Application of a *Strike Limit Algorithm* based on adaptive Kalman filtering to the eastern North Pacific stock of gray whales

EVA HLÍN DEREKSDÓTTIR AND KJARTAN G. MAGNÚSSON

Science Institute, University of Iceland, Dunhagi 3, 107 Reykjavík, Iceland

Contact e-mail: kgm@hi.is

ABSTRACT

The application of a *Strike Limit Algorithm* (*SLA*) based on Adaptive Kalman Filtering techniques to the eastern North Pacific (ENP) stock of gray whales is described. This *SLA* is a modification of an earlier one which was designed for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. Extended Kalman filters are used to estimate the present stock size and posterior probability distributions for Maximum Sustainable Yield (*MSY*) and *MSY*-rate (*MSYR*). A catch control law selected from a one-parameter family of such rules is then used on the conditional estimates of stock size. These conditional strike limits together with the posterior distributions of the various combinations of *MSYR* and *MSY*, give a cumulative distribution function for the strike limit. The eventual strike limit is then determined as a pre-specified percentile of this distribution. The *SLA* can be tuned to varying degrees of risk by the choice of internal model parameters – so-called tuning parameters. The procedure is tested based on a set of trials specified by the IWC Scientific Committee Standing Working Group on Aboriginal Whaling Management Procedures, designed to test the performance of potential *SLA*s for the ENP gray whale stock.

KEYWORDS: WHALING-ABORIGINAL; MANAGEMENT; GRAY WHALES; MODELLING; MSY RATE

INTRODUCTION

Aboriginal whaling refers to subsistence hunting of large whales by native communities. In the context of aboriginal subsistence whaling a fishery type 2 as defined in IWC (2000) is a case where a substantial amount of information exists about the stock in question. An example of such a fishery is the aboriginal harvesting of the eastern North Pacific (ENP) stock of gray whales (*Eschrictius robustus*). This stock has been studied extensively (e.g. Buckland and Breiwick, 2002; Butterworth *et al.*, 2002a; b; Wade, 2002; Witting, 2003; Punt *et al.*, 2004); stock identity is unambiguous, a series of abundance observations exists (from 1968) and estimates of various parameters are available. However, despite this good information it has not been possible to reconcile the catch history with the observed population increase in recent years using a simple density dependent population model (IWC, 1993). Various adjustments have been proposed in order to address this problem, see for example Butterworth *et al.* (2002b) and Witting (2003). This paper addresses the problem of determining strike limits for this stock such that the nutritional and cultural needs of the hunting communities (as recognised by the IWC) are satisfied without endangering the stock.

The term *Strike Limit Algorithm* (*SLA*) is used in connection with aboriginal whaling. An *SLA* is an input-output rule or algorithm where a data series – usually abundance data –is input into the algorithm which produces as output the total number of whales which can be struck in any one year or block of years. An *SLA* based on Adaptive Kalman Filtering (AKF) applied to the Bering-Chukchi-Beaufort Seas (BCB) stock of bowhead whales *(Balaena mysticetus*) has been presented earlier (Dereksdóttir and Magnússon, 2001; 2003). This *SLA* is fairly general and is applicable with suitable modifications to a range of type 2 fisheries. This paper describes in detail the Adaptive Kalman Filtering *SLA* – hereafter referred to by the acronym AKF-*SLA* – applied to the eastern North Pacific stock of gray whales. This *SLA* forms one of the two component *SLA*s that make up the gray whale *SLA* – known by the acronym GUP, which stands for 'Grand Unified Procedure' – recommended by the Scientific Committee of the International Whaling Commission (IWC, 2005) and subsequently adopted by the Commission. The other component is the Johnston-Butterworth *SLA*, which uses a penalised likelihood method (see IWC, 2005 for a technical description). The strike limit produced by the GUP is the arithmetic average of the strike limits produced by each of the component *SLA*s.

The next section describes the AKF-*SLA*, starting with a general description of the basic principles, followed by a detailed mathematical description of the various components, which make up the *SLA*. Finally, some results of testing the performance of the procedure on a set of simulation trials specified by the Standing Working Group on Aboriginal Whaling Management Procedure (IWC, 2005) are given, together with some explorations of its flexibility.

THE ADAPTIVE KALMAN FILTER *STRIKE LIMIT ALGORITHM* (AKF-*SLA*)

General description

The state estimation part of the AKF-*SLA* applies the techniques of the Kalman filter (Kalman, 1960), which is a mathematical tool to obtain estimates of the state of stochastic dynamic systems with noisy observations, i.e. systems with both 'process noise' and 'observation noise'. In the case of a linear system, the estimate obtained is optimal in the sense that the mean square estimation error is minimised. In order to apply Kalman filtering methods, a mathematical model of the dynamics and the relationship between the observations and the true state – i.e. the abundance in this case – is required. The way the Kalman filter works is that the most recent state estimate is projected forward in time (a prediction) until a new observation becomes available. The prediction is then compared to the observation and the state estimate corrected. The correction or update is proportional to the difference between the prediction and the observation. A large difference results in a large correction and a small difference results in a correspondingly small update in the estimate. The proportionality constant, known as the Kalman gain, depends on the magnitude of the measurement noise and the noise in the dynamics. If the measurement noise is large and the level of confidence in the observation therefore low, the gain is small, thus giving a small correction in the model prediction. On the other hand, if the measurement noise is small relative to the process noise and the level of confidence therefore high, the gain will be high and the update thus large. The two extremes are to follow the observation exactly (corresponding to zero observation noise) or to ignore the observation completely and use only the model to obtain the state estimate (corresponding to infinite observation noise). The updated estimate of the state is then projected forward in time until a new observation is made. In the Kalman filtering application presented here, the state of the system is the size of the stock and the observations are the census estimates of the stock size.

The underlying model used in the *SLA* based on AKF is a simple population dynamics model, together with a linear model for the relationship between observed stock size and true stock size. The model contains both process noise and observation noise which are taken to be Gaussian and additive after a log transformation of the variables.

The stock dynamics model and the observation model contain a number of unknown parameters. In the basic application of the AKF-*SLA* to gray whales, two of the parameters, i.e. Maximum Sustainable Yield Level (*MSYL*) and annual survival rate *S* are fixed. The remaining two parameters, i.e. *MSY*-rate (*MSYR*) and *MSY* are estimated by Bayesian methods in conjunction with the Kalman filtering estimation scheme as described below in detail. Each of the two parameters ranges over a sequence of discrete values giving a two-dimensional grid of parameter values. A prior probability distribution is given to the parameter combinations in the grid and a Kalman filter is associated with each combination. Other choices for the parameter grid are possible in variants of the AKF-*SLA*; for example a bias filter can be added giving a three-dimensional parameter grid (see Dereksdóttir and Magnússon (2001) for an application to the BCB stock of bowhead whales).

The probability associated with each parameter combination in the grid is updated by Bayesian methods each time a new survey estimate becomes available. The estimate of the state associated with each of the combinations is updated at the same time by the corresponding Kalman filter. Thus, for each (*MSYR*, *MSY*) combination in the grid, there corresponds a posterior probability for this particular combination and an estimate of the state (i.e. stock size) conditional on this particular parameter combination. This combination of Kalman filtering and Bayesian methodology is known as AKF. The overall estimate of the present state (stock size) is then obtained by summing all the stock estimates corresponding to the different parameter combinations, weighted by the respective probabilities. This overall stock estimate is not used in the *SLA* described here.

