A METHOD OF ANALYZING CATCHES AND ABUNDANCE INDICES FROM A FISHERY

M. L. Parrack
NMFS Southeast Fisheries Center
75 Virginia Beach Drive, Miami, FL 33149, U.S.A.

SUMMARY

A procedure of calibrating virtual population analysis estimates of stock sizes developed from age specific catches to observed indices of abundance is described and demonstrated in application. The procedure allows separate appraisal of each individual index of abundance and develops least squares estimates of stock sizes and fishing mortalities. The method is applicable if one or more abundance indices exist that were not sampled for age composition and that may overlap in the years or ages encompassed.

REFERENCES

Se describe y demuestra prácticamente un procedimiento para calibrar estimaciones de tamaños de población, hechas por el método de VPA, desarrolladas a partir de capturas por edad hasta obtener índices observados de abundancia. Ese procedimiento permite una evaluación separada de cada uno de los índices de abundancia y desarrollar estimaciones de los tamaños de población y mortalidades por pesca por mínimos cuadrados. El método es aplicable si hay uno o más índices de abundancia que no han sido muestrados por composición demográfica, y que pudieran solaparse en los años y edades considerados.

RESUME

Une méthode de calibrage aux indices d'abondance observés des estimations de la taille du stock par l'analyse des populations virtuelles, élaborées à partir de prises spécifiques de l'âge, est décrite et son application illustrée. La méthode permet une évaluation individuelle de chaque indice d'abondance, et élaborer des estimations par les moindres carrés de la taille du stock et de la mortalité par pêche. Elle est applicable s'il existe un ou plusieurs indices d'abondance qui n'ont pas été échantillonnés quant à la composition démographique, et qui pourraient se recouper pour les années ou âges compris.
INTRODUCTION

It is well known that although virtual population analysis (VPA) calculations of stock size (see Dershavin 1922, Fry 1946, Gulland 1955, Murphy 1965) are indicative of relative abundance trends, they are often very biased estimates of abundance magnitudes. This calculation sequence (See Appendix 1) transforms an age-year matrix of catch into two 2-like matrices, one of abundance and one of fishing mortality rates, given two assumptions: 1) an age-year natural mortality rate matrix and 2) an age-specific fishing mortality rate vector with one element for each cohort in the catch matrix. It is recognized that unless the catch matrix is derived from actual direct observations of age, results may be very biased (Rosenberry and Reddington 1985, Sainsbury 1988). Likewise, unrealistic assumptions as to natural mortalities can result in biased results (Jones 1981). It is commonly known, however, that the assumption regarding the fishing mortality vector is critical. Rather small changes in the levels of its elements can impact VPA estimates of stock magnitude substantially. Results are so sensitive to these levels that unless very accurate estimates of F exist for each cohort, often calculations of abundance magnitudes are too biased to be of use.

In recent years several attempts have been made to overcome this limitation by combining VPA techniques with catch-per-unit effort (CPUE) indices or other observations indices of stock abundance (Gray 1977, Palombo 1980, Doubleday 1981, Anon. 1983). The underlying concept of these techniques is that of the infinite number of possible population abundance and corresponding fishing mortality trajectories defined by a catch matrix, the correct one is that which matches an accurate abundance index. The rationale involved in that the catch matrix is the single most complete, comprehensive, and accurate information that exists on the trajectory of the resource and the exploitation of it through time; it is the "ground truth" data. The working assumption is that if both the temporal trend (i.e., trajectory) in the observed abundance index (CPUE) and in a particular VPA stock abundance matrix are the same, then the VPA trajectory is reasonably correct. Although visual graphic methods have been used in the past, a mathematical technique that can be used to actually make the selection on an objective quantitative bases is least squares. Most investigations thus far have dealt with situations where age-specific abundance indices were available (See ICES 1983 for a synopsis). This manuscript is concerned with the common situation where the age structure of the annual CPUE (or other index of abundance) was not regularly sampled so age-specific CPUE is unavailable. CPUE is only one of many kinds of abundance indices, however, since it is the most common one, the term "CPUE" is used interchangeably with "abundance index" in this manuscript.

This study, thus, develops a least squares system for calibrating VPA estimates of stock size and fishing mortality trajectories to observed indices of abundance in situations where the age structure of the indices is unsampled.

