On assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (Balaena mysticetus) using a Bayesian approach

A.E. PUNT* and D.S. Butterworth†

Contact e-mail: andre.punt@marine.csiro.au

ABSTRACT

This paper explores a number of issues surrounding the current assessment of the Bering-Chukchi-Beaufort (B-C-B) Seas stock of bowhead whales and provides a 'preferred' set of specifications for this assessment. A Bayesian approach appears to be preferable. However, the Bayesian Synthesis method is subject to the Borel paradox. Reverting to a 'standard' Bayesian approach which places all 'indirect' information in priors (rather than representing this information as likelihoods) would overcome this problem. The basis for the prior distributions used should be documented clearly, and the sources of information for the B-C-B bowhead stock divided into 'indirect' and 'direct'. Simulation results and 'in principle' arguments support the choice of a current population size rather than the pre-exploitation equilibrium size for the parameter to scale the population size (i.e. a 'backwards' rather than a 'forwards' approach). Arguments are presented that the most appropriate choice for a productivity-related parameter, for which a prior has to be specified, is the maximum steady rate of increase. A method for treating the $N_0$ and $P_a$ estimates as relative indices of abundance, allowing for prior information about the relationship between absolute abundance and those estimates, and accounting for the correlation among the indices of relative abundance derived from the $N_0$ and $P_a$ data is developed. Two 'preferred approaches' for assessing the resource both lead to estimates for the lower 5th percentile of the replacement yield that are greater than the current annual strike limit of 67 for the B-C-B stock.

KEYWORDS: BOWHEAD WHALE; POPN ASSESSMENT; TRENDS; BIOLOGICAL PARAMETERS; MODELLING; WHALING-ABORIGINAL; ARCTIC

INTRODUCTION

Bowhead whales (Balaena mysticetus) of the Bering-Chukchi-Beaufort (B-C-B) stock are subject to subsistence whaling in Alaska and Chukotka. Thus the assessment of this stock is important for providing management advice to the International Whaling Commission (IWC), the intergovernmental body that establishes catch limits. The present regulations state that the total number of landed whales for seasons 1998-2002 shall not exceed 280, with no more than 67 struck in any one year (IWC, 1999).

Recent assessments of this stock have been conducted using both conditioned maximum likelihood (e.g. Butterworth and Punt, 1992; 1995; Punt and Butterworth, 1996; 1997a) and Bayesian methods (e.g. Givens et al., 1995; Givens and Thompson, 1996). The Bayesian assessments have been based on Bayesian Synthesis (e.g. Raftery et al., 1995a) and standard Bayesian methods (e.g. Punt and Butterworth, 1997a; Breiwick, 1997). These Bayesian analyses involve the development of a coherent joint posterior distribution for seven population model parameters: the total (1+) pre-exploitation size of the resource, $K_{1+}$; $MSYR$; $MSYL$; the age-at-maturity, $a_w$; the survival rate of adults in the absence of exploitation, $S_{adul} = \exp(-M_{adul})$; the survival rate of juveniles in the absence of exploitation, $S_{juv} = \exp(-M_{juv})$; and the greatest age at which juvenile natural mortality applies, $a$. The assessment conducted by the IWC Scientific Committee (hereafter 'Scientific Committee') at its 1994 meeting (IWC, 1995) used pre-model distributions 1 for each of these parameters, as well as pre-model distributions for the recent rate of population increase (ROI), the 1988 (1+) population size ($P_{1988}$), the maximum pregnancy rate ($f_{max}$), and the proportion of mature animals and calves in the population from 1985 to 1992.

There are three main reasons for using a Bayesian approach for stock assessment: (a) it provides a relatively straightforward means to represent the full range of uncertainty (both parameter uncertainty and model-structure uncertainty); (b) information based on 'expert opinion' and inferences about other stocks/species can be incorporated explicitly into the stock assessment within a statistically defensible framework; and (c) the output of the analysis is exactly the information needed to parameterise operating models for evaluating alternative candidate management procedures (viz. the probability of alternative states of nature). Thus, unlike the situation for maximum likelihood approaches, it is not necessary to argue that the joint distribution obtained for parameter estimates can be assumed to represent these probabilities, because it is exactly these probabilities which a Bayesian approach provides.

The principles underlying Bayesian Synthesis have been criticised as this method is subject to the Borel paradox (Wolpert, 1995; Bravington, 1996). Put simply, the Borel paradox arises because there are (through the relationships provided by the population dynamics model) two different prior distributions for the same quantity (Raftery and Givens, 1997). Concern has also been expressed within the Scientific Committee about some of the prior distributions selected for the 1994 assessment (IWC, 1995) - see

---

1 The term 'pre-model distribution' will be reserved in this paper for references to Bayesian Synthesis applications. The more common terms 'prior' and 'likelihood' will be used when discussing issues related to standard Bayesian assessments.

2 The term 'pregnancy rate' refers to the fraction of females past the age-at-first-parturition that give birth in a year (Punt, 1996; 1999). This definition differs from usage in some other earlier papers (e.g. de la Mare, 1989; Punt and Butterworth, 1991) in that it applies to births of both sexes rather than to females only.

---

* Division of Marine Research, CSIRO Marine Laboratories, GPO Box 1538, Hobart, Tas 7001, Australia
† Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa
Bayesian Synthesis or Bayesian Analysis

Bayesian stock assessment and risk analysis methods have been applied in the fisheries field for several years (e.g. Walters and Hilborn, 1976; Bergh and Butterworth, 1987; Sainsbury, 1988; Collie and Walters, 1991; Thompson, 1992; Hilborn et al., 1994; McAllister et al., 1994; Walters and Ludwig, 1994; Walters and Punt, 1994). The assessment method applied to the South African fur seal population by Butterworth et al. (1987) and more recently by Givens et al. (1993; 1995) for the B-C-B bowhead stock differs from other Bayesian assessments because it is based on a Bayesian Synthesis approach rather than a standard Bayesian analysis. As such, these assessments are subject to the Borel paradox rather than a standard Bayesian analysis as discussed.

Bravington’s (1996) appraisal of Bayesian Synthesis highlights the Borel paradox and suggests that sensitivity to this paradox can be explored through relabelling of model inputs and outputs. This suggestion is both sensible and adequate but it is required only if the assessment has been provided with more priors than are actually needed. (As detailed in the following section, one of the two sources of the Borel paradox in the 1994 B-C-B assessment was removed by the Scientific Committee’s decision in 1997 not to include a prior on $S_{juv}$ (IWC, 1998c).)

Bayesian analysis deals with priors and likelihoods in different ways. However, the 1994 B-C-B bowhead assessment treats some priors (e.g. that for the maximum pregnancy rate) as likelihoods. This practice is dangerous and can readily be shown to lead to erroneous results (e.g. Bravington, 1996). Raftery and Poole (1997) and Poole and Raftery (1998) provide suggestions on how to combine priors in a manner that overcomes this problem. However, for the B-C-B stock of bowhead whales, the most obvious solution to this problem is to place all of the ‘indirect’ information into the prior distributions and to represent all of the data for the B-C-B bowhead stock in the form of a likelihood function. In this situation (which we will refer to as a ‘standard’ Bayesian assessment), the Borel paradox is not a concern provided the joint prior is of the same dimension as the parameter vector. Naturally, one cannot use a ‘standard’ Bayesian assessment if there really is ‘indirect’ information about both model inputs and outputs. However, we will argue below that the basis for some of the priors used in the 1994 B-C-B bowhead assessment is so weak that it is perhaps better to ignore certain of these priors and thus be able to take advantage of adopting a ‘standard’ Bayesian approach.

THE ‘REFERENCE’ ANALYSIS

In 1997, the Scientific Committee specified a ‘reference case’ for comparing alternative approaches to the assessment of the B-C-B bowhead stock (IWC, 1998b and see Tables 1 and 2). The 1994 B-C-B bowhead assessment (IWC, 1995) incorporated priors for $MSYR_{mud}$, $MSYR_{mud}$, $a$, $S_{sadult}$, $S_{juv}$ and $f_{max}$. However, given values for any six of these seven parameters, the value for the seventh can be derived from the BALEEN II population dynamics model (Punt, 1999). The use of all seven priors therefore leads to an instance of the Borel paradox. The specifications of the ‘reference case’ resolve this problem as no prior is placed on $S_{juv}$ and instead
the values for the parameters $S_{\text{adult}}$, $f_{\text{m}}$, $a$, $\text{MSYR}$, $\text{MSYL}$ and $f_{\text{ma}}$, and the relationships within BALEEN II are used to compute a value for $S_{\text{juv}}$. For ease of presentation, the analyses presented in this paper are all variants of this ‘reference case’. The results of the assessments are summarised by eight management-related quantities, the first seven of which were identified by IWC (1998c).

