Use of Subjective Prediction in Optimal Stratified Sampling with Application to Shrimp Surveys in the Barents Sea

Alf Harbitz

Norwegian Institute of Fisheries and Aquaculture Ltd.

9291 Tromsø, Norway

Abstract

The task of applying subjective knowledge in predicting the number of trawl samples per stratum that minimises the coefficient of variation (CV) of the abundance estimator was considered. The constraint was the vessel time available. It was assumed that the strata biomass means, arbitrarily scaled, are the only unknown parameters needed to find the optimal solution. The concept of a subjective prediction distribution of the unknown stratum means was introduced. The distribution was described as person-dependent and determined based on intervals $[L,U]$ for the minimum and maximum subjectively predicted biomass values compared with the true measured values found after the predictions. The approach assumed a constant subjective confidence level defined as the probability of covering the true value in a random interval. A pilot subjective prediction experiment was conducted during the 1998 shrimp survey in the Barents Sea. Based on 62 $[L,U]$ predictions of shrimp biomass in the next trawl haul combined with the true biomass, the subjective prediction distribution for the cruise leader was estimated. The distribution was applied to the stratum predictions for the next survey. 10 000 random predictions of true strata means were simulated from the distribution. For each simulation, CV values of the abundance estimator were estimated based on relative strata means predicted from historical data as well as the subjective predictions. A significant CV reduction was obtained based on a combination of subjective prediction and historical data, compared to the use of historical data alone.

Key words: abundance, adaptive sampling, Barent Sea, Bayesian analysis, optimal stratified sampling, shrimp, subjective prediction

Introduction

To get as precise an estimate as possible of shrimp abundance in large areas, effective sampling strategies are of vital importance due to limited vessel time. The case where the study area is divided into predetermined geographical areas (strata) was considered here. The problem was to predict the number of samples in each stratum that minimises the coefficient of variation (CV) of the abundance estimator of biomass in the entire study area, combining historical data and subjective prediction. In each stratum it was assumed that the trawl samples are independent observations that provide unbiased estimates of the unknown sample stratum mean and variance.

If the cost of providing a sample is neglected, a well-known result in stratified sampling is that the optimal number of samples in a stratum is proportional to stratum area and stratum standard deviation (Cochran, 1977, Thompson, 1992). This solution is often denoted as Neyman allocation, after Neyman (1934). For trawl surveys in large areas, the stratum-dependent sailing time between trawl stations is the main cost. In this case the optimal solution is not given by an explicit "plug-in" formula, but must be found by numerical methods (Harbitz *et al.*, 1998). It turns out, however, that the optimal solution in this case deviates negligibly from the Neyman allocation unless there are extreme relative differences between strata areas and/or between stratum standard deviations (Harbitz *et al.*, 1998).

Survey data supported the assumption that the strata standard deviations are proportional to the strata means with a proportionality constant independent of strata and year. In this case the parameters needed to determine the Neyman allocation are the relative strata means, i.e. the stratum mean in each stratum divided by the sum of strata means in all strata. If the estimated

relative stratum means appear to be quite similar from survey to survey, relative stratum means in future surveys may be predicted from historical data even if abundance varies between surveys. If scientists or others have qualified reasons to doubt that we can predict the relative strata means from past data, due to environmental changes or other reasons, the challenge faced is how to utilise such subjective knowledge to improve these predictions. A major goal of this paper is to outline a method to quantify the person-dependent ability of subjective prediction in statistical terms.

The approach in this study is to establish a link between subjective data and real measurements of biomass, based on subjective prediction experiments. Each subjective observation simply consists of a minimum, a most likely and a maximum predicted biomass value for a future parameter, e.g. the shrimp biomass in the next trawl haul or the stratum mean. A major assumption is that the subjective confidence level, i.e. the expected relative proportion of the data where the true parameter value falls within the interval limits, is a person-dependent constant independent of stratum and survey. The better subjective knowledge, the smaller interval width is predicted.

Based on data from experiments as described above, the concept of a subjective prediction probability distribution of the unknown parameter is introduced, where the term probability has a classical frequentistic interpretation. Based on subjective predictions for the strata means in the next survey, along with historical trawl data, a combined prediction for the unknown relative strata means is constructed. When the combined predictions are established, the corresponding optimal number of trawl stations in each stratum are calculated.

In order to assess the effect of applying subjective prediction, simulations based on a case study with real historical trawl data and subjective prediction data from shrimp surveys in the Barents Sea were performed. Based on subjective prediction intervals for a future survey, along with an established normal subjective distribution fitted to the intervals, the future values of true relative strata means were simulated. For each simulation, the CV of the abundance estimator was calculated based on: 1) proportional allocation (non-stratified sampling), 2) optimal sampling based on historical data, 3) optimal sampling based on historical data and subjective prediction, and 4) optimal allocation based on the true relative strata means. A comparison of the different CV distributions from the simulations were used in order to study the effect of the 3 first approaches, which are all applicable in

practice then compared to the CV distribution obtained from optimal allocation based on true strata means.