METHOD

Consider the event where, in addition to an accurate year by age catch matrix there exists one or more CPUE or other abundance index sets that were not regularly sampled for age composition. Age specific indices of abundance therefore are not available, rather general knowledge as to the sizes and thus ages
most likely encompassed in each index set is all that is available. A general
procedure, then, to calibrate (adjust) VPA abundance calculations to the abun-
dance indices is to minimize the squared difference between the observed
abundance and that calculated by VPA. Symbolically let \( k \) index the \( K \) different
indices of abundance, \( j \) index age, and \( i \) index years. The least squares proce-
dure is to minimize:

\[
SS = \sum_{k=1}^{K} \sum_{i=I(k)}^{J(k)} \left( X(k, i) - q(k) \left( \sum_{j=J(k)}^{J2(k)} N(i, j) \right)^{b(k)} \right)^2
\]

with respect to \( FF, M, a(k) \) and \( b(k) \) where \( k = 1 \) to \( K \). Here \( I1(k) \) and \( I2(k) \) are
the first and last year where abundance index set \( k \) is available, \( X(k, i) \) is the
observed magnitude of abundance set \( k \) at the beginning of year \( i \), \( J1(k) \) and
\( J2(k) \) are the youngest and oldest age represented in abundance index set \( k \), \( q(k) \)
and \( b(k) \) are an equation parameters unique to abundance index set \( k \), and \( N(i, j) \)
is the VPA calculation of stock size at the beginning of year \( i \) for age \( j \).

In cases where the abundance index is in terms of weight such as yield-per-
unit-effort, the right hand term of the sum of squares expression becomes

\[
q(k) \left( \sum_{j=J(k)}^{J2(k)} w(j) \right)^b(k)
\]

where \( w(j) \) refers to the average weight at age \( j \) on the first of the year.

The left hand side of the sum of squares (SS) expression is thus the observed
abundance index and the right hand side the index predicted by the model.

Recall that the magnitudes of the \( N(i, j) \) are a function of catches, natural
mortality and the vector of the age-specific fishing mortality rates that apply
at the starting point of VPA calculations for each cohort. The catch matrix is
considered to be accurately known. Natural mortality is assumed constant over
ages and years to facilitate explanation. The vector of \( F(i) \)'s are determined in
two ways. Let \( Y \) be a single designated year where a defined fishing pattern is
imposed the last year of catch data, \( FF = \) the maximum age specific fishing
 mortalities in year \( Y \), and \( F(j) = \) a fishing pattern constant for age \( j \)
in year \( Y \) such that:

\[
F(Y, j) = FF \cdot F(j).
\]

This defines an age-specific \( F \) for each cohort caught in year \( Y \) as required to
generate the \( N(i, j) \) for those cohorts. For cohorts represented in the catch
and not caught in year \( Y \), the average \( F \) over cohorts where calculations are pre-
sent from the above method for the year that the cohort was last caught may be
used.

The left hand term (in the sums of squares expression) is thus the observed
abundance level and the right hand term the estimated level from the model.
Least squares solutions for \( M, FF, q(k) \) and \( b(k) \) may be found thus pro-
viding estimates of the \( F(i, j) \) and \( N(i, j) \). The (independent) variables
required to develop the least squares solutions are the matrix of catch, the
P(j) and the X(k,l). The solution for FF and M is iterative due to the non-linear nature of the VAF equations, and solutions for the q(k) and b(k) are analytical (unless an additive error is assumed and b(k) is not assumed to be equal to one).

The parameters q(k) and b(k) define the form of the relation between the observed abundance index and stock size projected by the model. For instance, if b(k) is fixed at unity for abundance index CPUK set k, then the relation between observed and estimated CPUK is a straight line through the origin. If that parameter is not fixed, then the relation between the two may take on any of the general shapes below as determined by least squares (below). Of these, some are clearly inappropriate, some meaningless, and some quite useful. Cook and Reddington (1985) point out that if the logistics of fishing and the stochastic nature of fishery is considered, the parameter b can be greater than zero and less than unity. This implies that rather large decreases in stock abundance will not cause a pronounced decrease in CPUK until stock levels are quite low. The usual model, C = q'M, implies that b = 1 so that stock size and CPUK are proportional. If b = 0 then the CPUK index and abundance are independent. Likewise, values of b less than zero indicate that the CPUK set does not index abundance in any useful way. If b > 1 then the CPUK set is very sensitive to stock size changes, a very desirable trait.