Table 1

<table>
<thead>
<tr>
<th>Year</th>
<th>Catch</th>
<th>Year</th>
<th>Catch</th>
<th>Year</th>
<th>Catch</th>
<th>Year</th>
<th>Catch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1848</td>
<td>18</td>
<td>1868</td>
<td>168</td>
<td>1924</td>
<td>41</td>
<td>1962</td>
<td>20</td>
</tr>
<tr>
<td>1849</td>
<td>573</td>
<td>1887</td>
<td>240</td>
<td>1925</td>
<td>53</td>
<td>1963</td>
<td>15</td>
</tr>
<tr>
<td>1850</td>
<td>2,067</td>
<td>1888</td>
<td>160</td>
<td>1926</td>
<td>35</td>
<td>1964</td>
<td>24</td>
</tr>
<tr>
<td>1851</td>
<td>898</td>
<td>1899</td>
<td>127</td>
<td>1927</td>
<td>14</td>
<td>1965</td>
<td>14</td>
</tr>
<tr>
<td>1852</td>
<td>2,709</td>
<td>1890</td>
<td>136</td>
<td>1928</td>
<td>30</td>
<td>1966</td>
<td>24</td>
</tr>
<tr>
<td>1853</td>
<td>807</td>
<td>1891</td>
<td>284</td>
<td>1929</td>
<td>30</td>
<td>1967</td>
<td>12</td>
</tr>
<tr>
<td>1854</td>
<td>166</td>
<td>1892</td>
<td>346</td>
<td>1930</td>
<td>17</td>
<td>1968</td>
<td>27</td>
</tr>
<tr>
<td>1855</td>
<td>2</td>
<td>1893</td>
<td>180</td>
<td>1931</td>
<td>32</td>
<td>1969</td>
<td>32</td>
</tr>
<tr>
<td>1856</td>
<td>0</td>
<td>1894</td>
<td>234</td>
<td>1932</td>
<td>27</td>
<td>1970</td>
<td>48</td>
</tr>
<tr>
<td>1857</td>
<td>78</td>
<td>1895</td>
<td>117</td>
<td>1933</td>
<td>21</td>
<td>1971</td>
<td>25</td>
</tr>
<tr>
<td>1858</td>
<td>461</td>
<td>1896</td>
<td>118</td>
<td>1934</td>
<td>21</td>
<td>1972</td>
<td>44</td>
</tr>
<tr>
<td>1859</td>
<td>372</td>
<td>1897</td>
<td>130</td>
<td>1935</td>
<td>15</td>
<td>1973</td>
<td>51</td>
</tr>
<tr>
<td>1860</td>
<td>221</td>
<td>1898</td>
<td>309</td>
<td>1936</td>
<td>24</td>
<td>1974</td>
<td>42</td>
</tr>
<tr>
<td>1861</td>
<td>306</td>
<td>1899</td>
<td>234</td>
<td>1937</td>
<td>53</td>
<td>1975</td>
<td>32</td>
</tr>
<tr>
<td>1862</td>
<td>157</td>
<td>1900</td>
<td>148</td>
<td>1938</td>
<td>36</td>
<td>1976</td>
<td>74</td>
</tr>
<tr>
<td>1863</td>
<td>303</td>
<td>1901</td>
<td>55</td>
<td>1939</td>
<td>18</td>
<td>1977</td>
<td>99</td>
</tr>
<tr>
<td>1864</td>
<td>434</td>
<td>1902</td>
<td>162</td>
<td>1940</td>
<td>20</td>
<td>1978</td>
<td>17</td>
</tr>
<tr>
<td>1865</td>
<td>590</td>
<td>1903</td>
<td>116</td>
<td>1941</td>
<td>38</td>
<td>1979</td>
<td>23</td>
</tr>
<tr>
<td>1866</td>
<td>554</td>
<td>1904</td>
<td>86</td>
<td>1942</td>
<td>26</td>
<td>1980</td>
<td>38</td>
</tr>
<tr>
<td>1867</td>
<td>599</td>
<td>1905</td>
<td>105</td>
<td>1943</td>
<td>14</td>
<td>1981</td>
<td>26</td>
</tr>
<tr>
<td>1868</td>
<td>516</td>
<td>1906</td>
<td>69</td>
<td>1944</td>
<td>8</td>
<td>1982</td>
<td>14</td>
</tr>
<tr>
<td>1869</td>
<td>382</td>
<td>1907</td>
<td>96</td>
<td>1945</td>
<td>23</td>
<td>1983</td>
<td>16</td>
</tr>
<tr>
<td>1870</td>
<td>637</td>
<td>1908</td>
<td>123</td>
<td>1946</td>
<td>20</td>
<td>1984</td>
<td>16</td>
</tr>
<tr>
<td>1871</td>
<td>138</td>
<td>1909</td>
<td>61</td>
<td>1947</td>
<td>21</td>
<td>1985</td>
<td>14</td>
</tr>
<tr>
<td>1872</td>
<td>200</td>
<td>1910</td>
<td>37</td>
<td>1948</td>
<td>8</td>
<td>1986</td>
<td>22</td>
</tr>
<tr>
<td>1873</td>
<td>147</td>
<td>1911</td>
<td>48</td>
<td>1949</td>
<td>11</td>
<td>1987</td>
<td>29</td>
</tr>
<tr>
<td>1874</td>
<td>95</td>
<td>1912</td>
<td>39</td>
<td>1950</td>
<td>23</td>
<td>1988</td>
<td>28</td>
</tr>
<tr>
<td>1875</td>
<td>200</td>
<td>1913</td>
<td>23</td>
<td>1951</td>
<td>23</td>
<td>1989</td>
<td>25</td>
</tr>
<tr>
<td>1876</td>
<td>76</td>
<td>1914</td>
<td>61</td>
<td>1952</td>
<td>11</td>
<td>1990</td>
<td>14</td>
</tr>
<tr>
<td>1877</td>
<td>270</td>
<td>1915</td>
<td>23</td>
<td>1953</td>
<td>41</td>
<td>1991</td>
<td>47</td>
</tr>
<tr>
<td>1878</td>
<td>80</td>
<td>1916</td>
<td>23</td>
<td>1954</td>
<td>9</td>
<td>1992</td>
<td>46</td>
</tr>
<tr>
<td>1879</td>
<td>266</td>
<td>1917</td>
<td>35</td>
<td>1955</td>
<td>36</td>
<td>1993</td>
<td>51</td>
</tr>
<tr>
<td>1880</td>
<td>480</td>
<td>1918</td>
<td>27</td>
<td>1956</td>
<td>11</td>
<td>1994</td>
<td>38</td>
</tr>
<tr>
<td>1881</td>
<td>435</td>
<td>1919</td>
<td>33</td>
<td>1957</td>
<td>5</td>
<td>1995</td>
<td>57</td>
</tr>
<tr>
<td>1882</td>
<td>242</td>
<td>1920</td>
<td>33</td>
<td>1958</td>
<td>5</td>
<td>1996</td>
<td>45</td>
</tr>
<tr>
<td>1883</td>
<td>42</td>
<td>1921</td>
<td>9</td>
<td>1959</td>
<td>2</td>
<td>1997</td>
<td>62</td>
</tr>
<tr>
<td>1884</td>
<td>160</td>
<td>1922</td>
<td>39</td>
<td>1960</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1885</td>
<td>377</td>
<td>1923</td>
<td>12</td>
<td>1961</td>
<td>17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results for the ‘reference case’

Table 3 lists post-model-pre-data and posterior distributions for the ‘backwards’ and ‘forwards’ approaches. Results are shown in Table 3 for the ‘forwards’ approach for $Z_1 = 1,000,000$ (‘reference case’) and $Z_1 = 2,500,000$. The maximum weight for anyone draw for the reference case ‘backwards’ analysis (0.00066) suggests that $Z_1 = 250,000$ is more than sufficient to obtain an adequate numerical representation of the posterior. In contrast, the maximum weight for the reference case ‘forwards’ analysis (0.02286) is perhaps larger than desirable. Increasing $Z_1$ from 1,000,000 to 2,500,000 decreases the maximum weight to 0.00894, which seems adequate. The results for these two choices of $Z_1$, however, differ only marginally (Table 3).

The posterior distribution is approximated numerically using a variant of the Sampling-Importance-Resampling (SIR) algorithm. This involves drawing $Z_1$ sets of parameter values from the joint prior distribution and then calculating the likelihood corresponding to each vector. The likelihood is set equal to zero if the value for $S_{\text{juv}}$ is greater than that for $S_{\text{adult}}$, or if the population is rendered extinct. The posterior is then based on $Z_2 = 5,000$ draws (with replacement) from the $Z_1$ sets of parameter values, where the probability of selecting a given parameter set is proportional to its likelihood. The maximum weight (the ratio of the likelihood for most likely set of parameter values to the total likelihood over all sets of parameter values) is used to assess whether the SIR algorithm has converged adequately to the posterior distribution.

$K_{i+}$ - the pre-exploitation size of the 1+ component of the population.
$P_{1+}^{\text{post}}/K_{i+}$ - the ratio (expressed as percentage) of the size of the 1+ component of the population at the start of 1998 to $K_{i+}$.
$P_{1+}^{\text{post}}/K^i$ - the ratio (expressed as percentage) of the size of the mature female component of the population at the start of 1998 to the corresponding pre-exploitation size.
$\text{MSYR}_{1+}$ - $\text{MSYR}$ for uniform selectivity harvesting of the 1+ component of the population, expressed as a percentage.
$Q_0$ (1998) - the value of the quantity $Q_0$ (Wade and Givens, 1997) for 1998:

$$Q_0(1998) = \min(\text{RY}(1998))$$

where $\text{MSYR}_{1+}$ = $\text{MSYR}_{1+}$, $\text{MSYR}_{1+}$, $K_{i+}$.

Slope - the annual rate of increase of the 1+ population from 1978 to 1993, expressed as a percentage.

The results for the ‘reference case’

Table 3 lists post-model-pre-data and posterior distributions for the ‘backwards’ and ‘forwards’ approaches. Results are shown in Table 3 for the ‘forwards’ approach for $Z_1 = 1,000,000$ (‘reference case’) and $Z_1 = 2,500,000$. The maximum weight for anyone draw for the reference case ‘backwards’ analysis (0.00066) suggests that $Z_1 = 250,000$ is more than sufficient to obtain an adequate numerical representation of the posterior. In contrast, the maximum weight for the reference case ‘forwards’ analysis (0.02286) is perhaps larger than desirable. Increasing $Z_1$ from 1,000,000 to 2,500,000 decreases the maximum weight to 0.00894, which seems adequate. The results for these two choices of $Z_1$, however, differ only marginally (Table 3).

The posterior distributions differ markedly from the post-model-pre-data distributions (both in terms of precision and central tendency). The post-model-pre-data distributions for ‘backwards’ are more similar to the posteriors because the ‘backwards’ projections include the prior information.

<table>
<thead>
<tr>
<th>Post-model-pre-data</th>
<th>$K_0$</th>
<th>$R_Y$ (1998)</th>
<th>$Q_0$ (1998)</th>
<th>$P_{F1999}^{U}/K_{1+}$</th>
<th>$P_{F1999}^{L}/K_{1+}$</th>
<th>$P_{F1999}^{U}/MSTL_{1+}$</th>
<th>$MSTL_{1+}$</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forwards</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference case</td>
<td>13995</td>
<td>80 (180)</td>
<td>171 (174)</td>
<td>64.1 (63.9)</td>
<td>45.3 (43.7)</td>
<td>94.5 (93.0)</td>
<td>2.04 (2.05)</td>
<td>1.94 (1.95)</td>
</tr>
<tr>
<td>$Z_{&lt;2500,000}$</td>
<td>11547</td>
<td>110 (250)</td>
<td>204 (249)</td>
<td>47.5 (85.4)</td>
<td>36.1 (55.8)</td>
<td>69.9 (119.3)</td>
<td>1.12 (3.35)</td>
<td>1.09 (3.60)</td>
</tr>
<tr>
<td>With plus-group</td>
<td>11758</td>
<td>110 (250)</td>
<td>204 (249)</td>
<td>47.5 (85.4)</td>
<td>36.1 (55.8)</td>
<td>69.9 (119.3)</td>
<td>1.12 (3.35)</td>
<td>1.09 (3.60)</td>
</tr>
<tr>
<td>$S_{max} = 0.999$</td>
<td>11078</td>
<td>106 (242)</td>
<td>204 (249)</td>
<td>48.6 (90.7)</td>
<td>35.0 (53.5)</td>
<td>73.0 (125.8)</td>
<td>1.16 (1.47)</td>
<td>1.09 (3.60)</td>
</tr>
<tr>
<td>Alt Preg constraint</td>
<td>11425</td>
<td>115 (268)</td>
<td>109 (265)</td>
<td>48.5 (85.1)</td>
<td>36.4 (54.0)</td>
<td>72.2 (117.3)</td>
<td>1.18 (3.37)</td>
<td>1.09 (3.60)</td>
</tr>
<tr>
<td>Backwards</td>
<td>12631</td>
<td>211 (209)</td>
<td>209 (207)</td>
<td>71.2 (70.9)</td>
<td>45.9 (46.1)</td>
<td>100.2 (99.5)</td>
<td>2.59 (2.58)</td>
<td>2.49 (2.48)</td>
</tr>
<tr>
<td>Reference case</td>
<td>10992</td>
<td>141 (273)</td>
<td>136 (279)</td>
<td>53.5 (50.1)</td>
<td>37.8 (57.9)</td>
<td>76.8 (122.8)</td>
<td>1.51 (3.78)</td>
<td>1.39 (3.67)</td>
</tr>
<tr>
<td>With plus-group</td>
<td>11835</td>
<td>193 (191)</td>
<td>206 (204)</td>
<td>75.2 (74.9)</td>
<td>43.9 (44.4)</td>
<td>104.9 (101.4)</td>
<td>2.67 (2.66)</td>
<td>2.49 (2.48)</td>
</tr>
<tr>
<td>$S_{max} = 0.999$</td>
<td>12509</td>
<td>215 (212)</td>
<td>215 (212)</td>
<td>72.1 (71.8)</td>
<td>45.8 (46.2)</td>
<td>101.5 (100.9)</td>
<td>2.67 (2.66)</td>
<td>2.59 (2.58)</td>
</tr>
<tr>
<td>Alt Preg constraint</td>
<td>10843</td>
<td>145 (275)</td>
<td>139 (281)</td>
<td>54.6 (61.1)</td>
<td>38.0 (58.3)</td>
<td>78.5 (124.1)</td>
<td>1.55 (3.89)</td>
<td>1.43 (3.74)</td>
</tr>
</tbody>
</table>

Comparing alternative stock assessment methods

Punt and Butterworth (1997a) evaluated the relative performances of three alternative estimation procedures (two maximum likelihood methods and the 'forwards' Bayesian Synthesis approach) for the B-C-B bowhead stock by means of a Monte Carlo simulation exercise. The evaluation involved generating 100 sets of artificial abundance and 'proportion' data, applying each estimation approach to each data set, and then comparing point estimates (posterior medians for these Bayesian methods) with true values. The results of the simulation trials were summarised in Punt and Butterworth (1997a) by the biases and root-mean-square errors (RMSEs) (expressed in relative terms) of four quantities (indicated by $Q$ below) of interest to management:

$$\hat{\hat{Q}} = (1 + \beta)Q^{true.U} + \varepsilon$$

and

$$RMSE(Q) = \sqrt{\frac{1}{100} \sum_{U=1}^{100} (\hat{\hat{Q}} - Q^{true.U})^2}$$

where $Q^{true.U}$ is the true value of quantity Q in simulation U,$\hat{\hat{Q}}$ is the estimate of Q in simulation U,$\beta$ is the relative bias, and $\varepsilon$ is assumed to be a normally distributed random variable.
The analyses of this paper involve applying the simulation testing framework developed by Punt and Butterworth (1997a) to compare the estimation ability of the ‘backwards’ and ‘forwards’ approaches to Bayesian analysis. Punt and Butterworth (1997a) considered eleven trials, the first nine of which involved fixed values for the biological parameters. Here we consider only the remaining two trials, which involved generating true values for the biological parameters from the posterior distributions obtained from either the ‘forwards’ or the ‘backwards’ variants of the Bayesian assessment.

