Strike Limit Algorithm optimisation: a realistic example

GEOF H. GIVENS

Department of Statistics, Colorado State University, Fort Collins, CO 80523 USA
Contact e-mail: geof@lamar.colostate.edu

ABSTRACT

This paper illustrates a process for finding an improved variant of an aboriginal whaling management procedure *Strike Limit Algorithm* (*SLA*), applying the merging and optimisation approach of Givens (1997; 1999b). A modified version of the *SLA* developed by Punt and Butterworth (1997) was chosen as the procedure to be optimised for management of the Bering-Chukchi-Beaufort Seas stock of bowhead whales. The optimisation considers functions of the catch limit and other outputs from the nominal *SLA*, along with outputs from two other *SLA*s and estimates of certain population dynamics parameters. The result reduced the Bayes risk by over 90%, compared to the nominal procedure, and improved simulated *SLA* performance by usually allowing more strikes at less depletion risk. Such results suggest that this approach may be attractive in the general development of wildlife management procedures.

KEYWORDS: MANAGEMENT PROCEDURE; MODELLING; WHALING – ABORIGINAL; ARCTIC; BOWHEAD WHALE

INTRODUCTION

International Whaling Commission (IWC) management of aboriginal subsistence whaling will eventually rely on an aboriginal whaling management procedure (AWMP) chosen from a collection of candidate procedures after extensive simulation testing (e.g. see Donovan, 1999; IWC, 2000). An AWMP is a fully automatic algorithm designed to operate on the results of an assessment (i.e. a statistical estimation problem relying on sparse series of whale abundance data), and to produce a catch limit in each year of real (or simulated) management. The only inputs to an AWMP which vary over time are: (i) stock abundance estimates and corresponding coefficients of variation; (ii) point and variance estimates for the proportion of the stock that belongs to certain age or size classes; (iii) a number of whales 'needed' to be caught for each year up to the present; and (iv) numbers of whales actually caught during past years. As time progresses, new data on these quantities become available.

'Need' is usually expressed as a number of whales and is set periodically, often in 4-5 year blocks, by the IWC on the basis of 'need statements' submitted by individual IWC member nations. Evaluation of such need statements is a political rather than a scientific process. Nations requesting aboriginal hunting catch limits provide information about the operation and history of the aboriginal whaling fishery and their estimation of present-day nutritional and cultural needs to justify their requests. Thus, the setting of 'need' is the function of the political body of the IWC. However, that body instructs its Scientific Committee to assess whether a whale stock can safely sustain the established 'need' level. Since future 'need' is unknown, the IWC has asked the Scientific Committee to ensure that candidate AWMPs perform adequately within a range of potential future need trajectories. By explicitly providing this range, the IWC has enabled AWMP developers to avoid a complex and probably futile scientific debate over predicting future need and to instead focus on ensuring need satisfaction (IWC, 1999).

Although potential AWMPs may employ sophisticated modelling and estimation strategies, an AWMP can also be a completely arbitrary procedure (i.e. a 'black box'); ultimately, an AWMP will be judged primarily on how it performs in simulation. An AWMP should, as far as possible, meet the potentially conflicting IWC objectives of low chance of population extinction or severe depletion, high satisfaction of needed catch and high rate of population recovery (IWC, 1999). The key component of an AWMP is the *Strike Limit Algorithm* (*SLA*) that calculates catch limits from available data.

AWMPs are tested through extensive computer simulation. A population dynamics model is used to project a whale stock from a specific date in history (usually associated with pre-exploitation) to 100 years in the future. All available past data about the stock are available to the AWMP, as are simulated future abundance survey and other data, and future need. Simulations are repeated for a vast collection of scenarios that vary assumptions about whale biology and dynamics, hunting and the environment. Each scenario is replicated 100 times (this is termed a 'trial') so that different random future data sequences are encountered by the AWMP for each scenario.

An analogous management procedure (the 'Revised Management Procedure' – RMP) for commercial whaling has already been developed (e.g. IWC, 1994). One strength of the RMP development process was the concentrated effort on whale population dynamics and assessment models. This resulted in rigorous simulation comparison of five competing commercial whaling management procedures, many of which employed such models (e.g. see IWC, 1992).

A major difference in the development of the AWMP from the RMP is that the focus is on a case-specific rather than a generic approach, partially in recognition of the accomplishments of the RMP development process, but largely because aboriginal subsistence whaling is limited to a few stocks and areas with recognised 'need' but with quite different levels of biological knowledge (IWC, 1999). It is not enough for an AWMP to have the best average performance across implementations. Rather, performance must be optimised individually for each stock so that in each case aboriginal need is met to the greatest extent possible subject to risk and recovery limitations.

One idea proposed to address this stock-specific approach was a framework for embedding the development and selection of one of a small number of candidate *SLA*s in a larger statistical estimation problem. From this idea has grown a number of techniques (Givens, 1997; 1998; 1999a; b; c; Givens and Bernstein, 1998; Givens *et al.*, 1999):

(1) *H-optimisation*: a method for enhancing a nominal *SLA* to improve its performance.
(2) *H-tuning for equivalence*: a method for equivalencing several competing *SLA*s so that their performance may be fairly compared.
(3) *Merging*: a method for the optimal mathematical combination of several competing *SLA*s to produce strike limits that better meet management goals.

These terms are used by the IWC Scientific Committee Standing Working Group on the Development of an AWMP (IWC, 1999). In statistical terminology, all three methods amount to choosing a Bayes rule estimator within some class.

Much of the work with these methods has used relatively simple *SLA*s, limited trials, abbreviated result summaries, or abstracted examples to investigate the methods and illustrate their features. In this paper, a realistic application is presented: the development of an optimised *SLA* for the set of scenarios designed by the Scientific Committee (IWC, 1999) to resemble management of the Bering-Chukchi-Beaufort Seas stock of bowhead whales. The scope of this application is similar to the approach that the IWC Scientific Committee might take if it were to attempt a final optimisation or merging of candidate *SLA*s for this stock. The analysis that follows is the first comprehensive and realistic test of the optimisation and merging approach.

Terminology and labels

The AWMP development process has resulted in a long list of esoteric terms and labels; these are defined in appendix 2 of IWC (1999).

The term *Initial Exploration Trial* is relevant here, referring to a case-specific set of simulation assumptions used to test an AWMP *SLA*. In this paper, the 'fishery type 2' *Initial Exploration Trials* will be used. 'Fishery type 2' refers to a case where there is a relatively large amount of available information and the existing IWC guidelines for the management of aboriginal whaling have largely been met, such as the Bering-Chukchi-Beaufort Seas stock of bowhead whales. There are six type-2 trials; the assumptions of each are given in appendix 3 of IWC (1999). In this paper, only the trials denoted B3 and B7 will be used. Trial B7 is one of the most pessimistic trials, assuming very low productivity, a low recent stock abundance and high increasing 'need'. Trial B3 is one of the most optimistic, assuming a high productivity rate and high recent abundance.