The AKF method therefore comprises a set of Kalman filters – one filter for each parameter combination in the grid. The state estimates and the posterior probabilities associated with each point in the parameter grid and with the corresponding stock estimate are then updated every time a new survey estimate becomes available.

A catch control law selected from a one-parameter family of such rules is then used on the conditional estimates of stock size. These conditional strike limits together with the posterior distributions of the various combinations of *MSYR* and *MSY*, give a cumulative distribution function for the strike limit. The eventual strike limit is then determined as a pre-specified percentile of this distribution.

The AKF-*SLA*: mathematical description

The Kalman Filter

It is assumed that the population dynamics and observations are governed by the following equations:

$$N_{t+1} = \left(S(N_t - C_t) + (1-S)\left(1 + A\left(1 - \left(\frac{N_t}{N_\infty}\right)^z\right)\right) N_t \right) e^{u_t} \quad (1)$$

$$N_t^{obs} = N_t e^{v_t} \quad (2)$$

where N_t is the total population of animals 1 year and older (1+) in year *t*, C_t is the catch in year *t* and u_t and v_t are normal random variables with zero mean and variances q_t and r_t, respectively. This is the well-known Pella-Tomlinson (P-T) model with parameters: annual survival rate *S*, pre-exploitation population size (carrying capacity) N_∞ and the resilience parameter *A*, which is related to *MSYR* by *MSYR*=*A*(1-*S*)/(*S*(*z*(*z*+1))). The exponent *z* in equation (1) determines the *MSYL* according to *MSYL* = $(z+1)^{-1/z} N_\infty$. This model is a simplification of the usual P-T models since no delay in the dynamics is incorporated.

The state variable is defined to be $x=\ln(N)$ and the observation $y = \ln(N^{obs})$. The state and observation equations can therefore be written in the form:

$$x_{t+1} = f(x_t) + u_t \quad (3)$$

$$y_t = x_t + v_t \quad (4)$$

where:

$$f(x_t) = \ln\left(S(e^{x_t} - C_t) + (1-S)\left(1 + A\left(1 - \left(\frac{e^{x_t}}{N_\infty}\right)^z\right)\right) e^{x_t} \right) \quad (5)$$

The state of the system is estimated by the Extended Kalman Filter (the equation describing the dynamics is non-linear and hence the EKF – in which non-linear functions are linearised – must be used). In order to apply the Kalman filtering method a linearisation of *f(x)* is required:

$$F_t = \frac{\partial f}{\partial x}(x) = \frac{S \cdot e^x + (1-S)\, e^x \left(1 + A\left(1 - \left(\frac{e^x}{N_\infty}\right)^z\right) - A \cdot z \left(\frac{e^x}{N_\infty}\right)^z \right)}{S(e^x - C_t) + (1-S)\left(1 + A\left(1 - \left(\frac{e^x}{N_\infty}\right)^z\right)\right) e^x} \quad (6)$$

The estimate of the state at time *t*, using data up to *t*-1 is denoted by $x_{t|t-1}$ and is known as the prior estimate of x_t. The corresponding variance at time *t* is:

$$P_{t|t-1} = E((x_t - x_{t|t-1})^2) \quad (7)$$

When a new observation y_t becomes available, the estimate $x_{t|t-1}$ is updated according to:

$$x_{t|t} = x_{t|t-1} + K_t(y_t - x_{t|t-1}) \quad (8)$$

which is the posterior estimate of x_t i.e. the estimate of the state at time t using data up to t. Here K_t is known as the Kalman gain at time t. The term in brackets on the right hand side is the difference between the actual observation and the predicted observation at time t. Thus a large difference between the actual and predicted observations will give a large modification in the state estimate and a small difference results in a correspondingly small modification. The Kalman gain is given by:

$$K_t = P_{t|t-1}(P_{t|t-1} + r_t)^{-1} \quad (9)$$

The variance $P_{t|t-1}$ is updated by:

$$P_{t|t} = (1 - K_t)P_{t|t-1} \quad (10)$$

$P_{t|t}$ is the variance associated with the updated (posterior) estimate of the state at time t.

Finally, new prior estimators of the state and the variance at t+1, are obtained by the forward projection equations:

$$x_{t+1|t} = f(x_{t|t}) \quad (11)$$

$$P_{t+1|t} = F_t P_{t|t} F_t^T + q_t \quad (12)$$

where F_t is given by equation (6) and the linearisation is about the point $x=x_{t|t}$. The Kalman gain at time t+1 can then be calculated and hence the posterior estimate of the state at t+1 and so on.

Initial values for x_0 and P_0 (the state with an associated variance) are required to start the filter. The natural starting value is the pre-exploitation stock size N_∞, provided the catch history is fully known and the stock dynamics can be described by a standard density dependent model. This approach is adopted in the BCB bowhead version of the AKF-*SLA* (Dereksdóttir and Magnússon, 2003). However, neither of these conditions are fullfilled for the ENP gray whale stock so this does not work here. Since the first gray whale abundance estimate was in 1968 it would seem natural to start the filters in that year. However, this entails that all trajectories pass through the value of the 1968 estimate (i.e. 12,921), which might be regarded as an unreasonable constraint since this estimate is no more correct than subsequent ones. One way to avoid forcing the trajectories through the 1968 estimate is to start the filters earlier and either use this earlier starting value as an additional parameter to be estimated or simply start at an arbitrary value with some associated variance.

We have selected 1930 as the starting year for the filters. The initial condition for the Kalman filters is therefore a stock estimate for 1930 together with an associated coefficient of variation (*CV*). However, since no abundance estimate from 1930 exists, the starting value and the associated variance can be freely chosen and used as tuning parameters. The 1930 population size is normalised by the carrying capacity N_∞, i.e. the 1930 population size N_{1930} is defined by a tuning parameter, α, where $N_{1930} = \alpha N_\infty = (MSY/(0.6MSYR))$ with an associated *CV*, P_0, which is also used as a tuning parameter. The first update is made in 1968 when the first abundance estimate becomes available.