IWC (1997) noted that previous simulation evaluations had made no attempt to compare estimation methods for the B-C-B bowhead stock with respect to their estimates of precision. Both estimation procedures are Bayesian, and so can readily be applied to provide comparable 90% credibility intervals. The intervals are compared for each management quantity using three measures of performance: (a) the probability that the 90% credibility interval includes the true value, (b) the probability that the true value is smaller than the lower 90% limit and (c) the probability that the true value is larger than the upper 90% limit. If the estimation procedure performed ‘perfectly’, the values for these quantities would be 0.90, 0.05 and 0.05.

Table 4 lists the relative biases and RMSEs for the ‘forwards’ and ‘backwards’ approaches for six quantities of interest to management ($K_{1+}$, $MSYR_{1+}$, $Q_0$ (1998), $P_{T1998}/K_1$, $P_{T1998}/MSYL_{1+}$, and $RY$ (1998)). Figs 1 and 2 plot the actual and estimated values for four of the six quantities for the two trials. Results are not shown in these figures for $RY$ (1998) and $P_{T1998}/MSYL_{1+}$, because they are qualitatively the same as those for $Q_0$ (1998) and $P_{T1998}/K_1$ respectively. Not surprisingly, the performance of the ‘forwards’ estimation approach is better when ‘forwards’ rather than ‘backwards’ is used to generate the true data, although it remains poor in both cases. But importantly, whichever approach is used to generate the data, the ‘backwards’ estimation approach outperforms its ‘forwards’ counterpart in terms of both RMSEs and the absolute size of the bias. Both approaches tend to provide ‘conservative’ (i.e. negatively biased) estimates of the management quantities upon which catch limits would be based (Table 4; Figs 1 and 2).

In terms of coverage probability, the ‘backwards’ approach again performs better than the ‘forwards’ approach (Table 5 on p. 60). The poor performance of the ‘forwards’ approach is attributable to the estimate of the upper 90% credibility value is far too low for all of the quantities except $K_{1+}$, for which the lower limit is too high.

This comparison overestimates the confidence to be placed in the Bayesian credibility intervals because all the estimators assume the exact form of the true population dynamics model, and further because the assumption of deterministic dynamics made by all the estimation procedures is correct. Had the simulations allowed for process error effects (such as variation in the juvenile survival rate or uncertainty about historical catches), it is likely that the credibility intervals would have been shown to be overly narrow. Punt and Butterworth (1993) demonstrate that coefficients of variation estimated using bootstrap procedures for fake assessments can be negatively biased by some tens of percent when observation errors (but not process errors) are taken into account.

<table>
<thead>
<tr>
<th>Estimation procedure</th>
<th>Simulation trial</th>
<th>Forwards</th>
<th>Backwards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backwards</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{1+}$</td>
<td>15.9 (21.5)</td>
<td>4.7 (12.6)</td>
<td></td>
</tr>
<tr>
<td>$MSYR_{1+}$</td>
<td>-37.0 (36.1)</td>
<td>-17.2 (25.2)</td>
<td></td>
</tr>
<tr>
<td>$Q_0$ (1998)</td>
<td>-31.1 (15.5)</td>
<td>-31.5 (15.1)</td>
<td></td>
</tr>
<tr>
<td>$P_{T1998}/K_1$</td>
<td>-14.8 (16.5)</td>
<td>-8.3 (11.9)</td>
<td></td>
</tr>
<tr>
<td>$P_{T1998}/MSYL_{1+}$</td>
<td>-15.6 (18.0)</td>
<td>-7.7 (13.6)</td>
<td></td>
</tr>
<tr>
<td>$RY$ (1998)</td>
<td>-29.5 (30.6)</td>
<td>-13.6 (20.4)</td>
<td></td>
</tr>
<tr>
<td>Forwards</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{1+}$</td>
<td>8.2 (15.7)</td>
<td>-1.2 (10.2)</td>
<td></td>
</tr>
<tr>
<td>$MSYR_{1+}$</td>
<td>-28.3 (28.3)</td>
<td>-7.6 (27.3)</td>
<td></td>
</tr>
<tr>
<td>$Q_0$ (1998)</td>
<td>-23.6 (26.5)</td>
<td>-5.5 (24.9)</td>
<td></td>
</tr>
<tr>
<td>$P_{T1998}/K_1$</td>
<td>-9.8 (12.9)</td>
<td>-4.1 (10.1)</td>
<td></td>
</tr>
<tr>
<td>$P_{T1998}/MSYL_{1+}$</td>
<td>-10.8 (15.3)</td>
<td>-4.0 (13.2)</td>
<td></td>
</tr>
<tr>
<td>$RY$ (1998)</td>
<td>-22.9 (25.5)</td>
<td>-7.1 (22.3)</td>
<td></td>
</tr>
</tbody>
</table>

Priors and Likelihoods

The analyses presented above illustrate the need to identify those quantities for which some of the available information comes from ‘indirect’ sources (e.g. inferences from data for other stocks/species), and those for which all of the information comes from ‘direct’ sources. In a standard Bayesian assessment, the former quantities must be included as priors, while the information contained in the latter should form part of the likelihood. In the 1994 B-C-B bowhead assessment (IWC, 1995), two of the prior distributions (those for $P_{max}$ and $P_{1998}$) were treated as ‘data’ rather than priors in the analysis, even though they were based, in part, on ‘direct’ evidence. The contribution of $P_{1998}$ to the likelihood function included ‘indirect’ information for example, about whale numbers and behaviour - e.g. Raftery and Zeh (1991) as well as information from the ‘direct’ count data collected at Point Barrow, Alaska. The likelihood contribution for $P_{max}$ was not based on any direct information about the pregnancy rate of bowhead whales at very low population size, but rather on inferences about what this rate might be, taking account of perceptions/observations for other baleen whale species (IWC, 1992; 1995).

The first step used in this process is to list the various sources of information which could contribute to the assessment, and then to clarify which are data (and so should be incorporated into the likelihood function) and which constitute ‘indirect’ information (and must therefore form part of the (joint) prior distribution). Sainsbury et al. (1998) highlight the point that this step in the process of conducting a Bayesian assessment has often been missing.

<p>| Table 4 Percentage biases and root mean square errors (in parentheses) for two estimators and two management-related quantities (see text for details). |</p>
<table>
<thead>
<tr>
<th>Simulation trial</th>
<th>Forwards</th>
<th>Backwards</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_{1+}$</td>
<td>15.9 (21.5)</td>
<td>4.7 (12.6)</td>
</tr>
<tr>
<td>$MSYR_{1+}$</td>
<td>-37.0 (36.1)</td>
<td>-17.2 (25.2)</td>
</tr>
<tr>
<td>$Q_0$ (1998)</td>
<td>-31.1 (15.5)</td>
<td>-31.5 (15.1)</td>
</tr>
<tr>
<td>$P_{T1998}/K_1$</td>
<td>-14.8 (16.5)</td>
<td>-8.3 (11.9)</td>
</tr>
<tr>
<td>$P_{T1998}/MSYL_{1+}$</td>
<td>-15.6 (18.0)</td>
<td>-7.7 (13.6)</td>
</tr>
<tr>
<td>$RY$ (1998)</td>
<td>-29.5 (30.6)</td>
<td>-13.6 (20.4)</td>
</tr>
<tr>
<td>Forwards</td>
<td>-8.2 (15.7)</td>
<td>-1.2 (10.2)</td>
</tr>
<tr>
<td>$MSYR_{1+}$</td>
<td>-28.3 (28.3)</td>
<td>-7.6 (27.3)</td>
</tr>
<tr>
<td>$Q_0$ (1998)</td>
<td>-23.6 (26.5)</td>
<td>-5.5 (24.9)</td>
</tr>
<tr>
<td>$P_{T1998}/K_1$</td>
<td>-9.8 (12.9)</td>
<td>-4.1 (10.1)</td>
</tr>
<tr>
<td>$P_{T1998}/MSYL_{1+}$</td>
<td>-10.8 (15.3)</td>
<td>-4.0 (13.2)</td>
</tr>
<tr>
<td>$RY$ (1998)</td>
<td>-22.9 (25.5)</td>
<td>-7.1 (22.3)</td>
</tr>
</tbody>
</table>

6 It is often computationally more efficient to update priors based solely on ‘indirect’ data with the ‘direct’ data for the parameters concerned - this does not impact the final results at all.

7 The analyses presented above illustrate the need to identify those quantities for which some of the available information comes from ‘indirect’ sources (e.g. inferences from data for other stocks/species), and those for which all of the information comes from ‘direct’ sources. In a standard Bayesian assessment, the former quantities must be included as priors, while the information contained in the latter should form part of the likelihood. In the 1994 B-C-B bowhead assessment (IWC, 1995), two of the prior distributions (those for $P_{max}$ and $P_{1998}$) were treated as ‘data’ rather than priors in the analysis, even though they were based, in part, on ‘direct’ evidence. The contribution of $P_{1998}$ to the likelihood function included ‘indirect’ information for example, about whale numbers and behaviour - e.g. Raftery and Zeh (1991) as well as information from the ‘direct’ count data collected at Point Barrow, Alaska. The likelihood contribution for $P_{max}$ was not based on any direct information about the pregnancy rate of bowhead whales at very low population size, but rather on inferences about what this rate might be, taking account of perceptions/observations for other baleen whale species (IWC, 1992; 1995).

The first step used in this process is to list the various sources of information which could contribute to the assessment, and then to clarify which are data (and so should be incorporated into the likelihood function) and which constitute ‘indirect’ information (and must therefore form part of the (joint) prior distribution). Sainsbury et al. (1998) highlight the point that this step in the process of conducting a Bayesian assessment has often been missing.

‘there should be much more careful documentation of the steps involved in successive updating (i.e. the initial definition of the prior, the information used to calculate a posterior that in turn is the prior for the next iteration of the analysis)’.