The IWC Scientific Committee will ultimately judge the *SLA*s by assessing how they accomplish depletion avoidance, need satisfaction and stock recovery for the management of a simulated stock across 100 replicate trial simulations of 100 years length, over a wide variety of trials. No formulaic combination of these criteria is intended. Two very important statistics used by the Scientific Committee and later in this paper are final depletion (denoted $D1_{1+}$) and total need satisfaction (denoted $N1$). Final depletion is the ratio of final abundance of whales aged 1 or older after 100 years of simulated management to the number of whales aged 1 or older in an unexploited, equilibrium population. Total need satisfaction is the total number of permitted harpoon strikes divided by the total 'needed' strikes over 100 years of simulated management. Across replications, these statistics have probability distributions; percentiles such as the 5% and 50% points are usually reported. The precise definitions of all performance evaluation statistics used by the Scientific Committee are given in appendix 3 of IWC (1999).

Many *SLA*s rely on terms in common use by the Scientific Committee to refer to notions related to density-dependent population dynamics including maximum sustainable yield (MSY), MSY level (MSYL), MSY productivity rate (MSYR), and replacement yield (RY). Hereafter, these terms should be interpreted as referring to the population component aged 1 year or older, unless otherwise specified.

Brief review of H-optimisation and merging

The optimisation and merging approach empirically adjusts an *SLA* by estimating a parameterised function of the nominal *SLA* outputs. The parameter estimation is set up in a manner for which the solution is an admissible Bayes rule, and hence has certain statistically desirable properties. The resulting optimal catch limits are functions of the output from one or more nominal *SLA*s. An informal review of this approach follows; a more formal presentation of the approach is given by Givens (1997; 1999b).

Let $\boldsymbol{\theta}$ denote the values of unknown parameters such as MSYR and carrying capacity (K) which constitute the assumptions of a particular *Initial Exploration Trial*. At a particular point in time, in the ideal situation where $\boldsymbol{\theta}$ is known, let an idealised strike limit be denoted $H(\boldsymbol{\theta})$ and let N represent aboriginal need at this time, both expressed as a number of whales.

H is an artificial construct used to shape or improve *SLA* simulation performance that may be discarded after use so that the IWC Scientific Committee may evaluate *SLA*s on whatever basis it desires. The Scientific Committee's current suggestion for H is given by IWC (1999) as the minimum of aboriginal need and the quantity $H^*(\boldsymbol{\theta})$ defined at time t as:

$$H^*(\boldsymbol{\theta}) = \begin{cases} 0.9\text{MSY} & \text{if } P_t/K \geq \text{MSYL} \\ 0.8\text{RY} & \text{if } 2000/K \leqslant P_t/K < \text{MSYL} \\ 0 & \text{if } P_t < \text{MSYL} \end{cases} \qquad (1)$$

In application, the idealised strike limits, $H(\boldsymbol{\theta})$, are never known because $\boldsymbol{\theta}$ is unknown. However, the Scientific Committee would be thrilled to obtain an *SLA* with the performance characteristics of $H(\boldsymbol{\theta})$. Therefore, $H(\theta)$ can be used to obtain a simple performance goal that circumvents individual consideration of dozens of univariate performance statistics.

Given data $\boldsymbol{X}$ available at this time point and arising from likelihood function $\ell(\boldsymbol{X}|\boldsymbol{\theta})$, the strike limit calculated by a candidate *SLA* at this time is denoted Q. Usually $\boldsymbol{X}$ would include a series of past abundance estimates and a catch history. Although one hopes to improve strike limits by using H-optimisation or merging, this *SLA* could be used as is; it is therefore called a *nominal SLA*. This *SLA* may rely upon some tuning parameters whose values are chosen by the developer, or for which several alternative values are used by the person doing the optimisation, as shown later in this paper.

Note that subscripts for time, trial scenario and trial replication number have been omitted from $H(\boldsymbol{\theta})$, N, $\boldsymbol{X}$ and Q here. The suppression of nuisance subscripts is continued hereafter where possible.

If a developer wished to improve the performance of one or more nominal *SLA*s on a variety of trials to which s/he assigned weights $p(\boldsymbol{\theta})$, the H-optimisation or merging approach would be to estimate ideal strike limits using a set of predictor variables derived from the nominal *SLA*(s). This estimation proceeds as follows:

(1) Identify the ideal strike limits, $H(\boldsymbol{\theta})$. For a collection of possible values of $\boldsymbol{\theta}$, calculate $H(\boldsymbol{\theta})$ over a sample of replicate datasets ($\boldsymbol{X}$) and time points. Recall that *SLA*s are tested via computer simulation. Thus, this step consists of simulating a replicated collection of trial scenarios and recording the ideal catch limits as each simulation progresses. The values of $H(\boldsymbol{\theta})$ are actually known here because specification of a trial scenario entails specification of $\boldsymbol{\theta}$.

(2) Identify and observe potentially useful predictor variables, $\boldsymbol{Y}_0$. To predict $H(\boldsymbol{\theta})$, the obvious predictor variables to employ would be strike limits from nominal *SLA*(s). If more than one nominal *SLA* were used, these *SLA*s might be different tunings of the same procedure or one or more tunings of different procedures. With more than one nominal *SLA*, the H-optimisation process has been called merging because it extracts and combines the best information from several candidate *SLA*s. Other predictor variables might include intermediate calculations of nominal *SLA*s, particularly estimates of interesting biological parameters that are internal to the *SLA*s.

Denote the values of the complete collection of potential predictors at a particular time point as $\mathbf{Y}_0 = \{Y_1,\ldots, Y_p\}$.

In practice, the $\mathbf{Y}_0$ and $H(\boldsymbol{\theta})$ can be collected simultaneously during the simulation of a replicated collection of trial scenarios.

(3) Select predictors and a model class for optimisation. Of the p potential predictors (which may include interactions, polynomial terms, etc.) a subset may be selected for use in a model to predict $H(\boldsymbol{\theta})$. Suppose the selected predictors, $\mathbf{Y}$ are a subset of the potential predictors. The choice of $\mathbf{Y}$ is a statistical model selection problem: which predictors are believed to be most useful for predicting $H(\boldsymbol{\theta})$? This step is analogous to selecting the predictors in a linear regression model before solving for the optimal estimated regression coefficients.