For each cohort in the catch matrix, the starting point calculation must be determined that is applicable to the particular cohort. The parameter vectors (N and P) for the cohort are determined by the starting point calculation. In this regard, a cohort may fall into one of three categories or "cases". Parameter vectors are calculated either directly by least squares or indirectly from least squares estimates made previously. The first two cases are used for cohorts encompassed in one or more of the indices of abundance; these determine the magnitude of the sum of squares. In the below description of each case, symbols are as before except that l is used to index cohorts so that l = 1-j, i.e. the cohort index is the year minus the age.

Case 1. Cohort l occurs in the X (i.e. abundance indices), was caught in year Y, and P(Y-l) is set greater than zero. The starting point calculation is:

$$F(Y, Y-l) = F(Y-l)$$

Case 2. Cohort l occurs in the X but was not caught in year Y, (or equivalently if it was caught in year Y, then P(Y-l) was set equal to zero). Averages are used to determine the starting point calculation. The starting point calculation is in year 1-J(l) at age J(l) where J(l) refers to the oldest age in the catch of cohort l. A simple average F is employed as follows:

$$F(1+J(1), J(1)) = \frac{1}{n} \sum_{j=m}^{m+2} F(1-j, j)/n$$

Although it is probably less desirable (see Anon 1985), an average weighted by stock-size may also be used. This is the average of all F's over a defined
bracket of ages, \( m_1 \) to \( m_2 \), during the year the cohort was last caught. If case 1 calculations are made first and then calculations for the remaining cohorts are sequentially made in decreasing order of the cohort index, a sufficient number of F's will exist in each year to compute these averages; in fact it is usually calculated over all fully recruited ages except the oldest one.

Case 3. Cohort 1 does not occur in any of the abundance indices and was last caught at age \( m_1 \) or older. Such cohorts are outside the minimization system so they do not affect the sum of squares but their parameter vectors can be calculated by averaging F's estimated by least squares. Calculation is that described for Case 2.

The sequence of events then, is to determine the calculation method for each cohort, determine the minimum sum of squares and resulting population parameters over Case 1, and 2 cohorts, and then calculate the trajectories for case 3 cohorts if they are of interest. It is interesting that any number of CPUE sets may be used simultaneously and that any number of these may overlap in regards to the span of years or of ages encompassed. If the system becomes undetermined inadvertently, multi-minima will become obvious in the results. This, of course, is easily avoided by insuring that the number of parameters being estimated does not exceed the number of differences entering the sum of squares.

APPLICATION

When applying this methodology, the procedure is to first fit each available CPUE set to the catch matrix separately in isolation from all other sets, and closely inspect the results. Each CPUE set can be rejected or accepted as a useful indicator of stock abundance in this manner by appraising the sum of squares surface, residual plot, and results of regression.

Sum of Squares response Surface. This is visually appraised from a plot of the sum of squares surface on FF (vertical axis) and the loss rate due to other causes (i.e. horizontal axis). The resulting isopleth should exhibit a single definite and sharp minima over a reasonable range of FF and M. If the minima occurs outside of this bound then the catch matrix or CPUE set is in question.

Residuals. A scatter plot of the observed abundance (CPUE) minus that predicted by the model must be examined in the usual way. Improper model selection, temporal changes in the CPUE and stocksize relation, excessive "noise" in the CPUE and other problems will be reflected in such plots. It is important to note that if M is fixed, rather than estimated by least squares, inaccurate levels of M will result in a temporal trend in residuals.

Residual Mean Square. This statistic (RMS) is the best criteria for choosing between the conventional CPUE-stocksize relation (CPUE \( \propto sM \)) or the power function (CPUE \( \propto s^b \)). If the latter is used and the estimate of \( b \) is near one then the results are nearly the same, however, this does not often occur so the choice must be made as to whether to estimate the additional parameter (b) or restrict it to unity (conventional model). If the RMS is not
decreased appreciably by estimating the additional parameter, then there is no justification for doing so.

Regression Results. Often, a simple scatter plot of estimated stocksize on observed CPUE is most informative when considered with the parameters of the model. In the case where the conventional model is used (b=1), points will hopefully appear clustered about a straight line through the origin. If the power model is imposed (b estimated) the plot must be considered in conjunction with the estimate of b. As shown in the method section, the estimate of the parameter b reveals a great deal about the CPUE set:

1) If b < 0.0 the CPUE set is useless.
2) If b > 1.0 the CPUE is a sensitive index of stock abundance.
3) If b = 1 the CPUE is proportional to abundance with sensitivity determined by the remaining parameter (q).
4) If 0.0 < b < 1.0 the CPUE is insensitive to abundance unless abundance is at very low levels.