8 It is often computationally more efficient to update priors based solely on ‘indirect’ data with the ‘direct’ data for the parameters concerned - this does not impact the final results at all.
Fig. 1. Plots of predicted and true values for $K_{t+}$, $MSY_{t+}$, $P_{est}/K'$, and $Q_0$ (1998) for the ‘backwards’ simulation trial. Results are shown for (a) the ‘forwards’ and (b) the ‘backwards’ estimation approaches.
Fig. 2. Plots of predicted and true values for $K_{t+}$, $MSY_{t+}$, $F_{(1998)/K'}$, and $Q_0$ (1998) for the 'forwards' simulation trial. Results are shown for (a) the 'forwards' and (b) the 'backwards' estimation approaches.
approach requires a prior distribution for 
assessments have been a focus for discussion. The ‘forwards’
are invariant to such transformations.
highly sensitive to it because a Bayesian assessment is not
approaches, this choice can be very important and the results may be
transformations of the model parameters. However, for a Bayesian
current biomass) is irrelevant because the likelihood is invariant to
the choice of this parameter (whether, for example,
methods based on (conditional) maximum likelihood estimation,
rate in any year as this scaling parameter. For stock assessment
avoids the
realisations from a prior distribution for a recent estimate of
this prior distribution. In contrast, the ‘backwards’ approach uses
priors
function.
approaches to including the abundance data in the likelihood
statistical procedures, but does comment on alternative
approaches to including the abundance data in the likelihood
function.
Absolute abundance
All stock assessments must incorporate a parameter that scales the
overall abundance. Punt and Hilborn (1997) note that this
parameter is of particular importance in most assessments, but that
data for other stocks/species can rarely be used to construct an
informative prior for it. In many stock assessments, this parameter
is chosen to be \(K\) (the pre-exploitation equilibrium biomass),
although it is possible to select the biomass/numbers/exploitation
rate in any year as this scaling parameter. For stock assessment
methods based on (conditional) maximum likelihood estimation,
the choice of this parameter (whether, for example, \(K\) or
the current biomass) is irrelevant because the likelihood is invariant
to transformations of the model parameters. However, for a Bayesian
approach, this choice can be very important and the results may be
highly sensitive to it because a Bayesian assessment is not invariant
to such transformations.

Two approaches to the B-C-B bowhead Bayesian Synthesis
assessment have been a focus for discussion. The ‘forwards’
approach requires a prior distribution for \(K_{1+}\) and projects
population trajectories forwards from realisations generated from
this prior distribution. In contrast, the ‘backwards’ approach uses
realisations from a prior distribution for a recent estimate of
absolute abundance to essentially extrapolate trajectories back to
1848; thus it provides an implicit distribution for \(K_{1+}\), and so
avoid it
need for an explicit specification of a prior distribution for this
parameter. Punt and Butterworth (1997a) show that the results of
the ‘backwards’ approach are not notably sensitive to the year
(within the period of the past two decades) for which abundance
estimates are available which is selected to provide the recent
estimate of abundance.

Considerable attention has been directed towards identifying
the reason for the difference in the results for these two
approaches (see Table 3). Much of the debate initially centred on
the justification of the basis used to provide the ‘direct’
component of the prior distribution for \(K_{1+}\). For the 1994
assessment (IWC, 1995), this was based on an application of the
DeLury (1947) estimation procedure to historical (1849-1870)
catch per unit effort (CPUE) data. Butterworth and Punt (1995)
criticised the derivation of this prior because the DeLury method
effectively assumes that \(MSY = 0\) and because the general
acceptance of CPUE as an index of abundance has proved
problematic in the past in the Scientific Committee (IWC, 1988,
p.35; IWC, 1989). Punt and Butterworth (1996) and Givens and
Thompson (1996) show, however, that including the ‘direct’
component of the \(K_{1+}\) prior in the likelihood when applying the
‘backwards’ method has virtually no impact on the results.
Subsequently, Raftery and Poole (1997) showed that the reason
for the differences between the results for ‘forwards’ and
‘backwards’ is attributable to differences in the joint region of
support for \(P_{1993}, K_{1+}\), and \(MSY\) for the two approaches.

This is a case in which there are two priors for the same
quantity (the parameter that scales the population). However, although
the full pooling approach of Raftery and Poole (1997)
and Poole and Raftery (1998) removes the associated problem of
the Borel paradox, in doing so it introduces a new one, namely
how to choose the pooling weight that is to be placed on the two
prior (\(K_{1+}\) and \(P_{1993}\)) when conducting full pooling. Clearly
results are sensitive to the weight chosen, as demonstrated by the
differences in results for the two extreme choices for this weight
(corresponding to ‘backwards’ and ‘forwards’) which are shown
in Table 3. In addition, the assumption of a priori independence
between \(K_{1+}\) and \(MSY\) underlyng the inclusion of ‘forwards’ in
full-pooling is violated because, prior to inclusion of the data in
the assessment, some combinations of \(K_{1+}\) and \(MSY\) can be
rejected as implausible (see below for further details). Rather than
attempting to combine these two priors, we prefer instead to
choose the more appropriate of the two.

This choice initially seems rather arbitrary because both seem
reasonable a priori. However, the simulation results of Table 4 are
available to guide a choice in this matter. These suggest that
the assumption underlying the ‘backwards’ approach is more
appropriate as it leads to lower MSEs and less biased 90%
credibility intervals. There are also two ‘in principle’ reasons for
preferring a current rather than a historical population size as the
parameter which scales the

<table>
<thead>
<tr>
<th>Simulation trial</th>
<th>Forwards</th>
<th>Backwards</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K_{1+})</td>
<td>0.62</td>
<td>0.38</td>
</tr>
<tr>
<td>(MSY_{1+})</td>
<td>0.59</td>
<td>0.00</td>
</tr>
<tr>
<td>(Q_{1+})</td>
<td>0.63</td>
<td>0.00</td>
</tr>
<tr>
<td>(P_{Y_{1993}}/K)</td>
<td>0.65</td>
<td>0.00</td>
</tr>
<tr>
<td>(P_{Y_{1993}}/MSY_{1+})</td>
<td>0.69</td>
<td>0.01</td>
</tr>
<tr>
<td>(RY_{1998})</td>
<td>0.66</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The 1994 application of the Bayesian Synthesis method (IWC,
1995) was based on seven input pre-model (prior) distributions
and five terms in the likelihood function. The basis for each of the
prior distributions based on ‘indirect’ data is discussed below,
including comments on some of the updates reflected by the
‘reference case’ specifications of ‘IWC (1998c). This section does
not deal in detail with the derivation of the ‘direct’ data (e.g. the
‘proportion’ data) because they were derived using standard
statistical procedures, but does comment on alternative
approaches to including the abundance data in the likelihood
function.

7 Logarithmic pooling, which is the approach generalised in full
pooling, given two priors \(P_1\) and \(P_2\) for the same parameter, provides a
pooled prior proportional to \(|P_1|^{1-a}|P_2|^{1-a}\) where the pooling weight a
reflects the relative reliability accorded to the two sources of
information that underlie the two priors specified. Raftery and Poole
(1997) argue for \(a=0.5\) when pooling priors for \(P_{Y_{1993}}\) and \(K_{1+}\) on the
basis of invariance under relabeling of inputs and outputs (initial and
current population sizes), suggesting also that two priors that are
agreed by the same expert (the IWC Scientific Committee) should
accordingly be deemed equally reliable. We disagree with these views,
judging for the various reasons put forward in this paper that the (or
indeed any) prior advanced for \(K_{1+}\) in the particular case of the B-C-B
bowhead population is much less reliable than that for \(P_{Y_{1993}}\).
population size. The first is that there is no need when applying the ‘backwards’ approach to specify that the population is not currently extinct, because this is incorporated implicitly in the prior for current population size. If a ‘forwards’ approach is taken, it is necessary to place a prior distribution on current depletion (or current population size) and incorporate it as a bound in the likelihood function. The second reason is that, in the absence of data, the ‘backwards’ approach does not update any of the prior distributions (i.e. the post-model-pre-data distributions for $M_{\text{Syr}}$, $M_{\text{Syl}}$, $a_m$, $a$ and $S_{\text{adult}}$ are identical to their priors\(^8\)). In contrast, the ‘forwards’ approach updates the joint prior distribution substantially (inter alia because combinations of low $K_{1+}$ and low $M_{\text{Syr}}$ correspond to extinction and can thus be excluded). Intuitively, it would seem undesirable to update prior distributions in the absence of direct data. The greater difference between the post-model-pre-data distribution and the prior for $M_{\text{Syr}}$ for ‘backwards’ than ‘forwards’ (contrast the prior in Table 2 with post-model-pre-data distributions in Table 3) would seem to contradict this. However, there are two factors that determine the post-model-pre-data distribution. The first is the impact of rejecting parameter combinations that give rise to juvenile survival rates exceeding $S_{\text{finite}}$, and the second is the impact of the effect just discussed. The post-model-pre-data distribution for $M_{\text{Syr}}$ for ‘backwards’ reflects only the first while that for ‘forwards’ reflects both. As expected, the post-model-pre-data distribution for ‘forwards’ gives greater probability to higher values for $M_{\text{Syr}}$.

Another seemingly undesirable property of the ‘forwards’ approach is that once a single abundance estimate becomes available, the joint distribution for the biological parameters (including $M_{\text{Syr}}$) is updated. This seems intuitively undesirable because a single abundance estimate does not provide any information about $M_{\text{Syr}}$. The past application of the ‘Hitting with fixed MSYR’ methodology by the Scientific Committee constitutes specific concurrence with this assertion. Even though an estimate of absolute abundance is available, all values for $M_{\text{Syr}}$ remain equally likely\(^9\) because there is always some combination of $M_{\text{Syr}}$ and $K_{1+}$ which ‘hits’ the estimate of abundance exactly. As a consequence, the likelihood is the same for all values of $M_{\text{Syr}}$, but the associated prior is updated in the Bayesian integration under the ‘forwards’ approach. This is because the non-linearity of the relationship between $M_{\text{Syr}}$ and $K_{1+}$ (given a single estimate of abundance) means that equal intervals on the $K_{1+}$ axis do not correspond to equal intervals on the $M_{\text{Syr}}$ axis (Butterworth and Punt, 1997, illustrate what amounts to this point). In contrast to this situation for the ‘forwards’ approach, the effect of a single estimate of abundance on the ‘backwards’ approach is to update the prior for current abundance without having any impact on the distribution for $M_{\text{Syr}}$.

A concern with the ‘backwards’ approach is its use of the $N(7800; 1300^2)$ prior for $P_{1993}$. The basis for this prior is unclear because there is no obvious independent information that could be used to construct an informative prior for $P_{1993}$. However, the ‘backwards’ approach can be applied using an ‘uninformative’ $U[0, \infty)$ prior for $P_{1993}$. As the likelihood for $P_{1993}$ would be very informative compared to any sensible ‘uninformative’ prior for $P_{1993}$, the results are unlikely to be very sensitive to the choice of an ‘uninformative’ prior for $P_{1993}$.

**MSYR**

The second parameter that all stock assessments have to incorporate is one that determines the overall productivity of the resource. In the BALEEN II population dynamics model, this is the resilience parameter ($A$). The 1994 B-C-B bowhead assessment (IWC, 1995) placed a prior distribution on $M_{\text{Syr}}$ rather than on $A$ (presumably because scientists are familiar with values for $M_{\text{Syr}}$, which facilitates comparison among stocks/species, unlike the situation for $A$). Other possible choices for the productivity parameter include the increase rate at low population size, $\lambda_{\text{max}}$, and the current rate of increase, ROI (Punt, 1999).

The $U[1\%; 7\%]$ prior for $M_{\text{Syr}}$ selected by the Scientific Committee (IWC, 1995) is consistent with that used in the development of the Catch Limit Algorithm for commercial whaling. Some concern has been expressed over the validity of the approach taken and its consistency in previous discussions about $M_{\text{Syr}}$ (IWC, 1995, p.148). Butterworth and Punt (1995) point out that the upper 2.5%ile of the posterior for $M_{\text{Syr}}$ under the ‘backwards’ approach suggests that values of $M_{\text{Syr}}$ higher than the upper bound of 7% permitted by its prior above are not incompatible with the other information available. Gelman et al. (1995) suggest that all plausible values for the model parameters should be assigned non-zero prior probability. One reason for this is that if the prior assigns zero probability to the value of some parameter, this value is always assigned zero probability in the posterior distribution. Punt and Butterworth (1996) argue that any prior for $M_{\text{Syr}}$ for the B-C-B bowhead stock should be viewed as ‘uninformative’, because the dissimilarities of bowheads and other baleen whale species render inferences for bowheads drawn from those other species questionable. Consequently, they advocate that the prior be chosen to be uniform and over a wider range than specified by IWC (1995). This suggestion was implicitly accepted by IWC (1998c) where a $U[1\%, 7\%]$ prior for $M_{\text{Syr}}$ (corresponding to an upper bound for $M_{\text{Syr}}$ considerably larger than 7%) is specified (see Table 2).