Let $g(\mathbf{Y}, \boldsymbol{\alpha})$ represent a way to combine the selected Y_i to form a strike limit – an example for three predictors might be:

$$g(\mathbf{Y},\boldsymbol{\alpha}) = \min\left(N, \max\left(0, \alpha_0 + \alpha_1 Y_1 + \alpha_2 Y_2 + \alpha_3 Y_1^2 + \alpha_4 Y_3\right)\right) \quad (2)$$

where $\boldsymbol{\alpha} = \{\alpha_0,\ldots, \alpha_3\}$ and, for example, $Y_3 = Y_1 Y_2$. The α_i are parameters to be chosen through the optimisation process.

(4) Define a Bayesian estimation context. In order to establish an optimality criterion for fitting models like equation (2), we must pose the solution of $\boldsymbol{\alpha}$ as an estimation problem. This requires choosing a weighting, $p(\boldsymbol{\theta})$, of the trials. Also define a loss function to characterise the penalty accrued if $g(\mathbf{Y}, \boldsymbol{\alpha})$ does not equal the ideal strike limit. Denote the loss $L(g(\mathbf{Y}, \boldsymbol{\alpha}), H(\boldsymbol{\theta}))$. A typical choice would be squared error loss, namely $L(g(\mathbf{Y},\boldsymbol{\alpha}), H(\boldsymbol{\theta})) = (g(\mathbf{Y}, \boldsymbol{\alpha})\text{-}H(\boldsymbol{\theta}))^2$.

(5) Optimise. Optimisation amounts to minimising the average posterior expected loss, called the Bayes risk. In other words, the α_i are chosen to minimise:

$$\int_{\boldsymbol{\theta}}\int_{X} \sum_{\text{time}} L(g(\boldsymbol{Y},\boldsymbol{\alpha}), H(\boldsymbol{\theta}))\, p(\boldsymbol{\theta}|X)\, dX\, d\boldsymbol{\theta} \quad (3)$$

where $p(\boldsymbol{\theta}|\boldsymbol{X}) \propto p(\boldsymbol{\theta})\ell(\boldsymbol{X}|\boldsymbol{\theta})$. If $\hat{\boldsymbol{\alpha}}$ minimises the Bayes risk, then $g(\boldsymbol{Y},\hat{\boldsymbol{\alpha}})$ is the optimal *SLA* within the class of *SLA*s characterised by a model class $g()$ and the inferential context established by $p(\boldsymbol{\theta})$, the likelihood and the loss function. By 'optimal' we mean that it is the estimated Bayes rule. Givens (1999b) also notes that it is an admissible estimate.

In practice, an estimated $\hat{\boldsymbol{\alpha}}$ can be obtained by replacing the integrals in (3) with summations over the simulated trials and replicated data. The objective function is still a smooth function in $\boldsymbol{\alpha}$ that depends only on a fixed set of constants (instances of $\boldsymbol{X}$ and $\boldsymbol{\theta}$).

Full technical details about H-optimisation and merging are provided by Givens (1997; 1999b).

Of course, an *SLA* with reduced mean Bayes risk does not necessarily have superior performance with respect to univariate evaluation statistics such as final depletion or total need satisfaction. However in the examples to date examined (Givens, 1997; 1998; 1999b), reduced risk did translate to enhanced performance. The application discussed next is no exception: a 90% reduction in Bayes risk resulted in an *SLA* that generally allowed more strikes at less depletion risk to the stock. This improvement in performance was possible because the optimised *SLA* made more efficient use of the available data.

OPTIMISATION OF THE PUNT-BUTTERWORTH *SLA*

The analysis that follows is organised into sections corresponding to the steps outlined above.

Identification of the ideal strike limits, $H(\boldsymbol{\theta})$

The choice of $H(\boldsymbol{\theta})$ given in equation (1) was used because it reflects IWC Scientific Committee performance preferences.

Identification of potentially useful predictors, $\mathbf{Y}_0$

The Punt-Butterworth SLA

The goal of this paper is to improve the *SLA* described by Punt and Butterworth (1997). A slight variation on their procedure will hereafter be called the 'nominal' *SLA*. For setting a strike limit Q^{pb} in year t, their procedure is roughly as follows.

The *SLA* is based on a modified version of the estimator underlying the *Catch Limit Algorithm* (*CLA*) of the RMP for commercial whaling (IWC, 1994). Define Z_1 to be the q_1^{pb} posterior percentile of P_{t+20}/K, where P_t is the total stock size in year t, K is the corresponding carrying capacity, and the posterior distribution is the one calculated by their version of the estimator component of the *CLA*. Define Z_2 to be the q_2^{pb} percentile of P_{t+20}/P_t with respect to the same posterior. Projections made at time t about future P_{t+n} assume that future hunting mortality will remain constant for n years at the level it was in year t.

Punt and Butterworth (1997) made some alterations to the estimator at the core of the *CLA*; since these are superseded below, they are not mentioned further here.

The *SLA* estimates the highest level of catch which leads to at least one of $Z_1 \geq$ MSYL and $Z_2 \geq 1$ being satisfied. If this catch level is less than aboriginal need, then let Q^{pb} equal this catch level. Otherwise, let Q^{pb} equal aboriginal need. Since the estimator underlying the *CLA* does not employ an age-stratified dynamics model, MSYL, P_t, and K should be

interpreted here as referring to the total population component rather than the component aged 1 year or older.

The estimation relies on a variety of tuning parameters whose values were provided by Punt.

Modifications to the Punt-Butterworth SLA and predictors derived from it

To find an optimal variant of the Punt-Butterworth *SLA*, some small modifications were introduced to the original procedure. This modified version was used in the merging and optimisation framework.

For convenience, the full computation of Q^{pb} was done only each time a new abundance estimate was obtained. If aboriginal need changed in the interim, Q^{pb} was changed to equal the minimum of the most recent catch bound and the new need level. Otherwise, Q^{pb} remained constant until the new abundance estimate was available. There was no phase-out rule. All of these convenient omissions would be remedied in an implementable *SLA*.

The estimator used in the modified Punt-Butterworth *SLA* was revised from that proposed by Punt and Butterworth. Their estimator was based on that used in the *CLA* (IWC, 1994). In that original algorithm, the log-likelihood of the data was downweighted relative to the log prior by a multiplicative factor of

$$k = \frac{1}{16}.$$

The modification introduced here is that the downweighting factor was taken to be a time series starting at value k in the first year and ending at $k/(\beta^{pb})^2$ in the final year of the 100-year simulated management period. The change in this factor was not linear with time; rather β^{pb} changed linearly with time. Positive values of β^{pb} were allowed.

The variables Z_1, Z_2 and Q^{pb} from this *SLA* (at various tunings) were used as potential predictors.