Another very informative statistic of the regression is the significance level for the hypothesis that the correlation (r) between observed CPUE and estimated stock abundance is greater than zero, i.e. Pr [ r > 0.0]. This statistic provides a very powerful tool for accepting or rejecting a particular CPUE set.

In order to test the CPUE sets (as outlined above) the P (vector) must be established. The fishing pattern year (Y) is best set equal to the last year in the CPUE set. An identity (unit) vector may be used for P in these trials by defining the CPUE set to pertain only to ages where the regression parameters (q and b) may be assumed reasonably constant over those ages (such as ages fully recruited into the CPUE set in question).

After the list of useful CPUE sets is established, further investigation can occur. Here, all CPUE sets judged indicative of stock abundance may be entered into the estimation procedure for further analysis with the final goal of finding population trajectories (age-year specific P's and N's) that explain observations of catch and CPUE. Since this is basically a VPA system, starting point calculations should be made at the last year of catch (see Jones 1981, Pope 1972, Hietang 1977) to reduce the error or estimation and eliminate forward projections of stocksize smaller than observed catches. The variable P is therefore the last year of catch; hopefully that is also the last year of CPUE but that is not necessary. Next, the P vector must be considered. The conventional theory that fish recruit gradually as they age until they become fully recruited is based upon catch curve analysis of numerous fisheries, notably the North Sea plaice trawl fishery (Beverton and Holt 1957) although numerous other examples exist. If this assumption of the P(j) increasing with age until a maximum is reached holds, then establishing a reasonable P vector is simplified, however, it is not necessary to enforce this idea. The vector may be established by independent estimation, by "tuning", or a combination of two. This is accomplished by first using a unit vector, then changing the vector repeatedly and observing the magnitude of the resulting RMS (or sum of squares) until the result is judged satisfactory.
AN EXAMPLE

The data of Martin and Fry (1973) presented by Pololkein (1980) is used to demonstrate a simple application. These data are the catch at age for the Lake Opeongo Lake trout fishery for 1936-57. The natural mortality rate was estimated by Fry from mark-recapture data to be 0.385. The catch table (below) indicates recruitment was complete by age eight. Catches of fish older than age 12 were not employed as determined by Pololkein. The ratio of total annual catch to hundreds of boat hours (CPUE) were used as an index of abundance.