Bayesian estimation of model parameters

Equation (1) contains four unknown parameters, *S, A, z,* and N_∞. Two of those, *z* and *S* are fixed at 2.39 (corresponding to the standard choice of *MSYL* = 0.6 N_∞) and 0.97 (this value lies well within the likelihood range obtained in Butterworth *et al.* (2002b) and in Wade (2002)), respectively and the others – i.e. the resilience parameter *A* and the carrying capacity N_∞ – estimated by Bayesian methods, or rather, the equivalent parameters *MSYR* and *MSY*=*MSYR*·0.6 N_∞ are estimated. The reason for estimating *MSYR* and *MSY* rather than *MSYR* and N_∞ is that the latter two parameters are usually highly negatively correlated. Each of the two parameters range over a sequence of discrete values giving a 2-dimensional grid of ($MSYR_i$,MSY_j), i=1,..., *I*; j=1,..., *J* values. To each of the *IJ* pairs there corresponds an extended Kalman filter. In the ENP gray whale version the (*MSYR*, *MSY*) parameter grid is made up of *MSYR* values 1%, 2%,..., 5% and 6% and *MSY* ranging from 100 to 2176 in increments of 12, i.e. 6 values of *MSYR* and 174 values of *MSY*, giving a total of 1,044 parameter combinations and the same number of filters. A few words about the range and the increments in the grid selected are appropriate here. Obviously, the number of filters should be kept low for computational reasons. That being said, there are two criteria to consider: the range of values should be sufficiently large for parameter values outside the range to have negligible probability; and the grid fine enough for the calculated probability distribution functions to be reasonably smooth and without ‘gaps’. This question will be addressed below, but we note that the range of *MSYR* values in Butterworth *et al.* (2002b) and Wade (2002) is within the 1-6% range used here. Furthermore, those authors consider a carrying capacity greater than 60,000-70,000 to be unlikely. A maximum value of 60,000 for carrying capacity and 6% for *MSYR* gives (assuming *MSY*=0.6), *MSY*=0.6×60,000×0.006=2160 which is very close to the maximum *MSY* value in the grid. However, the ultimate test of the size and fineness of the grid lies in the calculated posterior distributions, which will be presented below. Since there is no prior information on the values of the parameters *MSYR* and *MSY*, the prior distribution for the parameter set ($MSYR_i$, MSY_j), i=1,2, .., *I*; j=1,2, …, *J*, is assumed to be discrete uniform on the specified grid. This probability distribution is updated every time a new census estimate becomes available.

Whenever a new observation becomes available, the conditional stock estimate $x_{t|t-1}$ ($MSYR_i$, MSY_j), is updated as described above and the posterior probability distribution $p(MSYR_i, MSY_j|Y_t)$ is updated for each of the pairs of parameters by Bayesian methodology. Here Y_t is the set of observations up to and including time t. The probability distribution is updated as follows.

Let κ denote the vector of parameters (*MSYR*,*MSY*). There are *IJ* possible values of κ corresponding to the *IJ* pairs ($MSYR_i$, MSY_j). A prior distribution, $p(\kappa_k)$ for the vector κ is given and each time a new observation becomes available, a posterior distribution, $p(\kappa_k|Y_{t=1})$ is updated according to:

$$p(\kappa_k \mid Y_t) = \frac{p(Y_t \mid \kappa_k)\, p(\kappa_k)}{p(Y_t)} \quad (13)$$

where the conditional distribution $p(Y_t|\kappa_k)$ is given by the recursive formula:

$$p(Y_t \mid \kappa_k) = \frac{1}{(2\pi)^{1/2}(P_{t|t-1}+r_t)^{1/2}} \exp\left(-\frac{(y_t - x_{t|t-1})^2}{2(P_{t|t-1}+r_t)}\right) p(Y_{t-1} \mid \kappa_k) \tag{14}$$

where $x_{t|t-1}$, and $P_{t|t-1}$ depend on κ_k and are obtained by the Extended Kalman Filter method. A 'small' prediction error $y_t - x_{t|t-1}$, gives a 'high' value of $p(Y_t|\kappa_k)$. Finally, $p(Y_t)$ is calculated by:

$$p(Y_t) = \sum_{k=1}^{IJ} p(Y_t \mid \kappa_k)\, p(\kappa_k)\,. \tag{15}$$

To each abundance observation there is an associated estimate of the *CV*. The variance of the measurement noise ν_t is given by:

$$r_t = Var(\nu_t) = \ln(1 + CV(N_t^{obs})^2) \tag{16}$$

The estimate of *CV*, CV_{est}, in (16) is probably an underestimate of the true *CV* of the abundance estimate. The historical observations of the abundance of gray whales with the given *CV* are not compatible with a standard density dependent population model and a constant CV_{add} is therefore added to all *CV* – estimates (historical and future) provided to the *SLA*. This value is treated as a tuning parameter.

This scheme described here for updating the state estimate and the conditional probability distribution associated with the parameter values is the AKF.

Catch Control Law

Applying a catch control law corresponding to each of the *IJ* pairs of (*MSYR*,*MSY*) to $x_{t|t-1}$ ($MSYR_i$,MSY_j) a sequence of *IJ* strike limits is obtained, together with the associated posterior probability distribution $p(MSYR_i,MSY_j|Y_t)$, i=1,2,…, I; j=1,..., J. Arranging all the *IJ* strike limits in an increasing sequence, the associated probability distribution makes it possible to construct the cumulative distribution function $F(C)$ for the strike limit. Once a percentile γ of this distribution is set, a provisional strike limit is determined by solving:

$$F(C_t) = p(C < C_t) = \gamma \tag{17}$$

for C_t. A one-parameter family of catch control laws is used. If the stock size N is less than *MSYL*, then the conditional strike limit is determined by the rule $C=\rho RY$, relating catch and replacement yield (*RY*) as calculated from equation (1), and by $C=\rho MSY$ if N is greater than *MSYL*. The multiplier ρ is a function of the conditional estimate of the stock size (i.e. conditional on *MSYR* and *MSY*) and is chosen from a family of continuous piecewise linear functions. This family is parameterised by β, the ρ-value at 0.5*MSYL*. The multiplier ρ depends on N as follows:

$$\rho = \begin{cases} 0 & N < 2000 \\ \dfrac{\beta}{(0.5MSYL - 2000)}(N - 2000) & 2000 < N < 0.5MSYL \\ \dfrac{(0.8-\beta)}{0.4MSYL}(N - 0.5MSYL) + \beta & 0.5MSYL < N < 0.9MSYL \\ \dfrac{N - 0.9MSYL}{MSYL} + 0.8 & 0.9MSYL < N < MSYL \\ 0.9 & MSYL < N \end{cases} \tag{18}$$

The parameter β is a measure of the steepness of the catch control law (Fig. 1) and is used as a tuning parameter. A strike limit is then set as:

$$SL_t = \min\left(C_t, Need_t\right) \tag{19}$$

where $Need_t$ is the pre-specified level of aboriginal need in year t. All components refer to the 1+ component of the population, i.e. the total number of animals one year and older.

A so-called 'Snap to Need' feature is incorporated whereby the strike limit is increased to need if the provisional strike limit resulting from the *SLA* exceeds 95% of need, and finally, a maximum of 20% change in strike limits between years is imposed. The strike limit is set for 5-year blocks at a time.

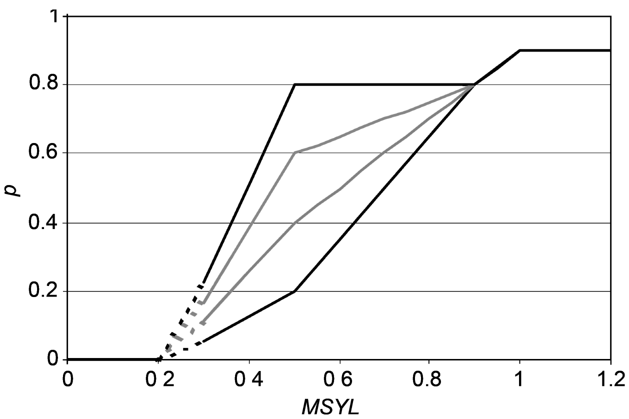

Fig. 1. A family of continuous piecewise linear catch control laws with β = 0.2, 0.4, 0.6 and 0.8. The parameter ρ is the fraction of replacement yield resulting from the catch control law, i.e. $C = \rho RY$.