Predictors from other SLA*s*

The next step is to exploit the merging idea (Givens, 1997; 1999b). Merging is the empirical combination of strike limits from several independent *SLA*s to produce a final strike limit which minimises posterior expected loss. The merged *SLA* never has worse Bayes risk than the best of the individual *SLA*s. In some early examples, modest improvements were achieved through merging (Givens, 1997; 1999b; Givens *et al.*, 1999).

For this paper, two additional *SLA*s were used to generate predictors. These *SLA*s were based on the catch control laws denoted Q_0 and Q_1 by Wade and Givens (1997). Q_0 was originally designed by Givens *et al.* (1996) to mimic the existing aboriginal whaling management protocol as defined in Sub-paragraph 13(a) of the International Convention for the Regulation of Whaling (IWC, 1995a) and reiterated by IWC Resolution 1994-4 (1995b) calling for AWMP development. Specifically, the catch control laws used here were:

$$Q_0 = \begin{cases} \beta_3 \text{MSY} & \text{if } P_t / K \geq \text{MSYL} \\ \min(\beta_1(\text{RY}-1), \beta_2 \text{MSY} & \text{if } 2000/K \leqslant P_t / K < \text{MSYL} \\ 0 & \text{if } P_t < 2000 \end{cases} \quad (4)$$

and

$$Q_1 = \begin{cases} \beta_3 \text{MSY} & \text{if } P_t / K \geq \text{MSYL} \\ \min(\beta_1 P_t \text{MSYR}, \beta_2 \text{MSY}) & \text{if } 2000/K \leqslant P_t / K < \text{MSYL} \\ 0 & \text{if } P_t < 2000 \end{cases} \quad (5)$$

where the β_i are tuning parameters chosen to reflect performance goals.

The q^{wg} posterior quantiles of Q_0 and Q_1 (at various *SLA* tunings) were used as potential predictors. The posterior used to derive quantiles is the one calculated by the estimator component of the modified Punt-Butterworth *SLA*.

Again, for convenience, the full computation was done only each time a new abundance estimate was obtained. If aboriginal need changed in the interim, the strike limit was changed to the minimum of the calculated limit and the new need level. Otherwise, the strike limit remained constant until the new abundance estimate was available. There was no phase-out rule. All of these convenient omissions would be remedied in an implementable *SLA*.

Predictors based on estimated population dynamics parameters

Another source of potentially useful predictors is intermediate calculations in *SLA* assessment models. The q^{bio} posterior quantiles of four biological quantities were used as potential predictors: MSYR, MSYL, MSY and RY. Quantiles were calculated with respect to the posterior derived from the estimator component of the modified Punt-Butterworth *SLA* (at various tunings). K was not used because it is linearly related to MSYL for the dynamics model used. However, the MSYL predictor was expressed as a number of whales rather than a fraction of K.

For convenience, these quantiles were calculated only when an abundance estimate was obtained.

The complete set of potential predictors, $\mathbf{Y}_0$

The eventual goal was to find the optimal α_i for models of the form:

$$Q = \min(N, \max(0, \alpha_0 + \alpha_1 Y_1 + \ldots + \alpha_n Y_n)) \quad (6)$$

where the Y_i were individual quantiles from or two-way multiplicative interactions between quantiles from the following list of variables: Z_1, Z_2, Q^{pb}, MSYR, MSYL, MSY, RY, Q_0, and Q_1.[1] The immediate problem, however, was to determine which predictors to use. Unlike some previous applications of the method, no polynomial predictors were considered and time was not allowed as a predictor. Superior results could be obtained if these restrictions were relaxed.

Many different versions of these potential predictors were considered by varying the values of k, β^{pb}, q_1^{pb}, q_2^{pb}, q^{bio}, q^{wg}, β_1, β_2 and β_3. Table 1 lists the values of these parameters used to generate each potential predictor. Within each block of this table, a full factorial crossing of relevant parameters (i.e. all possible combinations), and all possible multiplicative two-way interactions except those involving italicised entries, were used to generate potential predictors. Thus, for example, the first block of Table 1 describes

$$3 \times 4 \times 2 + \binom{24}{2} = 300 \text{ possible predictors } Y_i.$$

Table 1 therefore lists 28,436 potential predictors overall.

To consider all possible multiplicative two-way interactions between blocks would raise the total number of potential predictor variables to 11,103,828. If one calculated the predictors only every 5 simulation years, on 100 replicates of two simulation trial scenarios, this would amount to a dataset of over 44 billion numbers to be used for

[1] For Q^{pb}, both Z_1 and Z_2 were calculated using $q_1^{pb} = q_2^{pb}$ and the same values for K and β^{pb}. This reduces the number of possible predictors.

prediction of $H(\boldsymbol{\theta})$. Such effort is far beyond the scope of this paper, so some shortcuts are described in the section on predictor selection.

Software used

Data used in this paper (including values of predictors and $H(\boldsymbol{\theta})$) were generated using the August 1998 version of the IWC's AWMP simulation software, which implemented the simulation model, type-2 trials and summary statistics specified at the September 1997 Scientific Committee meeting. Since September 1997 the model, trials and statistics have been substantially modified. IWC software to implement these changes was not available at the time this paper was written. The most current version of the IWC AWMP simulation software can be found at *www.colostate.edu/~geof/iwcawmp.html*.

Selecting predictors and a model class

This section describes the process by which a subset of predictors, $\mathbf{Y}$, and a model class, $g(\mathbf{Y}, \alpha)$, were selected. Table 1 lists 28,436 potential predictors, and ideally all 11 million potential predictors arising from two-way multiplicative interactions should be added to that list. Without computing all of these, it is instructive to examine the available 28,436 predictors listed in Table 1 to look for hints about which ones might be involved in useful interactions. To limit computational effort, only a single investigative measure was used: the sample correlation of each of the 28,436 predictors with the ideal strike limits[2], pooled across the B3 and B7 trials. Based on these correlations, some of the predictors with the highest positive correlations with ideal strike limits were chosen, along with some which had strong negative correlations with ideal. Among both sets, the selection also focussed on maintaining a diverse list of potential predictors.

Table 1

Summary of predictor variables examined during correlation search strategy. Within each block, the results from all possible combinations of parameter values were correlated with ideal strike limits, as were the results from all two-way multiplicitive interactions except for those involving the italicised parameter value entries. Table 2 shows which between-block interactions were examined.