<table>
<thead>
<tr>
<th>Year</th>
<th>Effort</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1936</td>
<td>22</td>
<td>128</td>
<td>233</td>
<td>474</td>
<td>655</td>
<td>778</td>
<td>200</td>
<td>118</td>
<td>57</td>
</tr>
<tr>
<td>1937</td>
<td>22</td>
<td>34</td>
<td>198</td>
<td>500</td>
<td>1,025</td>
<td>556</td>
<td>176</td>
<td>38</td>
<td>4</td>
</tr>
<tr>
<td>1938</td>
<td>17</td>
<td>127</td>
<td>275</td>
<td>420</td>
<td>439</td>
<td>295</td>
<td>90</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1939</td>
<td>13</td>
<td>116</td>
<td>221</td>
<td>321</td>
<td>393</td>
<td>223</td>
<td>90</td>
<td>47</td>
<td>24</td>
</tr>
<tr>
<td>1940</td>
<td>11</td>
<td>82</td>
<td>234</td>
<td>434</td>
<td>354</td>
<td>250</td>
<td>46</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>1941</td>
<td>11</td>
<td>144</td>
<td>234</td>
<td>235</td>
<td>200</td>
<td>104</td>
<td>22</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>1942</td>
<td>9</td>
<td>46</td>
<td>117</td>
<td>217</td>
<td>101</td>
<td>93</td>
<td>28</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>1943</td>
<td>7</td>
<td>42</td>
<td>121</td>
<td>272</td>
<td>311</td>
<td>133</td>
<td>42</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>1944</td>
<td>9</td>
<td>38</td>
<td>114</td>
<td>197</td>
<td>202</td>
<td>198</td>
<td>93</td>
<td>44</td>
<td>31</td>
</tr>
<tr>
<td>1945</td>
<td>13</td>
<td>32</td>
<td>69</td>
<td>170</td>
<td>352</td>
<td>273</td>
<td>159</td>
<td>84</td>
<td>37</td>
</tr>
<tr>
<td>1946</td>
<td>17</td>
<td>78</td>
<td>116</td>
<td>240</td>
<td>345</td>
<td>217</td>
<td>92</td>
<td>47</td>
<td>26</td>
</tr>
<tr>
<td>1947</td>
<td>12</td>
<td>55</td>
<td>85</td>
<td>217</td>
<td>221</td>
<td>153</td>
<td>76</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>1948</td>
<td>15</td>
<td>47</td>
<td>122</td>
<td>311</td>
<td>315</td>
<td>303</td>
<td>8</td>
<td>47</td>
<td>32</td>
</tr>
<tr>
<td>1949</td>
<td>17</td>
<td>131</td>
<td>264</td>
<td>352</td>
<td>350</td>
<td>109</td>
<td>51</td>
<td>66</td>
<td>22</td>
</tr>
<tr>
<td>1950</td>
<td>25</td>
<td>179</td>
<td>286</td>
<td>429</td>
<td>297</td>
<td>184</td>
<td>68</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>1951</td>
<td>24</td>
<td>72</td>
<td>193</td>
<td>348</td>
<td>401</td>
<td>261</td>
<td>117</td>
<td>45</td>
<td>38</td>
</tr>
<tr>
<td>1952</td>
<td>16</td>
<td>53</td>
<td>99</td>
<td>300</td>
<td>208</td>
<td>194</td>
<td>81</td>
<td>50</td>
<td>15</td>
</tr>
<tr>
<td>1953</td>
<td>19</td>
<td>45</td>
<td>83</td>
<td>124</td>
<td>244</td>
<td>293</td>
<td>150</td>
<td>53</td>
<td>38</td>
</tr>
<tr>
<td>1954</td>
<td>20</td>
<td>48</td>
<td>128</td>
<td>300</td>
<td>373</td>
<td>359</td>
<td>187</td>
<td>73</td>
<td>29</td>
</tr>
<tr>
<td>1955</td>
<td>22</td>
<td>134</td>
<td>194</td>
<td>273</td>
<td>277</td>
<td>178</td>
<td>107</td>
<td>43</td>
<td>8</td>
</tr>
<tr>
<td>1956</td>
<td>32</td>
<td>149</td>
<td>239</td>
<td>217</td>
<td>232</td>
<td>121</td>
<td>101</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>1957</td>
<td>33</td>
<td>188</td>
<td>370</td>
<td>431</td>
<td>340</td>
<td>238</td>
<td>127</td>
<td>51</td>
<td>0</td>
</tr>
</tbody>
</table>

First, the CPUE was appraised. Age eight fish were fully recruited hence the CPUE was assumed to index ages 8+ with equal ability (i.e. J2(1) = 8 and J2(1) = 12; P(8) through P(12) = 1.0; Y = 1957). The regression of estimated stocksize on CPUE was first assumed linear through the origin, i.e. b = 1 thus CPUE = aN. The resulting sum of squares (SS) surface was continuous with a single definite minima at b = 0.53 and SS = 0.59530 (in 1957) and the correlation between estimated abundance and CPUE high (0.88, Fr [0.50] > .00001). Although an unevenness is indicated in the plot, the distribution of residuals is acceptable. (See figure 1). Reanalysis with the exponent parameter b estimated by least squares rather than set equal to one (CPUE = aN) changed the RMS very little. (The least squares solution for a was the same and the solution for b was 1.1, very near unity). Therefore, there was no justification for estimating the additional parameter in this case hence the conventional CPUE stocksize model (CPUE = aN) was appropriate.

These results are very similar to those of Pololkein's (1980) cohort analysis methods for the same assumptions (below). These calculations assumed the conventional CPUE stocksize model and considered ages 8-12 only.