RESULTS

A set of simulation trials – where each trial consists of 100 replicates simulated stochastically over a 100 year management period, starting in 2003 – for evaluating the performance of *SLA*s for the ENP gray whale stock have been developed; for a full description of all the trials see IWC (2005). The trials are conditioned on data for this stock, i.e. on the partial history of catches, past stock estimates, and parameter values. However, as mentioned above, this stock poses a problem since it is not possible to reconcile the catch history with the observed population increase in recent years (since 1968) using a simple density dependent population model (IWC, 1993). This problem is bypassed in the simulation trials by starting the population projections in 1930 – assuming a stable age distribution and ignoring the earlier catch history – rather than with a pre-exploitation stock size. The population rate of increase in 1930 is selected such that if the population dynamics model is projected from 1930 to 1968, the size of the 1+ component of the population in 1968 (the year of the first census) equals a pre-specified value, P_{1968}, selected from a probability distribution. Trials were also conducted with a so-called inertia model (Witting, 2003; IWC, 2004), which is quite different from the simple density dependent model. The performance of the various candidate *SLA*s was evaluated from a set of calculated performance statistics, designed to capture how well aboriginal need is satisfied, the risk to the stock as well as the stability of strike limits. For a full definition of all the performance statistics, see IWC (2003).

The criteria underlying the final choice of the variant and the tuning of the *SLA*, are of course the trial results, but we will also look briefly at the ability of the algorithm to detect the true *MSYR* value and to estimate the true stock size. We only present the Depletion and Need satisfaction statistics of a few key trials, GE01, GE04, GE10, GE16 and GE45 (Table 1). Note that need is set at 150 for 2003 and generally increases linearly over the management period to the value given in the column headed 'Final need'. Depletion is defined as the size of the population as a fraction of the carrying capacity and need satisfaction is the number of whales which can be struck as a fraction of the pre-specified aboriginal need. The present gray whale evaluation trials do not really pose a challenge to *SLA*s with a couple of exceptions. Need can be fully satisfied in most cases without depleting the stock unduly. The only exceptions are trials GE04 (high *MSYR* and negatively biased future observations) where need is not satisfied in spite of the stock being well above *MSY* level, GE16 (time varying bias on the historical observations, low *MSYR* and high need) and GE45 (time varying bias on the historical observations, low *MSYR* and the stock crashes in 1999/2000) where the stock may end up too depleted.

Tuning parameters and sensitivity

The gray whale version of the AKF-*SLA* contains the following tuning parameters:

(1) β: Height of the breakpoint at 0.5*MSYL* in the catch control law;
(2) γ: Percentile in the cumulative distribution function for the nominal catch limit;
(3) α: Stock size in 1930 as a fraction of carrying capacity N_∞, i.e. $N_{1930} = \alpha N_\infty$;
(4) P_0: Variance associated with N_{1930};
(5) CV_{add}: Additional variance added to the *CV* given to the *SLA*.

The variance in process error is not used in the tuning process, but is fixed at q=0.001 (corresponding to a *CV* of 3.2%). The values of all these tuning parameters were selected subjectively, rather than by attempting to optimise some function of the trial results. Tests show that the trial results are not sensitive to the value of β and this parameter was therefore fixed at 0.7 throughout. The cumulative distribution function for the nominal catch limit is shown in Fig. 2 for β=0.7. This function is 'nice and smooth' without the step function behaviour which occurred occasionally in the application of the AKF-*SLA* to the BCB stock of bowhead whales (Dereksdóttir and Magnússon, 2003). The results deteriorated as P_0 was increased and this parameter was therefore fixed at zero.

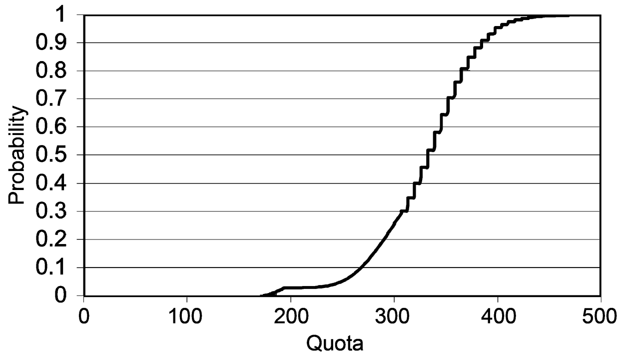

Fig. 2. The cumulative distribution function for the strike limit at the beginning of management (2003) for a set of (*MSYR, MSY*) filters.

Since the historical abundance estimates fluctuate rather wildly, in fact too much if the estimated *CV* is to be believed, it was considered necessary to increase this *CV* estimate to limit the consequent *SLA* fluctuations. However, the value of CV_{add} should not be set so high that the future observations are more or less ignored. In an attempt to achieve some balance between these two conflicting objectives, values of CV_{add}=0.10 and 0.15 were (subjectively) selected. Fig. 3 shows – for one simulation of trial GE16 – how the estimated trajectory tracks the observations more closely for small values of CV_{add}. The reason for the slight 'kink' in the true trajectory 1992-93 is that there was no hunting in those two years. There is considerable discrepancy between the true and estimated trajectories in the early part of the historical period 1968-2003. The estimated trajectory follows the observations, which are well below the true trajectory to begin with. The reason lies in the negative bias in the historical observations, which changes from 0.5 to 1.0 from 1968 to 2003. The agreement between the true and estimated trajectories from 2003 onwards is quite good.

The impact of the 1930 parameter α with values of 0.20, 0.25 and 0.30 together with the above two values of CV_{add} (0.10 and 0.15) and values of γ in the range 0.3-0.8 was investigated for two of the key trials GE16 and GE04 by plotting simultaneously the points for depletion statistic in the former and the need statistic in the latter (Fig. 4). Obviously, the further to the right (better depletion in GE16) and higher up (higher need satisfaction in GE04) the points lie, the better. Thus, a triple (α, CV_{add}, γ) of the three tuning

Table 1

Specifications for a few key gray whale evaluation trials for which results are reported.

Trial	Description	$MSYR_{1+}$	$MSYL_{1+}$	Final need	Survey freq.	Historical survey bias	Future survey bias
GE01	Base case	3.5%	0.6	340	10	1	1
GE04	Future negative bias	3.5%	0.6	340	5	1	1→0.5 in yr 25
GE10	$MSYR_{1+}$ = 5.5%	5.5%	0.6	340	10	1	1
GE16	$MSYR_{1+}$ = 1.5%; high need	1.5%	0.6	530	10	0.5 -> 1	1
GE45	GE16+40% die in 99/00	1.5%	0.6	340	10	0.5 -> 1	1