Predictor Variables	Parameters	Parameter Values
Z_1	q_1^{pb}	0.3, 0.4, 0.5
	k	$1, \frac{1}{4}, \frac{1}{16}, \frac{1}{36}$
	β^{pb}	$\sqrt{k}$, 1
Z_2	q_2^{pb}	0.3, 0.4, 0.5
	k	$1, \frac{1}{4}, \frac{1}{16}, \frac{1}{36}$
	β^{pb}	$\sqrt{k}$, 1
Q^{pb}	$q_1^{pb} = q_2^{pb}$	0.3, 0.4, 0.5
	k	$1, \frac{1}{4}, \frac{1}{16}, \frac{1}{36}$
	β^{pb}	$\sqrt{k}$, 1
Biological quantities	Name	MSYR, MSYL, MSY, RY
	q^{bio}	0.05, 0.10, 0.15, *0.20*, 0.25, *0.30*, *0.35*, *0.40*, *0.45*, *0.50*
	k	$1, \frac{1}{4}, \frac{1}{16}, \frac{1}{36}$
	β^{pb}	$\sqrt{k}$, 1
Wade-Givens strike limits	Catch control law	Q_0, Q_1
	β_1	*1.0*, 0.9, .075
	β_2	1.0, *0.9*, 0.75
	β_3	1.0, *0.9*, 0.75
	q^{wg}	0.05, *0.10*, 0.15, *0.20*, *0.25* *0.30*, *0.35*, *0.40*, *0.45*, *0.50*
	k	$1, \frac{1}{4}, \frac{1}{16}, \frac{1}{36}$
	β^{pb}	$\sqrt{k}$, 1

By examining the results in this manner, a list of 128 potential predictors was identified for closer examination of interactions; these are listed in Table 2. The new predictors consisting of all two-way multiplicative interactions between those listed in Table 2 were examined. Excluding those interactions already considered as part of Table 1, that amounts to another 8,128 potential predictors, for a total of 36,564 predictors that one might use to find an optimal *SLA*. This amounts to less than 0.33% of the potential predictors originally mentioned, and even a smaller percent of all possible useful predictors since predictors generated from other biological quantities, other *SLA*s, or various non-linear functions any predictor were never considered.

Despite having pared the problem down so much, a formidable search remained: which subset of the 36,564 remaining potential predictors was the best to use for estimating ideal strike limits? For the model $Q = \min(N, \max(0, \alpha_0 + \alpha_1 Y_1 + \ldots + \alpha_p Y_p))$, with $p \leq 36{,}564$, there would be $2^{36{,}564}$ different possible models – far too many to apply any standard statistical model selection technique. Again, the correlations between predictors and ideal strike limits were used to simplify the problem. Using the same reasoning as previously, the 28 candidate predictors shown in Table 3 were chosen for further study. These included *SLA*s based on Q_0, Q_1 and Q^{pb}.

The required optimisation is non-linear and it would still be computationally demanding to fit all 2^{28} possible models. However, H-optimisation models closely resemble linear regression models, and there are a variety of fast, efficient statistical model selection procedures that can be employed to identify the best and most parsimonious regression models. The 'leaps and bounds' approach of Furnival and Wilson (1974) was used to compare the Mallows (1973; Neter *et al.*, 1990) for all possible regression models. Mallows C_p is a popular measure of how poorly a model fits, with an additional penalty term for model complexity. Fig. 1 shows a plot of $\log(C_p)$ versus p, with several interesting models labelled. The best[3] regression model according to this criterion used the constant term plus 18 predictors: all those in Table 3 except c, d, g, h, j, k, n, q, za and zb. One model which fit much more poorly but stood out among all such simpler models was the model using only predictors e, f, m and z. The model using only these four predictors was adopted as $g(\boldsymbol{\theta}, \boldsymbol{\alpha})$.

Defining the Bayesian estimation context

To reduce computing demands, only two trials are considered: the optimistic B3 and pessimistic B7 scenarios. Thus, $p(\boldsymbol{\theta})$ assigns non-zero weight to these two scenarios. Equal weight was given to each. Two reasons why this crude approximation to a full integration over $\boldsymbol{\theta}$ may be adequate are as follows. First, the goal is to find a model class and estimation framework which allow the identification of good *SLA*s. Despite examining only two scenarios, the chosen

[2] The ideal strike limits, $H(\boldsymbol{\theta})$, are described later in the next subsection.

[3] This model was the simplest model for which C_p achieved a minimal value less than or equal to p.

Table 2

For the k and β^{pb} choices indicated in the side margin, the interactions between certain variants of each predictor listed in the top margin were investigated in a complete two-way factorial search. Thus, the two-way multiplicative interactions between all possible pairs of entries in this table were correlated with ideal strike limits. For 'biological quantities', all four of MSYR, MSYL, MSY and RY were used.

	Predictor Variables			
(k, β^{pb})	Z_1 when q_1^{pb} = ...	Z_2 when q_2^{pb} = ...	Biological quantities when q^{bio} = …	Q_0 and Q_1 when $(\beta_1, \beta_2, \beta_3, q^{wg})$ = …
$(\frac{1}{4}, \sqrt{k})$	0.3	0.3	0.05 0.10 0.15	(0.9, 0.75, 0.75, 0.15) (0.9, 0.75, 0.75, 0.05) (0.9, 0.75, 1.0, 0.15) (0.9, 0.9, 1.0, 0.05)
$(\frac{1}{4}, 1)$	0.3 0.5	0.3	0.05 0.10 0.15	(0.9, 0.75, 0.75, 0.05)
$(\frac{1}{16}, \sqrt{k})$	0.3 0.4	0.3 0.4	0.05 0.15 0.25	(0.9, 0.75, 0.75, 0.15) (0.9, 1.0, 0.75, 0.15) (1.0, 0.75, 0.75, 0.15)
$(\frac{1}{16}, 1)$	0.5	0.3 0.4	0.05 0.10 0.15 0.25	(0.9, 1.0, 1.0, 0.05) (0.75, 1.0, 1.0, 0.05) (0.75, 0.75, 1.0, 0.05) (0.75, 1.0, 0.75, 0.05) (0.9, 0.75, 0.75, 0.05) (0.75, 0.75, 0.75, 0.05) (1.0, 1.0, 0.75, 0.05)
$(\frac{1}{36}, \sqrt{k})$	0.3 0.4 0.5	0.3 0.4 0.5	0.25 0.50	(1.0, 0.75, 0.75, 0.25)
$(\frac{1}{36}, 1)$	0.5	0.3 0.4 0.5	0.10 0.15 0.25	(1.0, 0.9, 0.75, 0.10)

approach provides ample information about *SLA* performance and a flexible class of models. Therefore, it is likely that the optimum within this space will represent a significant improvement in terms of management performance, even if a broader exploration of trials might have found an even better *SLA*. Second, the apparent high-dimensionality of the scenario space indexed by $\boldsymbol{\theta}$ is somewhat misleading. For whales like the bowhead, population dynamics and management essentially boil down to a simple catch/productivity trade-off. The B3 and B7 trials effectively stake out the ends of this continuum, and all other trials lie between these. It is not critical to evaluate all corners of a high-dimensional $\boldsymbol{\theta}$ space if this space essentially maps onto a 1-dimensional catch/productivity continuum. Averaging performance at the endpoints of this continuum can capture most relevant performance features of an *SLA*.