<table>
<thead>
<tr>
<th>Year</th>
<th>F(1957)</th>
<th>M</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>1957</td>
<td>0.5820</td>
<td>0.5500</td>
<td>0.00017</td>
</tr>
</tbody>
</table>

The investigation of the CPUE thus showed an acceptable sum of squares surface and distribution of residuals, a very high probability that CPUE and estimated stocksize are positively correlated, and that the conventional CPUE...
stocksize model was best. The CPUE was therefore judged to be an acceptable measure of abundance hence the investigation was continued to established least squares estimates of annual stocksize and fishing mortalities. Since age 5-7 fish were but partially recruited, the CPUE stocksize relation was suspected to be different than for older fish. This idea was accomplished by redefining the same CPUE set as pertaining to ages 5-7, (i.e. $J_1(k) = 5$, $J_2(k) = 7$) thus, two CPUE sets entered the system; ages 5-7, 1936-57 and age 8+, 1936-57. Then an acceptable pattern developed by "tuning". For purpose of demonstration, only ten P vectors were tried during tuning with the below results. These results indicate that although the least squares solutions for full P in 1957 and M are 

<table>
<thead>
<tr>
<th>Trial</th>
<th>$\hat{5}$</th>
<th>$\hat{6}$</th>
<th>$\hat{7}$</th>
<th>$\hat{8}$</th>
<th>$\hat{9}$</th>
<th>$\hat{10}$</th>
<th>SS</th>
<th>M</th>
<th>$F(1957)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.0775</td>
<td>.53</td>
<td>.4990</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.03</td>
<td>0.10</td>
<td>0.40</td>
<td>0.70</td>
<td>0.95</td>
<td>1.00</td>
<td>.8791</td>
<td>.52</td>
<td>.9137</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
<td>0.10</td>
<td>0.50</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>.8166</td>
<td>.52</td>
<td>.7967</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
<td>0.10</td>
<td>0.50</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.9011</td>
<td>.52</td>
<td>.7880</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>0.10</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.0301</td>
<td>.50</td>
<td>1.8874</td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
<td>0.25</td>
<td>0.65</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.8608</td>
<td>.53</td>
<td>.5901</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>0.25</td>
<td>0.75</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.8669</td>
<td>.53</td>
<td>.5342</td>
</tr>
<tr>
<td>8</td>
<td>0.10</td>
<td>0.30</td>
<td>0.60</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.8973</td>
<td>.53</td>
<td>.5059</td>
</tr>
<tr>
<td>9</td>
<td>0.10</td>
<td>0.25</td>
<td>0.65</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.8955</td>
<td>.53</td>
<td>.5125</td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>0.30</td>
<td>0.65</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.8619</td>
<td>.53</td>
<td>.5277</td>
</tr>
</tbody>
</table>

somewhat insensitive to the 1957 fishing pattern, "tuning" the fishing pattern reduced the sum of squares significantly, 20 percent. The sum of squares surface exhibits a single definite minima. The residuals for both CPUE sets (i.e., for ages 8+ and for 5-7) are acceptably dispersed (Figure 2) and the correlation between estimated stocksize and CPUE high in both cases ($F_{(p > 0)} > .9999$). The corresponding calculations of fishing mortalities and stocksizes therefore seem a reasonable explanation for these observations.

DISCUSSION

This procedure is a useful investigative tool as shown in the example above. Likewise, it has been used elsewhere with the same favorable results. Its most useful quality is that it will not allow the investigator to employ an unsupported assumption unknowingly; statistical results clearly indicate problem areas. Areas which often are exposed include abundance indices that are too "noisy" to be useful or that do not reflect stock trajectories that can be explained from the known catch, unsupported catchability assumptions, questionable estimates of M, questionable estimates of exploitation levels, unrealistic partial P (fishing pattern) vectors, etc. A serious limitation of the method is that it does not, by itself, produce the "answer", rather it does verify with reasonable certainty if a particular hypothesized solution is supported by the data. However, at least in some cases, it can provide a criteria for choosing a particular solution as the best of a defined set.

As demonstrated, the method produced the same results for the Lake Opeongo lake trout data as the method of Bolker and the same assumptions. This method, however, does not require fishing effort nor does it require age specific CPUE as most others do (see Anes, 1953). This least squares technique is
applicable in cases where several dissimilar abundance indices of unknown annual age composition exist that may overlap in the years and/or ages encompassed. In addition, this technique provides a separate appraisal of each individual index of abundance, in isolation of all others, including criteria for accepting or rejecting the index as an accurate monitor of stock size. This appraisal provides an extremely powerful analysis tool.

