does not allow full coverage of the herring distribution area. Areas do not include a substantial part of the distribution area of the central Baltic herring, most of subdivision 29, the Gulf of Finland (subdivision 32), and the Gulf of Riga (part of subdivision 28) (Fig. 1). Consequently one herring population, the open-sea herring in the northeastern Baltic, has only partial coverage, while two populations (the Gulf of Finland herring and the Gulf of Riga herring), are not covered at all (Fig. 1). Therefore although there are considerable basic biological data on catch composition, there are essential gaps in the key data sets used for tuning of the assessments. In this sense, the Baltic herring fishery is data-limited. The limitation is not with the “amount” of data (which is voluminous) but rather with the limited geographic coverage. The main reason for poor coverage is high cost of acoustic surveys. In that respect, there is a clear need for alternative or supplementary data in order to fill the gaps in the existing tuning data and thus minimize the effect of inadequate acoustic estimates on assessment results.

CPUE data (catch per unit effort) from the commercial fishery may be a source of additional data for the herring stocks in the northeastern Baltic. The CPUE of the Finnish trawl fishery is routinely used in the assessments of herring stock in the Bothnian Sea (subdivision 30) and in the Bothnian Bay (subdivision 31; ICES 2001a, 2002, 2003). In theory, the CPUE of the commercial fishery should be proportional to the stock size (e.g., Gulland 1964), and consequent change in CPUE should track, at given gear efficiency, the dynamics in fish stocks. Such additional CPUE information could provide a low cost alternative or supplement to acoustic estimates used for tuning. The aim of the present paper was to explore if CPUE in commercial fishery could be a possible additional data source on stock status of the Baltic herring in the northeastern Baltic.

**Materials and methods**

Biological data for this study was collected from the Estonian commercial trawl and pound-net fishery during the routine data collection for assessment purposes. A sample of 100 fish was collected and analyzed every 10 days for all fishing seasons and fishing grounds in Estonian waters of subdivisions 28, 29, and 32 and in the Gulf of Riga (Fig. 1). Most of the data used in present work are from the period 1995-2002.

The pelagic trawl fishery takes about 95% of Estonian herring catches in most fishing areas, except the Gulf of Riga. In general, the fishing fleet and the gears have not been changed in the recent decade. Usually this fishery exploits the densest herring concentrations. Since no discarding or sorting occurs in the Estonian trawl fishery, catch data can be used to characterize the main trends in herring stocks. Approximately 5% of herring catches are taken by the pound-net fishery (but this can be up to 50% in the Gulf of Riga). The pound nets are big fixed trap-nets with open tops.
They are used in shallow coastal areas during the spawning season (second quarter of the year, April-June) to catch pre-spawning herring. Therefore only the spawning stock is represented in the pound-net fishery.

The CPUE data in trawl fishery (kilograms per hour) were compiled on a quarterly basis according to logbooks of fishing vessels. The respective data for herring pound-nets, in metric tons per check (the fishermen check the pound-net catches 3-5 times a week), were compiled on a monthly basis. The CPUE of the pound-net fishery was explored in the Gulf of Riga only (the main basin for that fishery). The VPA estimates of stock components were derived either from the reports of the ICES Baltic Fisheries Assessment Working Group (Gulf of Riga herring, central Baltic herring) or produced ad hoc, using input parameters (catch numbers, mean weights, and natural mortalities, etc.) available for the respective subdivisions in the same sources (ICES 2001a,b, 2002, 2003).

**Results and discussion**

The spawning stock biomass (SSB) of the complex central Baltic herring, including several local populations, has decreased by approximately 70% since 1974 (ICES 2003). Landings decreased from 300,000 t to below 200,000 t. Fishing mortality estimates have been in the range of 0.2-0.3; they were above the defined $F_{PA}$ ($F = 0.19$, ICES 2003) until 1994, increasing thereafter up to almost 0.5 in 2000 (Fig. 2). The separate assessments of different stocks and results of hydroacoustic surveys of different subunits (herring in subdivisions 25-27, subdivisions 28, 29, and 32 and the Gulf of Riga herring) have revealed variation within pooled assessment units. For instance, fishing mortality of the Gulf of Riga herring has decreased but there was a sharp increase in mortality in the northeastern Baltic (subdivisions 28, 29, and 32). At the same time stock abundance and biomass of the Gulf of Riga herring increased to record high levels in early 2000s also allowing higher catches (ICES 2002, 2003; Fig. 3).