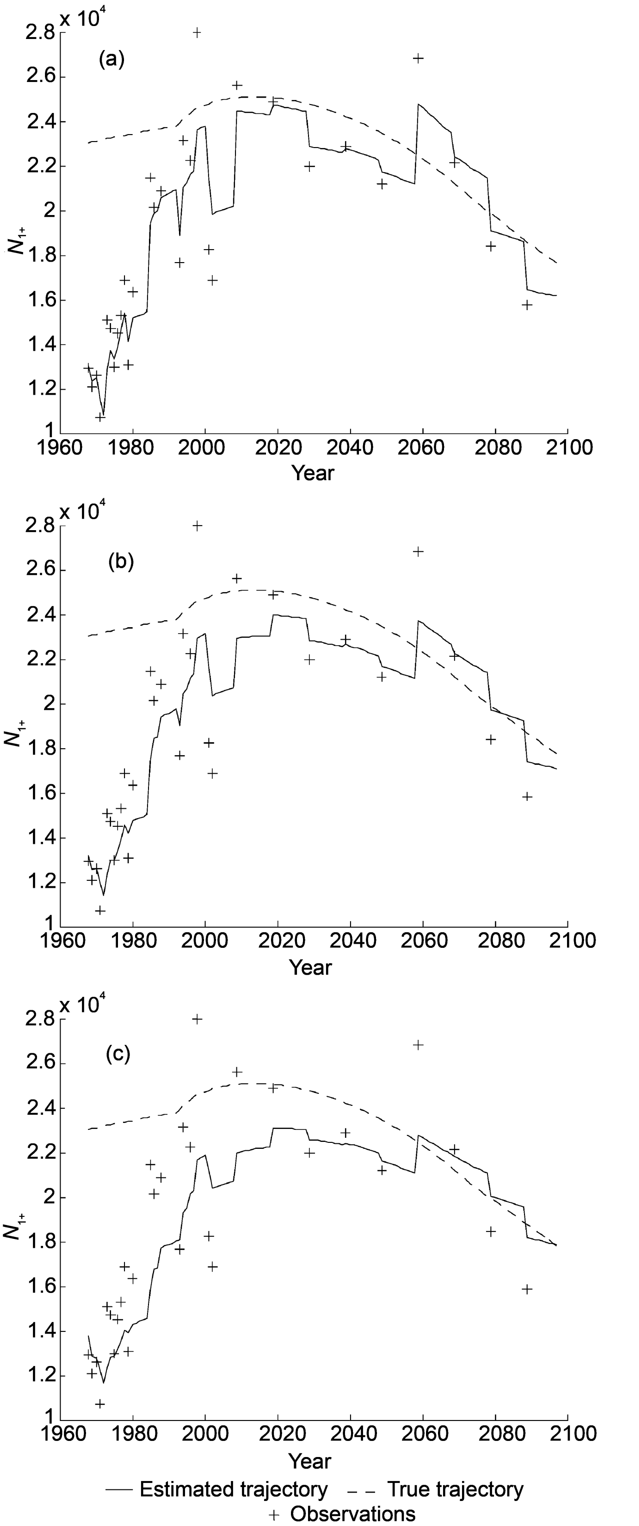

Fig. 3. Estimated and true stock trajectories from 1968 to 2100, along with observations for one simulation of trial GE16. In figure (a) CV_{add}=0, (b) CV_{add}=0.1 and (c) CV_{add}=0.2.

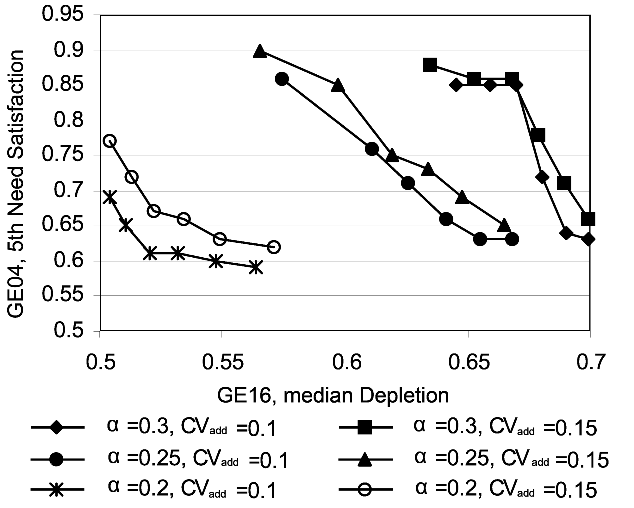

Fig. 4. 5th percentile Need satisfaction in GE04 and median Depletion in GE16 for three values of α and two values of CV_{add}. To construct each line γ ranges from 0.3 to 0.8 in increments of 0.1 and decreases from left to right for each pair of α and CV_{add}.

parameters, giving a point in the Depletion-Need plane lying both to the right and higher than a point corresponding to a different triple is clearly preferable since performance is better on both statistics.

Since the goal is to maximise both the depletion and need satisfaction statistics the most desirable tuning will provide results in the upper right hand corner of the figure. Two features are apparent in this figure. Firstly, tunings with CV_{add}=0.15 generally outperform CV_{add}=0.10 for values of γ within the range 0.3-0.8, since the curves corresponding to the former lie completely to the right and above the latter for fixed α-values in all cases; thus an increase in CV_{add} with a fixed α will move the point in the Depletion-Need plane up and to the right (some adjustment of γ may be required to maintain supremacy). However, the difference between CV_{add}=0.10 and CV_{add}=0.15 decreases as α and γ are increased and moreover the ability to detect sudden changes in stock size, such as in trial GE45, diminishes with increased CV_{add}. To obtain a good balance a value of 0.11 was selected for CV_{add}. Secondly, increasing improves performance since the curves move up and to the right for a fixed , except that there is a slight drop in need satisfaction between α-values 0.25 and 0.3 for the highest γ-value (0.8) since the top of the curve is lower for α=0.30. It would appear from Fig. 4 that α should be taken to be as high as possible, but need satisfaction in other trials starts to deteriorate with α=0.30. Thus, a version with α set to 0.25 was selected throughout.

We only present here the results of the tuning referred to as the D-M2 tuning in IWC (2005). This specific tuning of the AKF-*SLA* is one of two components in the GUP procedure. The values of the tuning parameters were set as follows:

$$\beta=0.7;\ \gamma=0.8;\ \alpha=0.25;\ P_0=0;\ CV_{add}=0.11.$$

The depletion and need satisfaction results for this tuning are given in Table 2 for the five selected trials. A complete set of trial results for the gray whale AKF-*SLA* with this tuning and for a higher tuning also are given in IWC (2005).

Table 2

Results for a few key gray whale evaluation trials of the D-M2 tuning of the AKF-*SLA*.

	D1: Final 1+ Depletion			N9: Need Satisfaction		
	5%	Med	96%	5%	Med	96%
GE01	0.854	0.885	0.911	1	1	1
GE04	0.881	0.898	0.922	0.86	0.99	1
GE10	0.902	0.914	0.924	1	1	1
GE16	0.425	0.548	0.695	0.92	0.94	0.96
GE45	0.369	0.498	0.574	0.92	0.94	0.95

Sensitivity to *S*, the survival rate was investigated by looking at values 0.95, 0.96, …, 0.99. There was very little difference in the trial results, except in trial GE16, when depletion improves slightly with higher values of *S* (Table 3). The value used in the algorithm (*S*=0.97) is somewhat arbitrary, but is similar to that estimated by Wade (2002).

Table 3
Results for trial GE16 using five different values for survival rate, *S*.