Punt and Butterworth (1997a) developed an approach to Bayesian analysis (‘less both’) that ignores both of the priors for $K_{1+}$ and $M_{\text{Syr}}$. It involves generating values for current population size and the current rate of increase of the population (ROI) from prior distributions, and then selecting values for $K_{1+}$ and $M_{\text{Syr}}$ to hit the values generated for $P_{1993}$ and ROI exactly. An assumption (not explicitly stated by Punt and Butterworth (1997a)) underlying this approach is that there is an ‘indirect’ prior for ROI which is $U(-\infty, \infty)$. The approach is thus equivalent to placing all of the direct information about ROI into the likelihood (for example, in the manner indicated in equation (3) following) and generating values for current population size from its prior and for ROI from $U(-\infty, \infty)$.

Punt (1999) outlines an approach for placing a prior on $\lambda_{\text{max}}$ instead of on $M_{\text{Syr}}$ when conducting a Bayesian assessment. Best (1993) provides estimates of annual increase rates at low population size for a range of severely depleted stocks of baleen whales. Ignoring the estimate for the B-C-B bowhead stock (to avoid using the abundance data

\(^8\) This is an oversimplification to better make the essential point, which relates in particular to the update of the $M_{\text{Syr}}$ prior under ‘forwards’. The reason it is not exactly correct as stated, however, is that even before the BALEEN II population model trajectories are computed, certain combinations of these biological parameters are impossible because of incompatibility with the demographics underlying the BALEEN II model, so that this aspect alone converts the independent priors into a joint distribution with some non-zero (but typically small) covariances.

\(^9\) For the purposes of simplicity of presentation, this argument has ignored the possibility of oscillatory trajectories.
for this stock twice in the analysis) and the estimate for the Eastern North Pacific gray whale (which is not currently at a small fraction of its pre-exploitation equilibrium size), and taking the lower rate of increase when more than one estimate is provided for a given population, leads to seven estimated rates of increase at low population size (Table 6). The mean of these estimated annual rates is 0.085 (SD 0.024).

Table 6

<table>
<thead>
<tr>
<th>Stock</th>
<th>Point estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>South African Right</td>
<td>0.068</td>
<td>[0.048, 0.086]</td>
</tr>
<tr>
<td>Argentine Right</td>
<td>0.073</td>
<td>[0.038, 0.108]</td>
</tr>
<tr>
<td>W. Australian Right</td>
<td>0.127</td>
<td>[0.076, 0.178]</td>
</tr>
<tr>
<td>NW Atlantic Humpback</td>
<td>0.094</td>
<td>[-0.12, 0.30]</td>
</tr>
<tr>
<td>W. Australian Humpback</td>
<td>0.088</td>
<td>[0.030, 0.146]</td>
</tr>
<tr>
<td>E. Atlantic Humpback</td>
<td>0.097</td>
<td>[0.06, 0.13]</td>
</tr>
<tr>
<td>NE Atlantic Blue</td>
<td>0.051</td>
<td>[0.026, 0.076]</td>
</tr>
</tbody>
</table>

Some account needs to be taken of the likely difference in productivity between bowheads and other baleen whales when using the information in Table 6 to develop a prior for the maximum steady rate of increase for the bowhead stock. Accordingly a range of alternative prior distributions for $\lambda_{\text{max}}$ are considered for the sensitivity tests of this paper. These prior distributions should bound most interpretations of the information.

(a) $N(0.085, 0.024^2)$ – Using the empirical distribution as summarised by a normal distribution.

(b) $U[0, 0.127]$ – A uniform distribution which covers the range of estimates and includes all non-negative values for $\lambda_{\text{max}}$ lower than the largest value in Table 6.

(c) $U[0, 0.051]$ – A uniform distribution with an upper bound equal to the lowest value in Table 6 – this reflects the perception that bowheads are among the least productive of the baleen whales.

(d) $U[0.005, 0.051]$ – A uniform distribution with an upper bound equal to the lowest value in Table 6 and a lower bound chosen to exclude the possibility of a very unproductive stock.

Note that some of the estimates in Table 6 pertain to increase rates for stocks that are probably not currently at ‘very low’ levels (e.g. Best (1993) reports that the West Australian humpback population is currently 16 – 21% of its pre-exploitation equilibrium level). Use of such estimates therefore leads to the prior being biased towards low values.

Table 7 lists results for the ‘reference case’ ‘backwards’ analysis and the four sensitivity tests that place a prior on $\lambda_{\text{max}}$ instead of on $M_{\text{SYL}}$. The results are notably sensitive to the choice of the prior for $\lambda_{\text{max}}$. This is not surprising because $\lambda_{\text{max}}$ is closely related to $M_{\text{SYL}}$, and it is well known that the results of the B-C-B bowhead assessment are sensitive to the choice of the prior (particularly the choice of its upper bound) for $M_{\text{SYL}}$. The sensitivity test which places a $N(0.082; 0.024^2)$ prior on $\lambda_{\text{max}}$ leads to more optimistic results (in terms of resource productivity) than the ‘reference case’, while the sensitivity tests which place an upper bound of 0.051 on $\lambda_{\text{max}}$ lead to less optimistic results. It is notable, however, that the lower 5%iles of the $R_Y$ and $Q_0$ distributions remain larger than 67, even for the most pessimistic assessment. The posteriors for the two most pessimistic cases suggest that the stock is most likely below $M_{\text{SYL}}$ in terms of the 1+ component of the population.

The prior distribution for $M_{\text{SYL}}$, upon which the ‘reference case’ is based was inferred from estimates of $\lambda_{\text{max}}$. The results in Table 7 suggest that considerable care needs to be taken in choosing species/stocks when constructing a prior for $M_{\text{SYL}}$ by inference because the results are very sensitive to which stocks/species are chosen.

With respect to the selection among $M_{\text{SYL}}$, $\lambda_{\text{max}}$, and $ROI$ as the parameter to choose (and with which probably to associate a uniform prior) to reflect the productivity of the resource, it should be noted that most of the estimates of $M_{\text{SYL}}$ for baleen whales that have been put forward (e.g., see summary in Butterworth and Punt, 1992) have been argued from inferences of increase rates at low population size. Such inferences depend implicitly on the values assumed for the biological parameters (such as $M_{\text{SYL}}$) (Butterworth and Best, 1990). To avoid this need, we prefer here to place a prior on $\lambda_{\text{max}}$ and let the population dynamics model make the link to $M_{\text{SYL}}$, whose quantitative relationship to $\lambda_{\text{max}}$ will vary across the ranges of values for the various biological parameters. On the other hand, bowheads have been argued to be dissimilar to other baleen whale species because of their unusually high age at maturity, which renders the defensibility of a prior on $M_{\text{SYL}}$ or $\lambda_{\text{max}}$ based on those other species somewhat questionable. Furthermore, the current replacement yield ($ROI$) is an output of the assessment of particular importance, and this is closely related to the product of the current population size and $ROI$. A uniform prior on $ROI$ would be less informative about $R_Y$ than would such a prior on $M_{\text{SYL}}$ or $\lambda_{\text{max}}$. However, there is no basis for

Table 7

Estimates of eight management-related quantities for the Bering-Chukchi-Beaufort Seas stock of bowhead whales based on the ‘backwards’ approach. The point estimates given are posterior medians, followed by posterior means in round parentheses. Posterior 90% credibility intervals are given in square parentheses. This table includes results for analyses that place a prior on $\lambda_{\text{max}}$ rather than on $M_{\text{SYL}}$.  

<table>
<thead>
<tr>
<th>Reference case</th>
<th>$R_Y$ (1998)</th>
<th>$Q_0$ (1998)</th>
<th>$P_{100}^{1000} / K_{100}$</th>
<th>$P_{175}^{1000} / K_{175}$</th>
<th>$P_{100}^{1000} / M_{\text{SYL}}^{1000}$</th>
<th>$M_{\text{SYL}}$</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>12631 (12863)</td>
<td>211 (209)</td>
<td>209 (207)</td>
<td>71.2 (70.9)</td>
<td>45.9 (46.1)</td>
<td>100.2 (99.5)</td>
<td>2.59 (2.58)</td>
<td>2.49 (2.48)</td>
</tr>
<tr>
<td>[10924 16531]</td>
<td>[141 273]</td>
<td>[136 279]</td>
<td>[35.3 90.1]</td>
<td>[37.8 57.9]</td>
<td>[76.8 122.8]</td>
<td>[1.51 3.38]</td>
<td>[1.39 3.67]</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$ – $N(0.082; 0.024^2)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12178 (12375)</td>
<td>216 (214)</td>
<td>221 (217)</td>
<td>74.1 (73.8)</td>
<td>47.2 (47.6)</td>
<td>104.8 (104.2)</td>
<td>2.82 (2.81)</td>
<td>2.72 (2.69)</td>
</tr>
<tr>
<td>[10687 15754]</td>
<td>[152 277]</td>
<td>[148 285]</td>
<td>[57.0 91.8]</td>
<td>[38.8 60.5]</td>
<td>[83.7 125.8]</td>
<td>[1.74 3.96]</td>
<td>[1.63 3.77]</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$ – $U[0, 0.127]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12631 (13109)</td>
<td>207 (203)</td>
<td>206 (201)</td>
<td>70.4 (69.8)</td>
<td>45.4 (45.8)</td>
<td>100.2 (99.2)</td>
<td>2.58 (2.52)</td>
<td>2.48 (2.42)</td>
</tr>
<tr>
<td>[10893 18579]</td>
<td>[123 272]</td>
<td>[118 276]</td>
<td>[50.4 89.4]</td>
<td>[36.9 79.7]</td>
<td>[75.6 125.3]</td>
<td>[1.27 3.81]</td>
<td>[1.16 3.67]</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$ – $U[0, 0.051]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13767 (14412)</td>
<td>193 (185)</td>
<td>183 (176)</td>
<td>64.1 (62.7)</td>
<td>42.8 (42.6)</td>
<td>90.8 (89.3)</td>
<td>2.10 (2.01)</td>
<td>2.02 (1.92)</td>
</tr>
<tr>
<td>[12365 21022]</td>
<td>[94 256]</td>
<td>[90 242]</td>
<td>[44.5 78.5]</td>
<td>[34.9 51.3]</td>
<td>[68.9 107.2]</td>
<td>[0.96 2.79]</td>
<td>[0.82 2.73]</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$ – $U[0.005, 0.051]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13755 (14326)</td>
<td>193 (187)</td>
<td>183 (177)</td>
<td>63.9 (62.9)</td>
<td>42.7 (42.6)</td>
<td>90.6 (89.3)</td>
<td>2.11 (2.02)</td>
<td>2.03 (1.93)</td>
</tr>
<tr>
<td>[12386 20129]</td>
<td>[97 254]</td>
<td>[93 241]</td>
<td>[45.1 78.6]</td>
<td>[34.9 51.4]</td>
<td>[69.1 108.1]</td>
<td>[0.99 2.79]</td>
<td>[0.85 2.71]</td>
</tr>
</tbody>
</table>
specifying a prior on ROI (which depends on the current status of the resource) so on balance we advocate placing a prior on $\lambda_{\text{max}}$.

The choice of a prior for $\lambda_{\text{max}}$ is complicated by a lack of data (see Table 6). We tentatively prefer a $\text{U}(0, 0.127)$ prior. This prior implicitly acknowledges the perception that bowheads are likely to be relatively unproductive compared to other baleen whales by including values for $\lambda_{\text{max}}$ lower than the lowest value in Table 6, but equally does not exclude higher values which are not incompatible with the data on the bowhead rate of increase.