The average CPUE data from the Estonian herring pelagic trawl fishery, recorded since 1995, is presented in the Table 1. The comparison of the dynamics of mean CPUE and the SSB estimates in respective assessment units show similar trends (Figs. 2 and 3). However, the degree of coherence varies—the closest associations were between subdivision 32 (Gulf of Finland) and for the combined unit subdivisions 28, 29, and 32 (Figs. 4 and 5). The fit between the mean CPUE and the SSB estimate in subdivision 29 was the poorest. The results indicate that for the Gulf of Riga and subdivision 32, the trends in CPUE of the commercial trawl fishery are quite similar to the SSB trends. Deficiencies in the quality of VPA estimates due to the complexity of herring in the area might be the reason for the poor fit observed for subdivision 29. The area is an essential feeding area and a migratory area for several local herring stocks that mixing there, particularly during the late autumn and wintering period.
Figure 2. Upper panel shows dynamics of SSB, landings, and fishing mortality in the central Baltic herring stock in the 1980-2000s. Lower panel demonstrates SSB and CPUE trends in subdivisions 28, 29, and 32 (ICES 2002, 2003).
Table 1. Mean catch per hour in the herring pelagic trawl fishery (kg).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf of Riga</td>
<td>1,009</td>
<td>707</td>
<td>637</td>
<td>667</td>
<td>700</td>
<td>697</td>
<td>804</td>
<td>820</td>
</tr>
<tr>
<td>29</td>
<td>1,035</td>
<td>892</td>
<td>802</td>
<td>671</td>
<td>993</td>
<td>836</td>
<td>771</td>
<td>419</td>
</tr>
<tr>
<td>32-1</td>
<td>715</td>
<td>572</td>
<td>373</td>
<td>271</td>
<td>441</td>
<td>394</td>
<td>300</td>
<td>184</td>
</tr>
<tr>
<td>32-2</td>
<td>629</td>
<td>498</td>
<td>410</td>
<td>241</td>
<td>417</td>
<td>370</td>
<td>272</td>
<td>538</td>
</tr>
<tr>
<td>32</td>
<td>672</td>
<td>559</td>
<td>377</td>
<td>267</td>
<td>438</td>
<td>392</td>
<td>297</td>
<td>264</td>
</tr>
<tr>
<td>Mean</td>
<td>868</td>
<td>697</td>
<td>498</td>
<td>348</td>
<td>589</td>
<td>559</td>
<td>483</td>
<td>482</td>
</tr>
</tbody>
</table>

CPUE data may not be normally distributed so several authors advocate using median CPUE instead of the arithmetic mean (Hamley and Howley 1985, Reed 1986, Järvik 1989). The comparison of CPUE data from the herring pound-net fishery in the Gulf of Riga has a poor fit, regardless whether the simple arithmetic mean or median of CPUE was used (Table 2, Fig. 6). The absence of correlation could be explained by the pattern of the pound-net fishery, concentrated mostly on major spawning grounds. The actual number of fish in pre-spawning schools at each particular spawning ground depends on environmental factors, such as wind direction and temperature, and therefore can be rather variable regardless of spawning stock size. However, the total annual pound-net catches were found to be a quite good indicator of Gulf of Riga herring SSB (Järvik and Raid 2000).

The commercial CPUE data used in the present study cover only a small fraction of the conventional assessment units. To determine if CPUE data obtained in the limited part of the assessment unit is representative as an indicator of stock developments, the arithmetic mean of commercial CPUE of Gulf of Finland herring in 1995-2002 was compared with the available set of respective SSB estimates for different herring stock units and their combinations. The results (Fig. 7) indicate that the commercial CPUE data, obtained from the remote northern part of the distribution area of the central Baltic herring stock complex provides a satisfactory description of the main SSB trends in both combined stocks (subdivisions 25-29 and 32, and subdivisions 28, 29, 32).

A much looser relationship was found between the mean CPUE in the Gulf of Finland herring fishery and the SSB estimate for the neighboring subdivision 29, which probably indicates again at the uncertainties in the SSB estimates for that subdivision.
Figure 3. Trends in SSB and CPUE of herring in the Gulf of Riga (ICES 2003), and in subdivisions 29 and 32.
Figure 4. Scatter plots of CPUE and SSB estimates in subdivision 32, subdivision 29, and Gulf of Riga assessment units (1995-2002).

Figure 5. Scatter plots of CPUE and SSB estimates in subdivisions 28, 29, and 32, and subdivisions 25-29 and 32 excluding the Gulf of Riga assessment units (1995-2002).
No relationship was revealed between the Gulf of Finland herring CPUE and SSB estimates of the Gulf of Riga herring. The absence of any correlation in that case is not surprising, because of the fundamentally different pattern in stock developments in the Gulf of Finland and the Gulf of Riga (Fig. 3).

Although there is a relatively good temporal coherence between commercial CPUE and herring SSB estimates, the use of commercial CPUE as an indicator for stock trend should be taken with caution because the CPUE estimation depends on the distribution pattern of fish, and sometimes also on the behavior of the fishing vessel captain. For example,
during daylight herring CPUE tends to be higher because herring are in dense shoals, whereas during the dark of night, herring are less dense and distributed more evenly. Hence, the source of CPUE data should be examined critically prior to including data in the assessment input.

**Conclusions**

Baltic herring fisheries are not limited by the *amount* of available data, but perhaps by the suitability of existing data. In particular, many of the Baltic herring fisheries do not have sufficient data to accurately describe relative trends in stock size. The tentative exercise of comparison of general trends in herring SSB estimates in different assessment units in the northeastern Baltic, to commercial CPUE data from the pelagic trawl fishery, show rather good fits in the case of units with less uncertain assessments. The CPUE data covering only a part of the remote section of the central Baltic herring stock follow the main trends in SSB estimates of the combined stock rather well, suggesting that further exploration of the possibilities of implementation of commercial CPUE data into assessment procedures of Baltic herring would be reasonable.
References