	D1: Final 1+ Depletion			N9: Need Satisfaction		
S	5%	Med	96%	5%	Med	96%
0.95	0.411	0.536	0.694	0.93	0.95	0.96
0.96	0.419	0.539	0.694	0.92	0.94	0.96
0.97	0.425	0.548	0.695	0.92	0.94	0.96
0.98	0.430	0.555	0.696	0.92	0.94	0.96
0.99	0.433	0.561	0.696	0.92	0.93	0.96

Sensitivity to *q*, the size of the process error was also investigated. Reducing *q* means that greater confidence is placed in the model and the effects of the observations are consequently down-weighted and *vice versa*. The main effect on the trial results of varying *q* was that the depletion improved in the 1.5% *MSYR* trials (GE16 and GE45) as *q* was reduced. This appears counter-intuitive at first, especially for GE45, where the stock crashes at the start of management, since a lower *q*-value will make the algorithm less responsive to the observations and hence slower to react to the population crash (Fig. 5 (a-c)). However, the reason is clear from the *MSYR* charts in Fig. 5 (d-f) showing the time evolution of the median, 5^{th} and 95^{th} percentiles from the 100 replicates of estimated *MSYR* (i.e. the expected value obtained from the posterior distribution): with a low *q*-value (i.e. high confidence in the model), the trajectory does not follow the rather steep rise exhibited in the biased historical observations as closely as with a higher *q* and the estimated *MSYR* is therefore lower, resulting in lower strike limits. It is worth noting from Fig. 5b that the algorithm with the selected *q*-value (0.001) is responding reasonably well to the population crash. It was therefore felt that this value of *q* strikes a reasonable balance between the model and observations.

Estimation of MSYR

Fig. 6 shows the time evolution of the probability distribution for *MSYR* between 1968 and 2002 (i.e. that based on the historical observations) and Fig. 7 shows the posterior probability distribution in 2002 for *MSYR* and *MSY*. Note that the probabilities of the *MSYR* values illustrated in Fig. 6 and those given in each column in Fig. 7 are the marginal probabilities, obtained by integrating over the *MSY* values (shown on the horizontal axis in Fig. 7). Based on the historical observations nearly all the posterior probability mass in 2002 is concentrated at *MSYR* 2% and 3%. It is worth pointing out that the probabilities for *MSYR* of 1% and 6% are practically zero as are probabilities for *MSY* higher than 1000. This confirms that the selected range of the parameter grid is sufficiently wide.

Fig. 8 shows the time evolution of the median, 5^{th} and 95^{th} percentiles from the 100 replicates of estimated *MSYR* over the subsequent 100 years for the five key trials GE01, GE04, GE10, GE16 and GE45. The *MSYR* estimate at the beginning of management is 2.5%. For GE01, which is a 3.5% trial the median *MSYR* declines to approximately 2.0%; for GE04, another 3.5% trial, but with a negative future bias, the medium *MSYR* rises slightly initially but levels off slightly above 2.5%; for the 5.5% trial GE10, the median *MSYR* stays level at 2.5% and for the 1.5% trials, GE16 and GE45 the median declines slightly to somewhere between 2.0 and 2.5%. It is evident that the algorithm is not particularly successful in picking up the true *MSYR* value.

Bias filters

The observations in some of the trials are biased (see Table 1). We carried out some explorations into the possibility of using filters with a bias, i.e. modifying the observation equation (2) as follows:

$$N_t^{obs} = B_t N_t e^{v_t}$$

where *B* is a possibly time-varying bias factor. We first added filters with a time-increasing historical bias, as in trials GE16 and GE45, but no future bias, thus using *MSYR* values 1%, 2%, …, 6% with and without a historical bias and thereby doubling the number of filters. This did not improve the depletion results in the two aforementioned trials, quite the contrary (in fact, median depletion went down to 0.454 and 0.428 in GE16 and GE45 respectively), the reason being that the bias filters with high *MSYR* values end up with most of the posterior probability mass in 2002, thus leading to higher future strike limits. Judging from the posterior values, it would thus appear that the best fit to the historical data with a simple density dependent model is for (time-increasing) biased observations and high *MSYR* values (median value 4.7% in 2002). We also looked at bias filters with a future bias as in GE04; i.e. added filters with a time-decreasing (1-0.5) and time-increasing (1-1.5) bias in the first 25 years of management, thus tripling the number of filters. This did not improve the results and will not be discussed further here. Additionally the use of future biases can be questioned since the possible scenarios are innumerable and one might be tempted to imitate the trials in order to improve performance. The conclusion is therefore that the addition of bias filters to the present version of the AKF-*SLA* is not a desirable option.

DISCUSSION

The present AKF-*SLA* for the ENP stock of gray whales has evolved from the version designed for the BCB stock of bowhead whales (Dereksdóttir and Magnússon, 2003). There are however a number of differences, of which the most notable are summarised below.

(1) The bowhead whale version uses a grid in the (*MSYR*, N_∞) parameter space, whereas the gray whale version uses an (*MSYR*, *MSY*) grid for the reason given above (i.e. the high negative correlation between *MSYR* and N_∞).

(2) The filters in the bowhead whale version were started in 1848 since it was assumed that the stock was at carrying capacity at that time. The variance associated with N_∞ was set to zero since (*MSYR*, N_∞) are simply points in the parameter grid which are given posterior probability values by the Bayesian methods described above, as abundance observations become available. The first update is made in 1978, the year of the first bowhead