**Natural mortality**

The derivation of the priors for $S_{\text{adult}}$ and $S_{\text{prev}}$ in 1994 (IWC, 1995) is poorly documented. The basis for the choice of the prior for $S_{\text{adult}}$ appears to be inferences from the capture of a very large animal at Wainwright in 1993 with two stone harpoons of a pattern generally out of use by the start of the 20th century, the age determination study by Nerini (1983), and estimates of natural mortality for adult right whales (IWC, 1995). IWC (1992) used values of 0.01yr$^{-1}$ and 0.02yr$^{-1}$ in Hitter-Fitter runs for the B-C-B bowhead stock primarily (it appears) because higher values were incompatible with estimates of the proportion of the population which is immature. Fig. 3 shows the marginal prior distribution assumed for $S_{\text{adult}}$ by IWC (1995), as well as the corresponding distribution for the number of further years for which 5% of the population will survive after maturing. The upper tail of this distribution (5% point = 562yrs) is clearly unrealistic. The Scientific Committee could, in its reconsideration of priors for the B-C-B bowhead assessment, perhaps consider a prior on longevity rather than on adult natural mortality because this is arguably a more ‘natural’ parameter. Any prior on longevity would exclude unrealistically long lifespans and hence (effectively) place an upper bound on $S_{\text{adult}}$. In suggesting the ‘reference’ priors in Table 2, IWC (1998c) explicitly dealt with this issue by placing an upper bound of 0.995 on $S_{\text{adult}}$ and by imposing a maximum age of 100 years.

Whitcher et al. (1996) and J.L. Laake (pers. commn) provide separate preliminary estimates of the survival rate of adult bowhead whales using data from aerial photographs of identified whales. These estimates are 0.970 (SD 0.054) [M-profile-1] and 0.995 (SD 0.055) [M-profile-2] respectively (see Fig. 4, upper panel). The prior distribution for $S_{\text{adult}}$ selected by IWC (1995) (see Fig. 4, lower panel, for its marginal$^{10}$) was not based (explicitly at least) on these data, so that this prior distribution and the aerial-photography-based distributions for the estimates can legitimately be considered to be independent. These aerial-photography-based distributions can therefore be included along with the abundance and ‘proportion’ information in the ‘direct’ data used when applying Bayesian methods. Naturally, because the Whitcher and Laake estimates are based on the same data, they cannot both be included in the same analysis.

The incorporation of these data can be achieved in two ways. The prior distribution of IWC (1995) for $S_{\text{adult}}$ can be updated using Bayes’ theorem, or this prior distribution can be left unchanged and an extra component added to the likelihood function. These two procedures will give identical results. The first is computationally more efficient and so has been applied here.

\[ 10 \text{This is a marginal distribution because IWC (1995) specified a joint pre-model distribution for } S_{\text{adult}} \text{ and } S_{\text{prev}}. \]
Butterworth (1995) raises the issue that the priors for age-at-maturity ($a_m$), natural mortality ($M$) and maximum fecundity should be correlated because, a priori, one expects that a high value for $a_m$ would be linked to low values for $M$ (e.g. Gunderson and Dygert, 1988). One of the impacts of the independence of the priors for $a_m$ and $S_{adult}$ (under the assumptions of IWC (1995)) is that there is a greater prior probability than seems reasonable that the age-at-maturity lies in the upper tail of its prior (say at the maximum of 26 years) and simultaneously that the adult survival rate lies in the lower tail of its prior (corresponding say, to an average age that a bowhead lives after maturity of only 8 years or less).

IWC (1992) notes that there is no direct evidence about $S_{juven}$. The prior selected by IWC (1995) is based on the (seemingly reasonable) assumption that juvenile natural mortality is less than adult natural mortality (although it should be noted that there are also arguments that the reverse applies for fur seals, at least, because of the stress placed on the adults by the cost of reproduction—L. Boyd, pers. comm). However, the actual prior for $S_{juven}$ is essentially arbitrary. This hardly surprising because there are precious few (if any) reliable estimates of $S_{juven}$ for baleen whales. Punt and Butterworth (1996) and Wade (1999) propose methods that avoid the need for specifying priors for both of $S_{adult}$ and $S_{juven}$. The method proposed by Wade (1999) involves generating values for $MSYR$, $MSYL$, $S_{adult}$, $a_m$ and $f_{max}$ from their priors and then calculating a value for $S_{juven}$. If the value for $S_{juven}$ is greater than the value for $S_{adult}$, the set of parameters is assigned zero likelihood. This approach was adopted by IWC (1998c), thus avoiding the need to specify a prior for $S_{juven}$ in Table 2. It also implicitly forces a relationship (and hence correlation) between $a_m$ and natural mortality and hence partially resolves the problem caused by the a priori assumption that $a_m$ and $S_{adult}$ are independent.

Including the data on $S_{adult}$ from analysis of aerial photographs of identified bowheads hardly impacts the results of the assessment. This is a consequence of the highly informative prior distribution assumed for $S_{adult}$ and the comparatively uninformative nature of the data. The lower panel of Fig. 4 shows the distributions for $S_{adult}$ obtained by updating its prior distribution using the likelihoods in the upper panel. The updated distributions are very similar to the original distribution, confirming the uninformative nature of the data given the prior assumed for $S_{adult}$.

George et al. (1998) provide estimates of age for 42 bowheads using the aspartic acid racemization technique. Four of the animals were estimated to be older than 100 years. These data should be used in future to update the prior for $S_{adult}$.

**Age-at-recruitment**

The assessment assumes that recruitment occurs at age 1 and that the historical harvest has been taken with uniform selectivity from the 1+ component of the population. This is equivalent to assuming a delta-function prior for the age-at-recruitment at age 1. However, the age-at-recruitment is likely to have changed over time because, during the early years of the commercial fishery, whalers presumably targeted large animals (IWC, 1992). Following the demise of this fishery, aboriginal exploitation targeted smaller animals (IWC, 1992). The current formulation of the Baleen II model cannot allow for changes to the age-at-recruitment explicitly. It can, however, divide catches into those from the mature component of the population and those from the recruited (in this case the 1+) component. The available data should be examined to see if an appropriate division of the historical harvests can be made between those to be assumed to be taken uniformly from the 1+ component and those taken similarly from the mature population, for a better reflection of the historical reality.

**Age-at-maturity**

The length at maturity for bowheads (averaged over both sexes) is approximately 13m (Withrow and Angliss, 1992). The prior assumed for $a_m$ by IWC (1995) is based on converting a length of 13m to age. Information from carbon isotope ageing of baleen plates suggests that bowheads reach 13m between 18 and 20 years of age (Schell et al., 1989). The prior for age-at-maturity, a normal distribution with mean 20 and standard deviation 3, constrained to lie between 13.5 and 26.5, was chosen to encompass the best estimates of 18-20 years and to incorporate a minimum value of 14 years (the lower limit for age-at-maturity obtained by Schell et al., 1989). It is therefore based primarily on 'direct' information from ageing. The prior is probably overly precise because no account was taken of uncertainty in ageing methods and of the assumption that the length at maturity is 13m. Such information would have to be reflected as 'indirect' information.

**Age of transition from juvenile to adult natural mortality rate**

The prior for the greatest age at which juvenile natural mortality is assumed to apply, $a$, is based on a suggestion by Givens et al. (1995). They argued that because nothing is known about this parameter, a discrete uniform distribution from 1-9 years would be appropriate. The upper limit for this prior was selected to be less than 10 years, the lower limit of the prior assumed by Givens et al. (1995) for $a_m$. When the Scientific Committee (IWC, 1995) selected a prior for $a_m$ which differed from that suggested by Givens et al. (1995), no changes were made to the prior for $a$. The available data are uninformative about the value for this parameter (Givens et al., 1995).

**Maximum pregnancy rate**

The prior for the maximum pregnancy rate in IWC (1995) (taken to be the maximum possible pregnancy rate by Punt and Butterworth (1996)) was modified by the Scientific Committee in 1997 (IWC, 1998c). The lower bound for $f_{max}$ of 0.25 was supported by evidence from photographically identified B-C-B bowheads. IWC (1998c) does not provide a basis for the assumption of a uniform distribution on $U_{f_{max}}$, nor for the upper bound of 0.4. IWC (1995, p.146) does, however, refer to a 3–4 year calving interval under optimal conditions, although it is not completely clear how this is intended to relate to the maximum pregnancy rate. The prior selected by IWC (1998c) is markedly more informative than that suggested by Givens et al. (1995), which had most of its mass between 0.14 and 0.5 but also had some mass between 0.5 and 1.

The BALEEN II population dynamics model assumes that density dependence acts on fecundity (de la Mare, 1989; Punt, 1999)\(^1\). It differentiates between a pregnancy rate, which is density-dependent, and a constant 0-year-old natural mortality rate that is pre-specified. It is unclear to what extent this has been considered in previous assessments.

---

\(^1\) A variant of the BALEEN II model exists which allows for the assumption that density dependence acts on natural mortality rather than fecundity (Punt, 1996), and has been applied in an assessment of the B-C-B bowhead stock (Punt and Butterworth, 1996).
The prior for this quantity is based on the arguments of Givens et al. (1995). It encompasses the range of values considered during the development of the CLA for commercial whaling. The prior chosen for \( MSYL_{\text{L}} \) (U[0.4; 0.8]) is centred on the Scientific Committee’s choice in general past practice of \( MSYL = 0.6 \). This choice was based primarily on empirical evidence (e.g. Fowler, 1981) that the per capita growth rate of large mammal populations as a function of population size has a negative second derivative -see Butterworth and Best (1994) for a more detailed historical summary. However, the evidence and justification for this view has more recently been called into question (de la Mare, 1994; IWC, 1994; MacCall and Tatsukawa, 1994). IWC (1995) specified this prior for \( MSYL_{\text{m}} \) but this was changed to \( MSYL_{\text{L}} \) in IWC (1998c), in the light of arguments reflected in IWC (1998a).

### Including the abundance data in the likelihood function

The data available for the assessment of the B-C-B bowhead stock are the estimates of the proportions of mature animals and calves in the population from 1985 to 1992 (the ‘proportion’ data) and the estimates of I+ abundance from surveys conducted at Point Barrow, Alaska (the ‘abundance’ data).

The ‘proportion’ data are included in the likelihood function by assuming that the observations are \( t \) distributed with 5 degrees of freedom (IWC, 1995). The model-estimates are taken to be average of the predicted proportions for 1988 and 1989. Only the proportions of calves and of mature animals are included in the likelihood function and these proportions are assumed to be independent of each other.

Zeh et al. (1995) present a series of estimates of abundance for the B-C-B bowhead stock based on visual and acoustic counts of bowheads off Point Barrow (reproduced here as Table 8)\(^{12}\). Bayes Empirical Bayes (B-E-B) estimates of abundance are available for 1988 and 1993 (Raftery and Zeh, 1991; 1998; Zeh et al., 1995). The B-E-B estimates are constructed from the data of the visual and acoustic surveys but also utilise prior information. For example, the B-E-B estimate for 1993 is based on a prior of 7,800 (SD 1,300) and a likelihood of approximately \( N(8,293; 626^2) \). For ease of presentation, the former will be referred to as the B-E-B prior and the latter as the B-E-B likelihood. The following discussion deals only with the B-E-B likelihood because the B-E-B prior does not comprise part of the likelihood function\(^{13}\).

Several alternative prescriptions are available to incorporate the abundance data in the likelihood function. Following past practice (IWC, 1995; 1998c), we assume that the \( N_0^P_A \) estimates in Table 8 provide information on relative abundance while the B-E-B estimate for 1993 is an estimate of absolute abundance. Prescriptions (b) and (c) examine the implications of assuming that the \( N_0^P_A \) estimates provide information on absolute abundance.