Although not demonstrated, the technique does not require $M$ nor CPUE and stock size relations to be constant over years and/or ages. A single CPUE series may be redefined as applicable to different ages such as one for fully recruited fish and one for pre-recruits. In such a case two different CPUE and stock size relations would be estimated, one for pre-recruits and one for full recruits. In addition, if the residual analysis or the regression results indicate that the relation between CPUE and stock size changed or became more variable from one time period to another, the CPUE series may be divided to accommodate the problem. In addition, $M$ may be defined differently between time periods or an age dependent model may be established and its parameters estimated by least squares. If these adjustments are attempted, however, caution should be taken to carefully compare the resulting RMS to avoid over parameterisation.

This manuscript did not encompass the results of a study of estimator performance. The sensitivity of the method to stochastic attributes of the data is therefore not yet demonstrated. It is beyond doubt that such a study would not only be informative but is acutely needed.

[FORTRAN 77 source is available upon request. This software was, however, written for a virtual memory machine with large word size (48 bytes) hence, modifications may be required to install it on a smaller machine. The methodology itself is rather simple so that computerization requires aptly exploiting the capabilities of the available machine rather than programming sophisticated mathematical techniques.]
LITERATURE CITED


Appendix: VPA Calculations

If it is assumed that the rate of change in numbers of a year class is proportional to abundance (i.e. \( \frac{dN_t}{dt} = N_t \alpha \)) then the reduction in numbers \((N)\) due to catches \((C)\) and all other unobserved causes is constant through time \((t)\) with the two sources of loss \((i.e. \text{observed catch and all other causes)}\) additive \((S=Nt)\). This assumption implies:

\[ \frac{dN_t}{dt} = \alpha N_t \]

\[ N_{t+1} = N_t e^{-\alpha t} \]

Thus

\[ N_t - N_{t+1} = \frac{N_t}{e^{\alpha t}} \]

and also that

\[ C_t = F_t N_t A_t \]

where \( A_t = 1 - e^{-\alpha t} \)

Substituting (2) in (3) gives

\[ C_t = F_t N_t A_t \left( 1 - e^{-\alpha t} \right) \]

The standard method developed from the ideas of Murphy (1965) and Gulland (1962), is to independently establish \( F \) during the last year of catch for the cohort (starting \( F \)), and the loss rate due to unobserved causes \((N)\), and then calculate \( F + N \) at each age in the life of the cohort beginning in the last year and proceeding back through time to the earliest year of interest as follows:

Step 1) Let \( t = \) the last year of recorded catches of the cohort and solve equation (3) for \( N_t \). (If, however, \( C_t \) refers to catches not only in year \( t \) but instead to the sum of catches in year \( t \) and all subsequent years, then \( N_t = N_t A_t \).)

Step 2) Let \( t = t - 1 \).

Step 3) If \( t \) is less than the age the cohort first appeared in catches (or a previous age of interest) stop, otherwise proceed.
Step 4) Solve equation (7) for $F_L$

Step 5) Solve equation (2) for $N_L$. Go to Step 2.

Notice that equation (4) offers no closed solution for $F_L$ hence iterative methods are required. Pope (1972), concerned with the absence of a closed solution developed an approximation that is within 4% for $0.00 x 0.3$ and $0.04 x 1.2$ and referred to it as "cohort analysis".

If catches are instantaneous at midyear rather than continuous then the equation

$$\frac{N_{t+1} - N_t}{N_t} = -0.5M$$

holds which, upon rearrangement is

$$N_t = N_{t+1} e^{-0.5M}$$

Rearrangement of (1) gives

$$F = \frac{M}{N_t} e^{-0.5M}$$

$$e = \frac{N_t}{N_{t+1}}$$

Substituting the expression for $N_t$ from equation (6) gives the approximation proposed by Pope:

$$F = \ln \left[ 1 + \frac{C_L}{N_{t+1}} \right]$$

This approximation maybe employed rather than equation (8) and the same steps used.
Figure 1(b). Regression results using a single CPUE set for fully recruited fish.

Figure 1(c). Sum of Squares Initials using two CPUE sets, one for ages 5-7 and one for ages 8-9.

RESIDUALS (OBS,CPUE-EST,CPUE)

CPUE (X-axis) vs ESTIMATED STOCKSIZE (Y-axis)

0.000000 0.045568 0.574568 1.461048
Figure 2(b). Regression results for the age 5+ CPUE.

Figure 2(c). Regression results for the age 8+ CPUE.