The main focus is optimal allocation, but the concepts developed can easily be extended to a Bayesian framework for inference about the unknown abundance, by applying the subjective prediction distribution as an *a priori* distribution. After the data from the survey are available, we can construct the *a posteriori* distribution of abundance conditional on trawl data, and e.g. calculate a credibility interval of specified level for the unknown abundance. Because the prior has a frequentistic interpretation based on data, contrary to the vague concept of subjective probability often used in Bayesian analysis, it is now meaningful to define the bias of the predicted parameter with respect to the prior. If it is unbiased, the posterior mean will be unbiased as well. Bayesian inference (Carlin and Louis, 1996) is becoming increasingly popular within fisheries science (McAllister and Kirkwood, 1998), not least due to the accelerating development of computer-efficient Markov Chain Monte Carlo techniques (Gilks, 1996). This makes the approach attractive also in cases where the subjective prediction distribution is used as a non-conjugate prior.

The subjective prediction approach can also be applied in an adaptive setting (Thompson and Seber, 1996), where new subjective predictions are made before each new stratum is reached, and an optimal reallocation of effort within the remaining strata is performed.

Material and Methods

Notations

i	=	subscript for stratum, $i = 1,..., m$
j	=	subscript for survey, $j = 1,...,n_{yr}$
l	=	summation index over strata
m	=	number of strata
Y	=	biomass in kg from trawl haul
$\overline{Y}$	=	empirical stratum mean of Y-values
m	=	stratum mean, EY
s	=	stratum standard deviation, std(Y)
s	=	1) empirical standard deviation, 2) subscript for subjective
t	=	available vessel time
t_0	=	trawl and handling time at each station
v	=	vessel sailing speed between trawl stations
$[L,U]$	=	subjective biomass prediction interval
M	=	modal subjective predicted value
α	=	lower subjective probability = $P(\mu < L)$
β	=	upper subjective probability = $P(\mu > U)$

ε	=	interval factor in the relationship $(U-L) = \varepsilon \cdot (U+L)/2$
r	=	subscript for relative
k	=	proportionality factor in relation $\sigma = k\mu$
Q	=	chi-square sum
Z	=	standardised subjective prediction variable = $(\mu - (L+U)/2)/(U-L)$
f_s	=	subjective prediction distribution of μ
N_i	=	number of samples in stratum i
n_j	=	total number of samples in survey j
A_i	=	area of stratum i
A	=	ΣA_i = total area of all strata
A_0	=	area covered by one trawl haul
μ_A	=	true abundance in total area
h	=	subscript for historical
*	=	superscript for estimator

The survey data

The shrimp biomass data were from annual trawl surveys in the Barents Sea in the period 1992–99. The area was divided in $m = 6$ fixed geographical areas, or strata (see Fig. 1 and lower row of Table 1). The total number of trawl stations (biomass observations) varied between a minimum of 92 (1999) to a maximum of 138 (1996). The minimum number of observations at one single stratum was 7, while the maximum was 59 (Table 1). The estimated stratum means and standard deviations are given in Table 2. Aschan and Sunnanå (MS 1997) give a description of survey design and trawl technology.

In 1998 a pilot subjective prediction experiment was conducted where each of 4 persons reported a minimum value, L, a most probable (modal) value, M, and a maximum value, U, for the biomass they predicted in the next trawl haul. The 4 persons included two scientists (the cruise leader and another biologist), the captain and the chief mate. All persons predicted independent of each other. As a basis for their prediction the biomasses from the previous haul and the previous year were easily available. Of particular interest was the cruise leader data, because this person had the authority to determine the effort allocation. A sample size of 62 (L,M,U) was reported by the cruise leader, and a similar number for the other persons.

The optimal number, N_i, of trawl stations per stratum

Let μ_A^* denote the abundance estimator for shrimp biomass abundance, μ_A, in the entire study area:

$$\mu_A^* = \sum_{i=1}^{m} \overline{Y}_i \cdot (A_i / A_0) \qquad (1)$$

where $\overline{Y}_i$ is the biomass sample mean from the N_i trawl stations in stratum i, A_i is the area of stratum i, and A_0 is the standardized area covered by a trawl station. All biomass samples were assumed to be stochastically independent.

The results in Harbitz *et al*. (1998) are applied in order to determine the N_i-value in each of m predetermined strata that minimise the CV of the abundance estimator for the entire area under the constraint of a given vessel time, t, available. For a given abundance, μ_A, this is synonymous with minimising the variance

$$\text{var}\left(\mu_A^*\right) = \sum_{i=1}^{m} \sigma_i^2 \cdot \left(A_i / A_0\right)^2 / N_i \qquad (2)$$