(a) The B-E-B likelihood for 1993 (or 1988) is assumed to provide information on absolute rather than relative abundance (IWC, 1992) and the \( N_0^P_A \) estimates are used to obtain an estimate of the rate of increase in I+ abundance. These two sources of information are then treated as being independent when constructing the likelihood function. The contribution of the abundance data to the negative of the logarithm of the likelihood function (excluding constants) in this case is given by:

\[
-\ln L = \frac{1}{2\sigma^2_\text{ROI}} \left( N_\text{est}^\text{BEB} - \bar{N}_y \right)^2 + \frac{1}{2} \ln \left( 1 + \frac{(\text{ROI} - \bar{\text{ROI}})^2}{\sigma^2_\text{ROI}} \right) \tag{3}
\]

where \( N_\text{est}^\text{BEB} \) is the B-E-B estimate of abundance for year \( y \) (assumed to apply to the I+ component of the population and to be equal to the mode of the likelihood at 8,293 for \( y = 1993 \)), \( \bar{N}_y \) is the estimate of the number of I+ animals at the start of year \( y \) from the population model, \( \sigma_\text{ROI}^2 \) is the standard error of \( \text{ROI}^{\text{obs}} \) (taken to be 626, the likelihood standard deviation, for \( y = 1993 \)), \( \text{ROI}^{\text{obs}} \) is the estimate of the 1978-93 rate of increase, which is assumed to have a \( t_y \) distribution (see Table 2), \( \bar{\text{ROI}} \) is an estimate of \( \text{ROI} \) based on a regression of \( \bar{N}_y : y = 1978, 79, ...93 \) on \( y \), and \( \sigma_\text{ROI} \) is the standard error of \( \text{ROI}^{\text{obs}} \).

(b) The data in Appendix A are assumed to provide independent estimates of absolute abundance; in terms of this assumption, the B-E-B likelihood would provide exact duplicate information to that already contained in the corresponding entry in the survey series, and hence is ignored. The contribution of the abundance data to the negative of the logarithm of the likelihood function (excluding constants) in this case is given by\(^{14}\):

\[
-\ln L = \frac{1}{2\sigma^2_\text{ROI}} \left( N_\text{est}^\text{BEB} - \bar{N}_y \right)^2 \tag{4}
\]

where \( N_\text{est}^\text{BEB} \) is the estimate of abundance for year \( y \) based on the survey and acoustic data

\(^{12}\) ‘Additional variance’ (IWC, 1997) will be ignored because Zeh et al. (1995) and Cooke (1996) report that these abundance estimates and their CIs are consistent with the assumption that any such variance is of negligible size.

\(^{13}\) Except when applying the ‘forwards’ approach to Bayesian Synthesis, for which the analysis is subject to the Borel paradox.

\(^{14}\) The choice of log-normal error is based on a suggestion by Buckland (1992).
(c) As for (b), except that account is taken of the covariance among the estimates of abundance (see Appendix A). The contribution of the abundance data to the negative of the logarithm of the likelihood function (excluding constants) is then given by (in vector-matrix notation):

\[ -lnL = \frac{1}{2} \left( tn_{survey} - \hat{tn} \right)^T \left( \Sigma_{survey} \right)^{-1} \left( tn_{survey} - \hat{tn} \right) \] (Equation 5)

where \( \Sigma_{survey} \) is the variance-covariance matrix for the logarithms of the survey estimates.

(d) The survey data are assumed to provide independent indices of relative abundance, with the B-E-B likelihood ignored for the same reason as in (b). The contribution of the abundance data to the negative of the logarithm of the likelihood function (excluding constants) in this case is:

\[ -lnL = \sum_1 \left( \frac{1}{2 \sigma^2_{survey}} \left( tn_{survey} - \hat{tn} \right)^2 \right) \] (Equation 6)

where \( b \) is the survey bias.

For this case, it is necessary to specify a prior for \( b \). In the absence of information about \( b \), an uninformative prior \( \ln b \sim U[-\infty, \infty] \) is assumed for this parameter.

(e) As for (d), except that account is taken of the covariance among the estimates of abundance. The contribution of the abundance data to the negative of the logarithm of the likelihood function (excluding constants) in this case is given by:

\[ -lnL = \frac{1}{2} \left( tn_{survey} - \hat{tn} \right)^T \left( \Sigma_{survey} \right)^{-1} \left( tn_{survey} - \hat{tn} \right) \] (Equation 7)

(f) The likelihood for the 1993 B-E-B estimate and the survey-based estimate for 1993 provide an estimate of the survey bias factor \( b \) of 0.936 (CV = 0.026)\(^{15} \). Following Butterworth et al. (1999) and Butterworth and Punt (1992), this estimate can be incorporated into the likelihood function as independent information about \( b \). For the case in which the covariance among the estimates is ignored, this leads to the following negative log-likelihood:

\[ -lnL = \frac{1}{2 \sigma^2_b} \left( \ln b - \hat{b} \right)^2 + \sum_1 \frac{1}{2 \sigma^2_{survey}} \left( tn_{survey} - \hat{tn} \right)^2 \] (Equation 8)

where \( \hat{b} \) is the estimate of the survey bias factor (0.936) and \( \sigma_b \) is the CV of \( \hat{b} \) (0.026).

(g) As for (f), except that account is taken of the covariance among the estimates of abundance. The contribution of the abundance data to the negative of the logarithm of the likelihood function (excluding constants) in this case is given by:

\[ -lnL = \frac{1}{2 \sigma^2_b} \left( \ln b - \hat{b} \right)^2 + \frac{1}{2} \left( tn_{survey} - \hat{tn} \right)^T \left( \Sigma_{survey} \right)^{-1} \left( tn_{survey} - \hat{tn} \right) \] (Equation 9)

\(^{15}\) This CV follows from the CVs for the \( N_{P1} \) and B-E-B estimates for 1993 assuming that these estimates are uncorrelated.

Approach (a) above forms part of the 'reference case', while approaches (b)-(g) provide increasingly sophisticated treatments of the data. Equation (9) provides the most complete treatment of the data as it assumes: that the data in Appendix A provide an index of relative rather than absolute abundance; that those estimates are correlated; and that the likelihood for the B-E-B estimate provides information on absolute abundance. The likelihood for the 1993 B-E-B estimate is not explicitly included in Equation (9) as much of the information underlying this likelihood is already included in the survey estimate for 1993. Zeh and Givens (1997) illustrate that including both the likelihood of the B-E-B estimate and the information corresponding to the data on trend in an analysis can lead to severely biased estimates of quantities of importance to management. If data were available on the likelihood for the 1988 B-E-B estimate of abundance, Equation (9) could be extended by including a second term related to the associated estimate of survey bias.

Table 9 presents results for analyses based on the 'backwards' approach. Results are shown in this table for the 'reference' method for incorporating the abundance data in the likelihood function (Equation 3) and six alternative methods (see Equations 4 to 9). The results based on Equations 4-9 indicate a slightly less productive population and hence lower values for RY (1998) and Qo (1998). It should be noted that the abundance estimates for these latter analyses are not identical to those upon which ROI is based (contrast the estimates in Tables 8 and A.1). However, this is not the only reason for the differences in Table 9 because Punt and Butterworth (1996) show that incorporating the \( N_{P}/P_d \) data in Table 8 into the assessment as absolute indices of abundance (cf. Equation 4) also leads to less optimistic results.

The results become slightly less optimistic if account is taken of the correlation among the estimates of abundance. Treating the abundance estimates as relative (Equations 6 and 7) rather than as absolute indices of abundance (Equations 4 and 5) or including a prior for the bias factor (Equations 8 and 9) increases the widths of the 90% credibility intervals slightly. However, the posterior means and medians are not impacted markedly by this change. Including a prior on the bias factor (our preferred approach) leads to results that are intermediate in terms of the widths of the 90% credibility intervals between those which treat all of the abundance estimates as absolute and those which treat all of the abundance data as relative. The posterior medians for Slope, RY (1998) and Qo (1998) for our preferred approach are also intermediate.

CONDUCTING THE POPULATION PROJECTION FOR RECENT YEARS ONLY

Assessments of the B-C-B bowhead stock have been conducted under the assumptions that, at the start of the catch series (1848), the population was at pre-exploitation equilibrium and that the carrying capacity of the bowhead population has not changed over time. An alternative to this set of assumptions is to assume instead that the population had a stable age-structure in some more recent year (see Punt (1999) for details of how this is implemented for the Baleen II model). The assessments of the Eastern North Pacific stock of gray whales are based on this latter assumption (Punt and Butterworth, 1997b; Wade, 1997; 1999).

One arguable advantage of this approach to conducting assessments of the B-C-B bowhead stock is that it becomes possible to place a (joint) prior distribution on \( K_1+ \), and the \( g_1+g_2 \).
Table 9  
Estimates of eight management-related quantities for the Bering-Chukchi-Beaufort Seas stock of bowhead whales based on the ‘backwards’ approach. The point estimates given are posterior means, followed by posterior means in round parentheses. Posterior 90% credibility intervals are given in square parentheses. The analyses in this table differ in how the abundance data are included in the likelihood function.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Reference case</th>
<th>Post-model-pre-data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10  
Estimates of eight management-related quantities for the Bering-Chukchi-Beaufort Seas stock of bowhead whales based on the ‘backwards’ approach. The point estimates given are posterior means, followed by posterior means in round parentheses. Posterior 90% credibility intervals are given in square parentheses. This table includes results for analyses that start the population trajectories from a more recent year (y_t) rather than from an assumed pre-exploitation equilibrium in 1848.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Reference case</th>
<th>y_t = 1930</th>
<th>y_t = 1950</th>
<th>Max. K_t = 50,000</th>
<th>y_t = 1960</th>
<th>y_t = 1970</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

population size in 1993, P_{1993}. Thus, the problem of having to choose between the ‘backwards’ and ‘forwards’ approaches is eliminated. Other advantageous aspects are that results are no longer dependent on values for catches during the early period of the fishery, which has had to be estimated in the absence of specific records, and that the possibility of regime shifts (tamountant to changes in K over time) is admitted. Four alternative choices for the first year in the analysis, y_1, are considered (1930, 1950, 1960, and 1970). The prior for K_{194} is taken to be that for the ‘reference case’, f(K_{194}) = U[fn7000, fn3100]. The sensitivity of the results to the choice for the upper end of the prior for K_{194} is explored by changing it from 31,000 to 25,000 and to 50,000.

Table 10 lists results for six sensitivity tests that do not change the population projections from deterministic pre-exploitation equilibrium in 1848, but assume instead that the population had a stable age-structure in some more recent year. For comparability with the ‘reference case’, the sensitivity tests are based on the ‘reference’ likelihood of Equation 3. These results are not particularly sensitive to the choice of y_t. However, there are notable differences between the results for the ‘reference case’ and those for the six sensitivity tests. For example, the posterior distributions for Q_{0} (1998) and RY (1998) have longer tails at high values and the posterior for the pre-exploitation size does not differ much from its prior. One consequence of the latter result is that the depletion- and population size-related results depend strongly on the upper bound of the prior assumed for K_{194}. For example, the posterior median for P_{1993}/K_{194} drops from 56.5 to 39.6% (as the upper bound for K_{194} is increased from 25,000 to 50,000). In contrast, the posterior medians for RY (1998), Q_{0} (1998), Slope, and MSTR_{194} do not depend notably on the upper bound for K_{194}.

Thus, not unexpectedly, dropping the assumption that the population was at its pre-exploitation level in 1848 (and that all the historical catches are known exactly) leads to much wider 90% credibility intervals for all quantities except Slope and MSTR_{194}. This effect is perhaps most notable for RY (1998) for which the 90% credibility interval is (roughly) [125, 400] for the sensitivity tests compared to [141, 273] for the ‘reference case’. It is perhaps notable that the lower 5% for Q_{0} (1998) hardly differs among the analyses.  

---

16 This needs to be a joint prior distribution because the a priori constraint that P_{1993} < K must be imposed (there being negligible probability that possible oscillatory behaviour of the population trajectory as a result of time-lags in the dynamics could see the population above K in recent years).
although the upper 95%ile differs markedly between the ‘reference case’ and the sensitivity tests. The posterior median for the Slope statistic for the sensitivity tests is slightly closer to the ‘observed’ value in Table 2.