The likelihood, $\ell(\boldsymbol{X}|\boldsymbol{\theta})$, is determined by the IWC's simulation framework (IWC, 1999) which generates simulated abundance estimates using a mechanism that includes lognormal errors with contamination from a complex process error model. A full discussion of this likelihood is beyond the scope of this paper; however $\ell(\boldsymbol{X}|\boldsymbol{\theta})$ need not be explicitly calculated to carry out the optimisation. Integration of equation (3) with respect to $\boldsymbol{X}$ is done via Monte Carlo by summing over the results from replicate simulated data series without requiring knowledge of the stochastic mechanism that generated these data.

The other aspect of Bayesian estimation is the loss fuction. Three different loss functions were used for *SLA* optimisation:

$$L_1(Q,H) = (H-Q)^2 \Big/ \sum_{\text{time}} \sum_{\boldsymbol{X}} H^2 \tag{7}$$

$$L_2(Q,H) = |H-Q| \Big/ \sum_{\text{time}} \sum_{X} |H| \tag{8}$$

$$L_3(Q,H) = (H-Q)^2 \tag{9}$$

Equation (7) matches the prescriptions of IWC (1999) except in one detail: H and Q are both calculated in year t assuming that past catches have been taken according to a nominal *SLA*. This differs from the IWC (1999) prescription that past catches should be taken according to H for calculation of an ideal strike limit and according to the *SLA* itself for calculation of Q. The reasons for this deviation are that: (i) the current version of the simulation control program does not implement what is prescribed; and (ii) huge computational demands are already imposed in this analysis without this added complexity. For the B3 trial used in this example, the deviation is irrelevant since both versions of H always equal need levels. In the remaining cases, the difference in the two versions of H or Q amounted to only a few whales. Therefore, the effect of this deviation from IWC (1999) should be quite small.

The optimisation: estimating $\hat{\boldsymbol{\alpha}}$ to minimise Bayes risk

Based on the results of the variable selection, the model:

$$Q = \min(N, \max(0, \alpha_0 + \alpha_1 Y_e + \alpha_2 Y_f + \alpha_3 Y_m + \alpha_4 Y_z)) \tag{10}$$

Table 3

The 28 predictors used at the final stage of model selection. Note that nominal *SLA*s Q^{pb}, Q_0 and Q_1 are listed as predictors w, zb and za respectively.

Label	Predictor
a	$\text{MSYR}_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q^{bio}=0.15\}} \times Z2_{\{k=\frac{1}{36},\beta^{pb}=1,q_2^{pb}=0.5\}}$
b	$\text{MSYL}_{\{k=\frac{1}{4},\beta^{pb}=1,q^{bio}=0.15\}} \times Z2_{\{k=\frac{1}{16},\beta^{pb}=1,q_2^{pb}=0.4\}}$
c	$Q0_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=0.9,\beta_2=1.0,\beta_3=1.0,q^{bio}=0.05\}} \times \text{MSYL}_{\{k=\frac{1}{4},\beta^{pb}=1,q^{bio}=0.05\}}$
d	$Q0_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=0.9,\beta_2=1.0,\beta_3=1.0,q^{bio}=0.05\}} \times Z2_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_2^{pb}=0.3\}}$
e	$\text{RY}_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q^{bio}=0.05\}} \times \text{MSY}_{\{k=\frac{1}{4},\beta^{pb}=\sqrt{k},q^{bio}=0.05\}}$
f	$\text{RY}_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q^{bio}=0.05\}} \times \text{MSYL}_{\{k=\frac{1}{36},\beta^{pb}=1,q^{bio}=0.25\}}$
g	$Q0_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=0.9,\beta_2=1.0,\beta_3=1.0,q^{bio}=0.05\}} \times Q0_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},\beta_1=0.9,\beta_2=0.75,\beta_3=0.75,q^{bio}=0.15\}}$
h	$Q0_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=0.75,\beta_2=1.0,\beta_3=1.0,q^{bio}=0.05\}} \times Q0_{\{k=\frac{1}{4},\beta^{pb}=\sqrt{k},\beta_1=0.9,\beta_2=0.75,\beta_3=0.75,q^{bio}=0.05\}}$

Label	Predictor	Label	Predictor
i	$Z1_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q_1^{pb}=0.4\}} \times Z1_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_1^{pb}=0.4\}}$	s	$\text{RY}_{\{k=\frac{1}{16},\beta^{pb}=1,q^{bio}=0.05\}}$
j	$Z1_{\{k=\frac{1}{16},\beta^{pb}=1,q_1^{pb}=0.5\}} \times Z1_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_1^{pb}=0.3\}}$	t	$\text{MSYR}_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q^{bio}=0.20\}}$
k	$Z2_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q_2^{pb}=0.3\}} \times Z2_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_2^{pb}=0.3\}}$	u	$\text{RY}_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q^{bio}=0.20\}}$
l	$Z2_{\{k=\frac{1}{16},\beta^{pb}=1,q_2^{pb}=0.3\}} \times Z2_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_2^{pb}=0.3\}}$	v	$Q^{pb}_{\{k=\frac{1}{4},\beta^{pb}=\sqrt{k},q_1^{pb}=q_2^{pb}=0.3\}}$
m	$Z1_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q_1^{pb}=0.3\}}$	w	$Q^{pb}_{\{k=\frac{1}{16},\beta^{pb}=1,q_1^{pb}=q_2^{pb}=0.4\}}$
n	$Z1_{\{k=\frac{1}{36},\beta^{pb}=1,q_1^{pb}=0.4\}}$	x	$Q1_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},\beta_1=1.0,\beta_2=0.75,\beta_3=0.75,q^{bio}=0.15\}}$
o	$Z1_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_1^{pb}=0.4\}}$	y	$Q0_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},\beta_1=1.0,\beta_2=1.0,\beta_3=1.0,q^{bio}=0.15\}}$
p	$Z2_{\{k=\frac{1}{16},\beta^{pb}=1,q_2^{pb}=0.3\}}$	z	$Q0_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=1.0,\beta_2=1.0,\beta_3=0.75,q^{bio}=0.05\}}$
q	$Z2_{\{k=\frac{1}{16},\beta^{pb}=\sqrt{k},q_2^{pb}=0.3\}}$	za	$Q1_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=0.9,\beta_2=0.9,\beta_3=0.9,q^{bio}=0.05\}}$
r	$Z2_{\{k=\frac{1}{36},\beta^{pb}=\sqrt{k},q_2^{pb}=0.3\}}$	zb	$Q0_{\{k=\frac{1}{16},\beta^{pb}=1,\beta_1=0.9,\beta_2=0.9,\beta_3=0.9,q^{bio}=0.05\}}$

was fit, where the *Y* variables are subscripted to refer to the labels in Table 3.