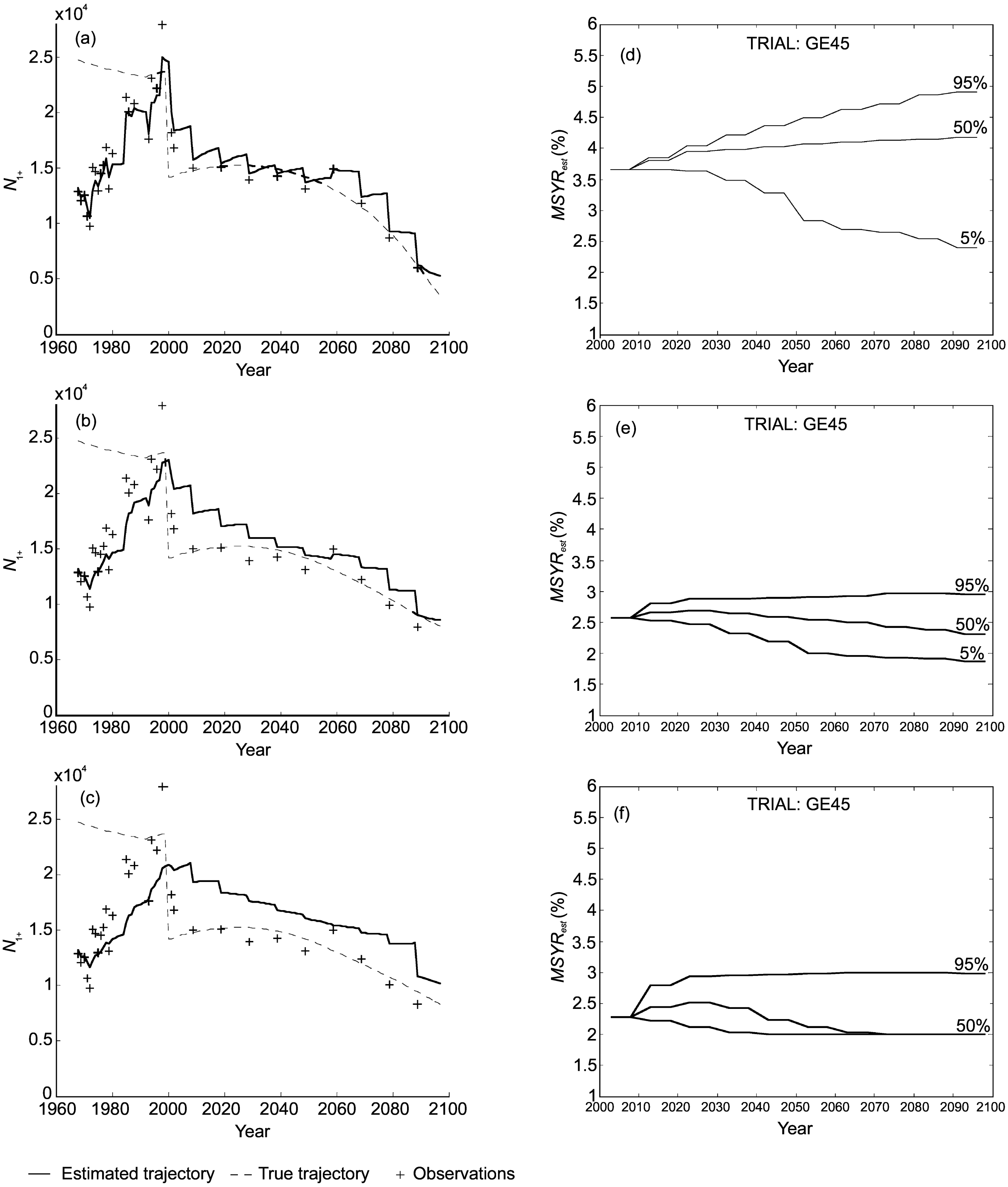

Fig. 5. Estimated and true stock trajectories from 1968 to 2100, along with observations for one simulation of trial GE45 are shown in the left column for three different values of q_t. In the column on the right the median, 5th and 95th percentiles for estimated *MSYR* are shown for the same three values of q_t. In figures (a) and (d) q_t=0.01, (b) and (e) q_t=0.001 and (c) and (f) q_t=0.0001.

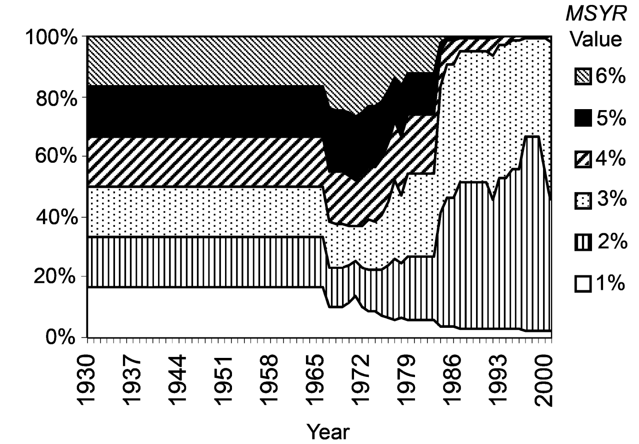

Fig. 6. The time evolution of the marginal probability distribution for each group of *MSYR* filters from 1930 to 2002.

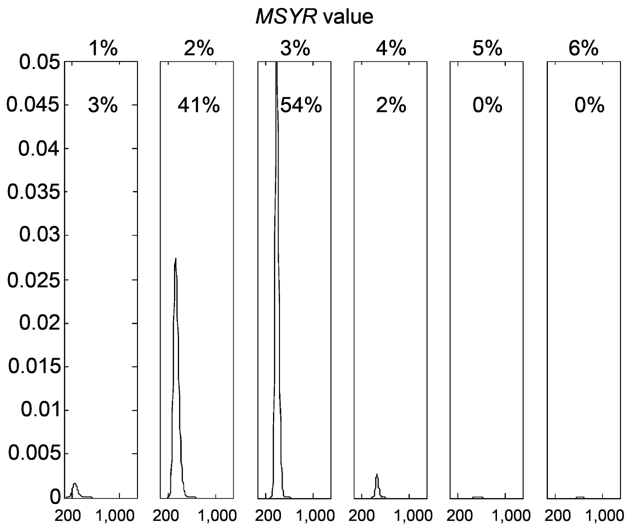

Fig. 7. The posterior probability distribution for *MSYR* and *MSY* in 2002 at the beginning of management. The numbers in each column are the probabilities of the six *MSYR* values.

abundance estimate. In the gray whale *SLA* it is not possible to start at a given year under the assumption that the stock is in its pristine state at that time for reasons given above. The first gray whale abundance estimate is from 1968, but to avoid forcing all trajectories through this estimate, the set of filters was started somewhat earlier, i.e. in 1930. Instead of using the 1930 population size as the starting value, the population is normalised by the carrying capacity N_∞, i.e. the 1930 population size N_{1930} is defined by a tuning parameter α, where $N_{1930} = \alpha N_\infty = \alpha(MSY/(0.6MYSR))$ with an associated *CV*, P_0, which is also used as a tuning parameter. The first update is made in 1968.

(3) The total number of filters in the bowhead whale version is 917 (7 *MSYR* values: 1%, 1.5%, 2%,…,4% and 131 values of N_∞ (from 10,000 to 23,000 in increments of 100). The number of filters in the gray whale version is slightly higher, 1,044. This is mainly due to the larger parameter range in the latter version. Note that the *MSYR* grid is coarser for the gray whales. This is mainly for computational reasons and the relative smoothness of the cumulative distribution function shown in Fig. 2 confirms that the grid is sufficiently fine.

(4) The estimated *CV*s in the abundance estimate, CV_{est}, provided to the *SLA* are used unchanged in the bowhead version. However, the historical observations of the abundance of gray whales with the provided *CV*s are not compatible with a standard density dependent population model. This *CV* is therefore likely to be an underestimate of the true *CV* in the abundance estimate. A constant CV_{add} is therefore added to all *CV* estimates (historical and future) provided to the *SLA*. This value is treated as a tuning parameter.

(5) The tuning in the bowhead *SLA* is two dimensional, the two parameters being β, the steepness of the catch control law, and γ, the percentile in the cumulative distribution function for the conditional strike limits. The tuning of the gray whale *SLA* is more flexible, with three additional tuning parameters (see the results section).

This list indicates how the AKF-*SLA* could be modified to apply to other aboriginal type 2 fisheries, i.e. changing the parameter grid, using different starting conditions for the set of filters, changing the tuning parameters, etc.

It is clear from Fig. 8 that the algorithm is not particularly successful in obtaining an estimate of *MSYR*. There is little difference between the *MSYR* estimates for GE01, GE10 and GE16, which are trials with *MSYR*, 3.5%, 5.5% and 1.5% respectively and the estimates change very little during the management period. This is to some extent due to the addition of CV_{add} but also due to the time-increasing bias in GE04 and GE16. However, it would also appear (at least by looking at one replicate in each of the trials GE16 and GE45 and also confirmed in other replicates) that the algorithm is tracking the true trajectory reasonably well and responding to the observations, but not unduly because of the (fairly) high value of CV_{add} (Fig. 3).

In addition to the percentile γ, α – the stock in 1930 as a fraction of N_∞ – is the parameter to which the trial results are most sensitive (Fig. 4). The reason is that the estimate of *MSYR* in 2003 decreases as α increases making the algorithm more conservative. It is interesting however, that depletion in trial GE16 and need satisfaction in trial GE04 both increase with increasing α.