OUR PREFERRED ANALYSIS

The advantages of starting the population projections in some more recent year are that it is not necessary to assume that the carrying capacity of the bowhead population has remained unchanged over the last 150 years and that all the historical catches are known exactly. We believe that analyses should be conducted for both options, i.e. assuming that the population was at its pre-exploitation level in 1848 and assuming that it had a stable age-structure in some more recent year. Results are presented for the case \( y_t = 1950 \) for the analyses that start the projections in a recent year as such results are not very sensitive to the choice for \( y_t \). The results for this assessment are shown for the ‘reference case’ prior for \( K_t \). This choice is essentially arbitrary, so little confidence can be placed in the results for the depletion-and-population size-related quantities because the posteriors for these quantities depend critically on the choice of a prior distribution for \( K_t \) which is here updated by the data (Table 10). However, Table 10 does also show that the key management-related quantities \( RY(1998) \) and \( Q_t(1998) \) are relatively insensitive to variations in the specification of a prior for \( K_t \), so that this approach retains utility.

We prefer the ‘backwards’ to the ‘forwards’ approach for three main reasons: the simulation test results show a clear preference for ‘backwards’, any assessment based on ‘forwards’ which also places a prior on the abundance in a recent year is subject to the Borel paradox, and ‘forwards’ updates the prior for \( MSTRy \) before any data are included in the assessment. As noted above, although the full pooling approach of Raftery and Poole (1997) and Poole and Raftery (1998) resolves the Borel paradox, in doing so it introduces a new problem of how to choose pooling weights (i.e. how to specify their relative reliabilities of the priors for \( K_t \) and \( P_{1993} \). \( \lambda_{\text{max}} \) (the maximum rate of population increase, which occurs at low population size) is our preferred choice for the productivity parameter, because nearly all the data available from other whale species upon which to base a productivity prior constitute observations of this quantity. We tentatively suggest a \([0, 0.127] \) prior for \( \lambda_{\text{max}} \) for reasons discussed above.

We prefer the likelihood defined by Equation (9) because it treats the abundance estimates in Table A.1 as relative indices of abundance and because it incorporates the (independent) information about the bias factor explicitly in the likelihood. Equation (9) is preferred to Equation (8) because it takes account of the correlation among the abundance estimates.

Table 11 lists results for the ‘reference case’ ‘backwards’ analysis and the two ‘preferred’ variants. In addition to providing results for the eight quantities listed above, results are also shown for \( S_{1UV}, S_{2UV}, a_r, f_{\text{max}}, P_{1993}, MSTRy/K_t, MSTRy_{\text{max}} \) and \( \lambda_{\text{max}} \). The results for \( MSTRy/K_t, MSTRy_{\text{max}} \) and \( \lambda_{\text{max}} \) are presented as percentages. As expected from the results of Table 9 which show that the use of the Equation 9 likelihood leads to notably lower estimates of productivity, the results for the ‘preferred’ analyses are less optimistic than those for the ‘reference case’. The lesser productivity of the reference is reflected by lower posterior medians for \( MSTRy \) and \( \lambda_{\text{max}} \). This appears to be a reflection primarily of lower values for \( S_{1UV} \) and \( MSTRy/K_t \) - the posterior medians for which drop from 0.94 and 0.72 respectively for the ‘reference case’ to 0.93 and 0.70/0.69 for the two ‘preferred’ variants. Although the results for the population size- and depletion-related quantities from the \( y_t = 1950 \) variant are unreliable owing to their sensitivity to the choice of a prior for \( K_t \), the posteriors for the biological parameters are remarkably similar for the two preferred analyses. The lower 5%iles for \( RY(1998) \) and \( Q_t(1998) \), although insensitive to the choice of \( y_t \), are notably lower than those for the ‘reference case’ (only about 80 compared to some 140). This is a consequence of a smaller current population size and lower productivity. However, these lower 5%iles are all larger than the current annual strike limit for the B-C-B stock of 67.

<table>
<thead>
<tr>
<th>Table 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates of the eight management-related quantities considered previously and eight further biological parameters/variables for the Bering-Chukchi-Beaufort Seas stock of bowhead whales based on the ‘backwards’ approach. The point estimates given are posterior medians, followed by posterior means in round parentheses. Posterior 90% credibility intervals are given in square parentheses. Results are shown in this table for the reference case and the two ‘preferred’ analyses.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantity</th>
<th>'Backwards' reference case</th>
<th>Preferred - ( y_t=1848 )</th>
<th>Preferred - ( y_t=1950 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_t )</td>
<td>12,631 (12,863)</td>
<td>13,393 (13,193)</td>
<td>14,126 (13,860)</td>
</tr>
<tr>
<td>( RY(1998) )</td>
<td>211 (209)</td>
<td>207 (206)</td>
<td>207 (206)</td>
</tr>
<tr>
<td>( Q_t(1998) )</td>
<td>209 (207)</td>
<td>189 (183)</td>
<td>189 (183)</td>
</tr>
<tr>
<td>( P_{1993}/K_t )</td>
<td>71.2 (70.9)</td>
<td>65.2 (64.4)</td>
<td>46.0 (45.2)</td>
</tr>
<tr>
<td>( P_{1993}/MSTRy )</td>
<td>45.9 (46.1)</td>
<td>42.9 (43.2)</td>
<td>32.6 (32.6)</td>
</tr>
<tr>
<td>( MSTRy )</td>
<td>2.59 (2.58)</td>
<td>2.27 (2.25)</td>
<td>2.22 (2.21)</td>
</tr>
<tr>
<td>Slope</td>
<td>2.49 (2.48)</td>
<td>2.19 (2.15)</td>
<td>2.24 (2.23)</td>
</tr>
<tr>
<td>( MSTRy/K_t )</td>
<td>0.72 (0.71)</td>
<td>0.70 (0.69)</td>
<td>0.69 (0.69)</td>
</tr>
<tr>
<td>( S_{1UV} )</td>
<td>0.937 (0.926)</td>
<td>0.928 (0.917)</td>
<td>0.928 (0.917)</td>
</tr>
<tr>
<td>( S_{2UV} )</td>
<td>0.988 (0.987)</td>
<td>0.986 (0.985)</td>
<td>0.986 (0.985)</td>
</tr>
<tr>
<td>( a_r )</td>
<td>0.31 (0.31)</td>
<td>0.31 (0.32)</td>
<td>0.31 (0.31)</td>
</tr>
<tr>
<td>( \lambda_{\text{max}} )</td>
<td>5.14 (5.16)</td>
<td>4.74 (4.78)</td>
<td>4.77 (4.77)</td>
</tr>
<tr>
<td>( P_{\text{max}} )</td>
<td>8,212 (8,195)</td>
<td>7,984 (7,988)</td>
<td>8,009 (8,016)</td>
</tr>
<tr>
<td>( MSTRy_{\text{max}} )</td>
<td>4.86 (4.89)</td>
<td>4.21 (4.22)</td>
<td>4.12 (4.14)</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

Judy Zeh (Dept. of Statistics, University of Washington, Seattle) and Jeff Laake (NMFS, Seattle) are thanked for providing the details of the aerial-photography-based distributions for adult survival rate as is Jeff Breiwick (NMML, Seattle) who supplied the catch data. Judy Zeh, Ann Cowling and an anonymous reviewer are thanked for their comments on an earlier version of this paper. Discussions with several colleagues at CSIRO Division of Marine Research are gratefully acknowledged.

REFERENCES


Appendix A  
CONSTRUCTING INDICES OF ABUNDANCE FOR THE B-C-B BOWHEAD STOCK

The data available for the construction of indices of abundance for the B-C-B bowhead stock include estimates of 
\( N_y \) (the observed number of bowheads passing within visual range) for ten years and estimates of \( P_y \) (the observed proportion passing within visual range) for five of these years (Table 8). Previous analyses of these data (e.g. Raftery et al. (1995b); Zeh et al. (1995)) have assumed that the proportion passing within visual range for the years for which estimates of \( P_y \) are not available is equal to the average over the years for which such estimates are available. Cooke (1996) criticised the methods used by Raftery et al. (1995b) and Zeh et al. (1995) because they treat the observation and process errors in an ad hoc manner. He proposed the following statistical model for the analysis of the data in Table 8:

\[
\begin{align*}
N_{y} & = P_{y} p_{y} e^{\nu_{y}} & \nu_{y} & \sim N(0; \sigma_{\nu_{y}}^{2}) \\
N_{y} & = P_{y} e^{\nu_{y}} & \nu_{y} & \sim N(0; \sigma_{\nu_{y}}^{2}) \\
\pi & = \pi_{0} e^{t} & \pi_{0} & \sim N(0; \sigma_{\pi_{0}}^{2})
\end{align*}
\]

(A.1)  

(A.2)  

(A.3)

where \( N_{y} \) is the estimate of \( N \) for year \( y \),  
\( P_{y} \) is the estimate of \( P \) for year \( y \),  
\( p_{y} \) is the size of the 1+ component of the population in year \( y \),  
\( \pi \) is the proportion of the 1+ component of the population within visual range in year \( y \),  
\( \sigma_{\nu_{y}}^{2} \) is the variance of the logarithm of \( N_{y} \) (reflecting observation error),  
\( \sigma_{\nu_{y}}^{2} \) is the variance of the logarithm of \( P_{y} \) (reflecting observation error), and  
\( \sigma_{\pi_{0}}^{2} \) is the process error variance.

The parameters of this model are \( P_{y} \) and \( p_{y} \) for each of the ten years, \( \pi \), and \( \sigma_{\pi_{0}}^{2} \). The estimates of the 22 parameters of this model are obtained by maximum likelihood. This involves finding the values for the parameters that minimise the following negative log-likelihood (after removal of constants):

\[
-\ln L = \sum_{y} \left( \frac{(\ln N_{y} - \ln(P_{y} p_{y}))^{2}}{2\sigma_{\nu_{y}}^{2}} + \sum_{y} \frac{(\ln P_{y} - \ln(\pi))^{2}}{2\sigma_{\nu_{y}}^{2}} + \frac{1}{2\sigma_{\pi_{0}}^{2}} \ln |I + \sigma_{U}| \right)
\]

where \( U \) is the reduced information matrix (which accounts for the impact of the estimation of the nuisance parameters - see Seber and Wild, 1989).

The first and third summations are over all ten years and the second summation is over the five years for which estimates of \( P_{y} \) are available. The maximum likelihood estimate of \( p_{y} \) for those years for which estimates of \( P_{y} \) are not available is \( \pi \).

Table A.1 lists the estimates of the \( P_{y} \), the estimates of the asymptotic standard errors of \( \ln P_{y} \) and the correlation matrix for \( \ln P_{y} \), that result from the implementation of Cooke’s approach.

<table>
<thead>
<tr>
<th>Year</th>
<th>Estimate</th>
<th>CV</th>
<th>Correlation matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>4,820</td>
<td>0.273</td>
<td>1.000</td>
</tr>
<tr>
<td>1980</td>
<td>3,900</td>
<td>0.314</td>
<td>0.166</td>
</tr>
<tr>
<td>1981</td>
<td>4,389</td>
<td>0.253</td>
<td>0.054</td>
</tr>
<tr>
<td>1982</td>
<td>6,572</td>
<td>0.311</td>
<td>0.168</td>
</tr>
<tr>
<td>1983</td>
<td>6,268</td>
<td>0.321</td>
<td>0.163</td>
</tr>
<tr>
<td>1985</td>
<td>5,132</td>
<td>0.269</td>
<td>0.126</td>
</tr>
<tr>
<td>1986</td>
<td>7,251</td>
<td>0.186</td>
<td>0.080</td>
</tr>
<tr>
<td>1987</td>
<td>5,151</td>
<td>0.298</td>
<td>0.175</td>
</tr>
<tr>
<td>1988</td>
<td>6,609</td>
<td>0.113</td>
<td>0.038</td>
</tr>
<tr>
<td>1993</td>
<td>7,778</td>
<td>0.071</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Table A.1  
Estimates, CVs (actually standard errors of the logarithms), and the correlation matrix for the indices of abundance for the Bering-Chukchi-Beaufort bowhead stock. Values are based on the estimation procedure described in Appendix A.