Optimisation relied on a quasi-Newton method with multiple starting values using the double dogleg step with the BFGS secant update to the Hessian (Dennis and Mei, 1979; Dennis *et al.*, 1981).

RESULTS

Table 4 shows summary results for the original Punt-Butterworth *SLA*, the modified version of it and the optimisations described above. The *SLA*s in this table were not intentionally equivalenced in any way. However, each of these *SLA*s has a median $D1_{1+}$ result on the B7 trial within 0.02 of the result achieved by H (namely 51.3). Therefore, it is fair to compare these *SLA*s since they are balanced with respect to depletion risk.

The results in Table 4 describe the final depletion ($D1_{1+}$) and total need satisfaction (N1) achieved in two type-2 trials: the pessimistic B7 trial and the optimistic B3 trial. High N1 values are desired (perfect = 100), as are $D1_{1+}$ values in the range 60-100. As $D1_{1+}$ increases above 60, higher values become less important than increasing N1. Considering a hypothetical bowhead stock of 9,000 whales and need ranging from 68-204, a ten-unit change in the depletion

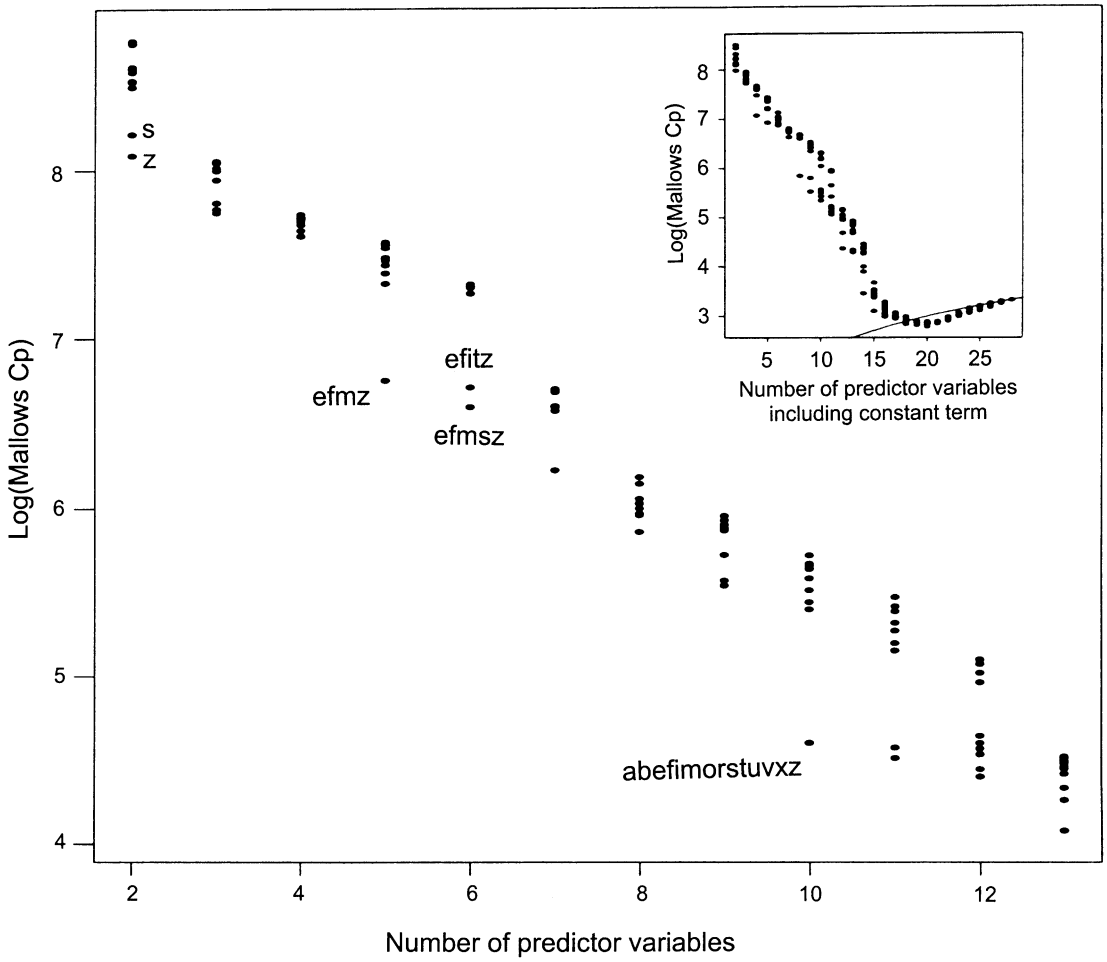

Fig. 1. A portion of the plot of Mallows C_p versus the number of predictors (including the constant term) for selecting a model for optimisation. This plot was made after omitting the 10 candidate predictors listed in the text. Several interesting models are labelled according to the annotation in Table 3. The inset shows the full C_p plot using all 28 candidate predictors; the reference 45° line is superimposed.

scores in Table 4 corresponds to 900 bowhead whales, and a ten-unit change in the need satisfaction scores corresponds to between about 7 and 20 strikes per year.

Table 4

Percentiles of the $D1_{1+}$ (final depletion) and N1 (total need satisfaction) summary statistics for the nominal SLA and several optimised variants. B7 and B3 are pessimistic and optimistic type-2 trial scenarios, as given by IWC (1999). Results are multiplied by 100. High N1 values are desired (perfect=100), as are $D1_{1+}$ values in the range 60-100.

				Strike Limit Algorithm				
Trial	Stat.	Perc.	P-B	Mod. P-B	Opt. L_1	Opt. L_2	Opt. L_3	Opt. other
B7	$D1_{1+}$	5%	35	48	41	41	40	41
		50%	50	51	51	50	50	50
	N1	5%	39	55	53	52	52	53
		50%	57	58	59	60	60	59
B3	$D1_{1+}$	5%	87	95	85	85	85	85
		50%	91	97	88	87	86	86
	N1	5%	86	64	76	80	82	78
		50%	93	68	92	95	96	95

The optimal *SLA* based on L_3 using the tuning in Table 4 was:

$$Q = \min(N, \max(0, 74.74 + 0.02137Y_e - 0.0002635Y_f + 37.50Y_m + 1.649Y_z)) \quad (11)$$

The optimised *SLA*s using L_2 and L_1 had similar coefficients. The optimisations reduced Bayes risk by 95%, 94% and 93%, respectively, when using L_1, L_2 and L_3, compared to the modified Punt-Butterworth *SLA*.