Other variants of the AKF-*SLA* were also investigated. Firstly, a variant where N_{1930} is used as a tuning parameter instead of α. Secondly, variants where N_{1930} (or α) is treated as the third parameter in a 3-dimensional grid of filters (*MSYR*,*MSY*, N_{1930}). These changes did not lead to any improvements on the trial results obtained by the version described above.

ACKNOWLEDGEMENTS

The development of this *SLA* has benefitted greatly from extensive discussions in the IWC Scientific Committee Standing Working Group on Aboriginal Whaling Management Procedures. The input from the group has been invaluable, which we gratefully acknowledge. We would also like to thank Andre Punt for a very thorough review of the manuscript and suggestions for improvements, as well as another anonymous reviewer. We also wish to thank Cherry Allison for all her work in setting up and running the trials. Finally, we thank the IWC for a contribution from the Developers' Fund.

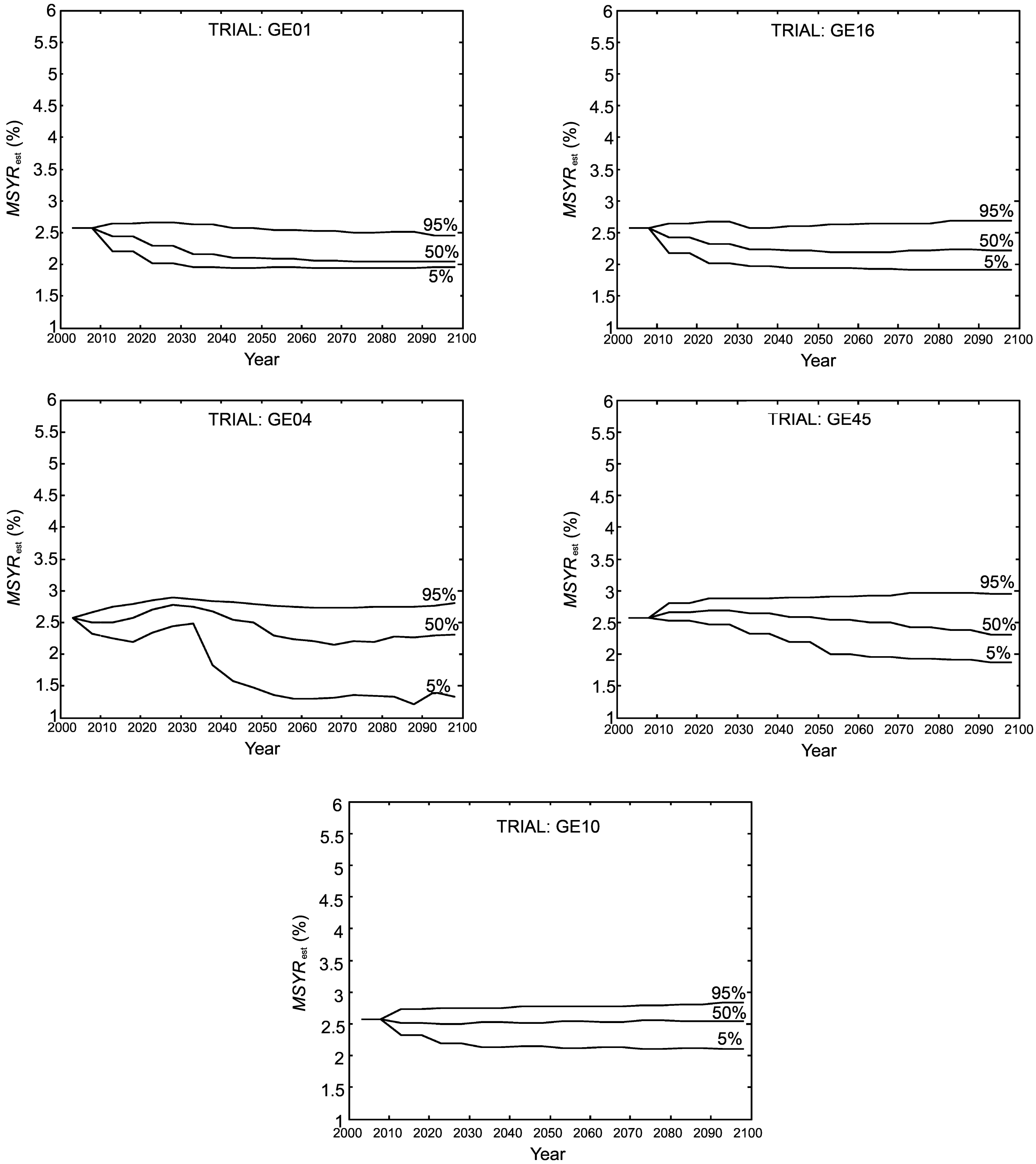

Fig. 8. Estimated *MSYR* results. The median, 5th and 95th percentiles of estimated *MSYR* for five key trials for the D-M2 tuning.

REFERENCES

Buckland, S.T. and Breiwick, J.M. 2002. Estimated trends in abundance of eastern Pacific gray whales from shore counts (1967/68 to 1995/96). *J. Cetacean Res. Manage.* 4(1):41-8.

Butterworth, D.S., Borchers, D.L. and Punt, A.E. 2002a. Dynamic response analysis for the eastern North Pacific gray whale population: An alternative approach. *J. Cetacean Res. Manage.* 4(1):77-83.

Butterworth, D.S., Korrûbel, J.L. and Punt, A.E. 2002b. What is needed to make a simple density-dependent response population model consistent with data for the eastern North Pacific gray whales? *J. Cetacean Res. Manage.* 4(1):63-76.

Dereksdóttir, E.H. and Magnússon, K.G. 2001. Investigations of an Aboriginal Whaling Management Procedure using Adaptive Kalman Filtering (AKF). *J. Cetacean Res. Manage.* 3(2):109-16.

Dereksdóttir, E.H. and Magnússon, K.G. 2003. A strike limit algorithm based on adaptive Kalman filtering with an application to aboriginal whaling of bowhead whales. *J. Cetacean Res. Manage.* 5(1):29-37.

International Whaling Commission. 1993. Report of the special meeting of the Scientific Committee on the assessment of gray whales. *Rep. int. Whal. Commn* 43:241-59.

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

International Whaling Commission. 2003. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 5:154-255.

International Whaling Commission. 2004. Report of the Scientific Committee. Annex E. Report of the Standing Working Group (SWG) on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 6:185-210.

International Whaling Commission. 2005. Report of the Scientific Committee. Annex E. Report of the Standing Working Group (SWG) on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 7:115-24.

Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82:34-45.

Punt, A.E., Allison, C. and Fay, G. 2004. An examination of assessment models for the eastern North Pacific gray whale based on inertial dynamics. *J. Cetacean Res. Manage.* 6(2):121-32.

Wade, P.R. 2002. A Bayesian stock assessment of the eastern Pacific gray whale using abundance and harvest data from 1967-1996. *J. Cetacean Res. Manage.* 4(1):85-98.

Witting, L. 2003. Reconstructing the population dynamics of eastern Pacific gray whales over the past 150 to 400 years. *J. Cetacean Res. Manage.* 5(1):45-54.

Date Received: February 2005
Date Accepted: May 2005