The column labelled 'Opt., other' in Table 4 shows results for an *SLA* not yet mentioned. For this variant, the Punt-Butterworth *SLA* was further modified so that a linear trend in hunting mortality, estimated from the last 20 years, was projected forward from current hunting mortality when making projections of P_{t+n}. The nominal method had been to assume constant future hunting mortality. This change alters the quantities Z_1, Z_2 and Q^{pb}. The model search and optimisation was repeated starting with a subset of the predictors in Table 3: a, b, e, f, i, l, m, o, p, r, z. These were chosen because: (i) they involved e, m, f, z, Z_1, or Z_2; (ii) they did not involve Q^{pb} which was previously much less useful than Z_1 and Z_2; and (iii) they were not among the 10 variables eliminated from Fig. 1 for their irrelevance given the remaining predictors. This search identified the model using b, e, f, l, m and p as one with an attractive Mallows C_p. This 6-predictor model was then optimised and tuned in the same manner as the rest. The results show that the change to linear projection of future hunting mortality does not appear to have improved *SLA* performance.

Clearly the modifications to the Punt-Butterworth *SLA* did not unequivocally improve its performance; therefore any superior performance for the optimised variants relative to this *SLA* should be attributed to the optimisation process rather than to the modifications themselves.

Table 4 shows that the optimised *SLA*s offered some improvements relative to the nominal *SLA*; in the risky scenario the optimised variants allowed more catch and simultaneously more protection for the stock, while in the safe scenario, they usually satisfied more need (although not at the 5th percentile). The optimised variants provided this superior performance despite having been built upon the apparently inferior, modified *SLA* rather than Punt and Butterworth's original procedure. The extra protection provided by the optimised variants usually did not incur a significant decrease in catch – in fact the extra protection was usually achieved while simultaneously allowing more catch. Such performance is possible because the optimised variants extracted more or better information from the available data than did the nominal *SLA*. Although the optimisation and merging strategy employed here is complex and computationally intensive, its ability to provide higher catch limits at less risk to the stock – achieved through a more efficient extraction of information from the available data and a reduction in estimation uncertainty – is highly desirable. Its success in this realistic example suggests that the approach may also be rewarding for the development of procedures for environmental and wildlife management in other settings.

ACKNOWLEDGEMENTS

This research was partially supported by Contract No. 99-181 from the North Slope Borough (Alaska), the State of Alaska (through the Alaska Department of Community and Regional Affairs), and the National Oceanic and Atmospheric Administration (through the National Marine Mammal Laboratory to the Alaska Eskimo Whaling Commission), Purchase Order #40ENNF900079 from the National Marine Fisheries Service and Purchase Order #40ABNF801829 from the National Marine Mammal Laboratory. Two anonymous referees are thanked for their comments which greatly improved the paper.

REFERENCES

Dennis, J.E. and Mei, H.H.W. 1979. Two new unconstrained optimization algorithms which use function and gradient values. *J. Opt. Theory and Appl.* 28:453-83.

Dennis, J.E., Gay, D.M. and Welsch, R.E. 1981. An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software* 7:348-83.

Donovan, G.P. 1999. Editorial. *J. Cetacean Res. Manage.* 1(1):ii-v.

Furnival, G.M. and Wilson, R.W. 1974. Regressions by leaps and bounds. *Technometrics* 16:499-511.

Givens, G.H. 1997. Separable multicriterion design and performance optimization of AWMPs. Paper SC/49/AS4 presented to the IWC Scientific Committee, September 1997 (unpublished). 21pp. [Paper available from the Office of this Journal].

Givens, G.H. 1998. AWMP development and diverse prototypes. *Rep. int. Whal. Commn* 48:483-95.

Givens, G.H. 1999a. Comparing methods for optimising *Strike Limit Algorithms*. Paper SC/51/AWMP5 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 18pp. [Paper available from the Office of this Journal].

Givens, G.H. 1999b. Multicriterion decision merging: competitive development of an aboriginal whaling management procedure. *J. Am. Stat. Assoc.* 94:1003-4.

Givens, G.H. 1999c. On statistical methods that permit and improve the use of simpler model-based *Strike Limit Algorithms*. Paper SC/51/AWMP4 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 12pp. [Paper available from the Office of this Journal].

Givens, G.H. and Bernstein, T.A.O. 1998. Further investigations of AWMP *Strike Limit Algorithms* and their optimisation. Paper SC/50/AWMP5 presented to the IWC Scientific Committee, April 1998 (unpublished). [Paper available from the Office of this Journal].

Givens, G.H., Zeh, J.E. and Raftery, A.E. 1996. Implementing the current management regime for aboriginal subsistence whaling to establish a catch limit for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. *Rep. int. Whal. Commn* 46:493-9.

Givens, G.H., Punt, A.E. and Bernstein, T.A.O. 1999. Equivalence tuning of *SLA*s. *J. Cetacean Res. Manage.* 1(3):289-95.

International Whaling Commission. 1992. Report of the sub-committee on management procedures, Appendix 2. Technical descriptions of revised management procedures. *Rep. int. Whal. Commn* 42:93-103.

International Whaling Commission. 1994. Report of the Scientific Committee, Annex H. The Revised Management Procedure (RMP) for Baleen Whales. *Rep. int. Whal. Commn* 44:145-52.

International Whaling Commission. 1995a. Chairman's Report of the Forty-Sixth Annual Meeting, Appendix 21. Amendments to the Schedule. *Rep. int. Whal. Commn* 45:52.

International Whaling Commission. 1995b. Chairman's Report of the Forty-Sixth Annual Meeting, Appendix 4. IWC Resolution 1994-4. Resolution on a Review of Aboriginal Subsistence Management Procedures. *Rep. int. Whal. Commn* 45:42-3.

International Whaling Commission. 1999. Report of the Scientific Committee. Annex F. Report of the Standing Working Group (SWG) on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 1:157-78.

International Whaling Commission. 2000. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 2:125-54.

Mallows, C.L. 1973. Some comments on Cp. *Technometrics* 15:661-75.

Neter, J., Wasserman, W. and Kutner, M.H. 1990. *Applied Linear Statistical Models*. Irwin, Boston. i-xi+1,181pp.

Punt, A.E. and Butterworth, D.S. 1997. Preliminary evaluation of a (generic) strike limit algorithm for aboriginal subsistence whaling with comments on performance statistics and simulation trials. Paper SC/49/AS7 presented to the IWC Scientific Committee, September 1997, Bournemouth (unpublished) 22pp. [Paper available from the Office of this Journal].

Wade, P.R. and Givens, G.H. 1997. Designing catch control laws that reflect the intent of aboriginal subsistence management principles. *Rep. int. Whal. Commn* 47:871-4.