where σ_i^2 is the stratum variance, i.e. the variance of biomass from a random trawl haul in stratum i. The general problem in stratified sampling is that the stratum variances are not known. It was assumed that the optimal solution deviates negligibly (Harbitz, *et al*., 1998) from the Neyman-allocation $N_i \propto A_i\sigma_i$. Further, a proportional relationship $\sigma = k\mu$ between stratum mean and standard deviation is found to be reasonable (see Fig. 2), where k is independent of stratum and year. The optimal solutions are then
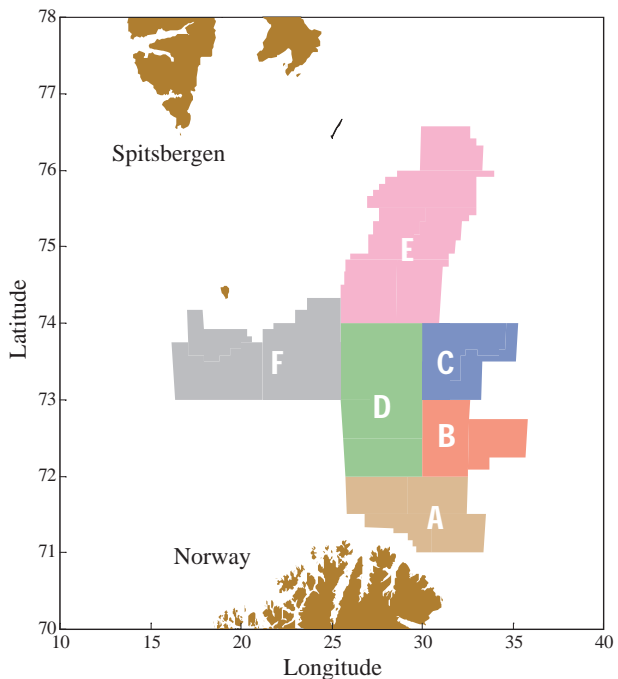

Fig. 1. The strata for the shrimp surveys in the Barents Sea 1992–99.

TABLE 1. Number of samples in each stratum. Barents Sea shrimp surveys 1992–99. Stratum areas in 1 000 naut. m² are shown.

Year	Stratum					
	A	B	C	D	E	F
1992	16	12	10	24	33	37
1993	22	11	12	26	23	33
1994	11	10	12	22	24	22
1995	14	8	8	22	29	22
1996	11	9	8	29	59	22
1997	14	7	9	10	40	15
1998	16	8	10	17	29	29
1999	8	11	14	14	30	15
Area [1 000 naut. m²]	6.73	4.71	4.00	9.34	11.48	9.58

TABLE 2. Sample strata means ($\mu*$) and standard deviations ($\sigma*$) in kg per trawl haul from Barents Sea shrimp surveys 1992–99. The standard deviations equal the square root of sample variances with n-1 as denominator.

Year		Stratum					
		A	B	C	D	E	F
1992	$\mu*$	21.8	37.1	56.8	28.6	65.7	28.7
	$\sigma*$	15.9	23.0	33.5	24.4	28.9	20.0
1993	$\mu*$	14.5	26.2	45.2	17.0	47.7	10.8
	$\sigma*$	12.9	13.1	33.3	12.4	26.3	6.8
1994	$\mu*$	19.9	11.0	13.3	8.8	30.4	13.1
	$\sigma*$	16.6	6.5	8.5	7.8	22.0	13.5
1995	$\mu*$	9.7	11.5	18.0	17.6	40.8	15.1
	$\sigma*$	6.5	9.5	20.1	19.9	29.7	14.6
1996	$\mu*$	23.0	10.8	32.5	24.3	68.6	28.1
	$\sigma*$	22.7	6.3	17.8	21.9	37.0	19.2
1997	$\mu*$	12.5	41.0	25.4	27.4	50.7	21.3
	$\sigma*$	12.5	36.7	10.5	26.1	47.4	14.3
1998	$\mu*$	16.9	39.4	53.7	23.3	81.4	58.1
	$\sigma*$	14.7	28.9	37.0	24.0	51.5	32.9
1999	$\mu*$	16.0	27.4	41.1	16.2	59.3	18.0
	$\sigma*$	10.5	21.0	26.0	15.4	41.8	13.9

proportional to the relative stratum means, $N_i \propto \mu_{ri} = \mu_i / \Sigma \mu_1$, as the only unknown parameter (see Appendix eq.(11)).

As a measure of precision the coefficient of variation, CV(μ_A^*) was applied:

$$cv(\mu_A^*) = \frac{std(\mu_A^*)}{\mu_A} \qquad (3)$$

where $\mu_A = \Sigma \mu_I \cdot (A_I / A_0)$. The equation above combined with simulations was used to study the effect of different approaches to predict the μ_{ri}-values, to be compared with the minimum CV value obtained by using true values for the μ_{ri}.

For brevity, the vector notation $\mu_r = [\mu_{r1},...,\mu_{rm}]$ is used to denote what was called the relative stratum mean vector. The main focus in this paper is to outline how reasonable predictors, μ_r^*, for μ_r can be constructed based on a combination of historical data and subjective predictions.

The historical predictor, μ_{rh}

For a given survey (year), the ratios $\mu_{rih}^* = \overline{Y}_i / \Sigma \overline{Y}_I$, $i = 1,...,m$, are the natural ratio estimators for the corresponding unknown relative strata means, μ_{rih}.
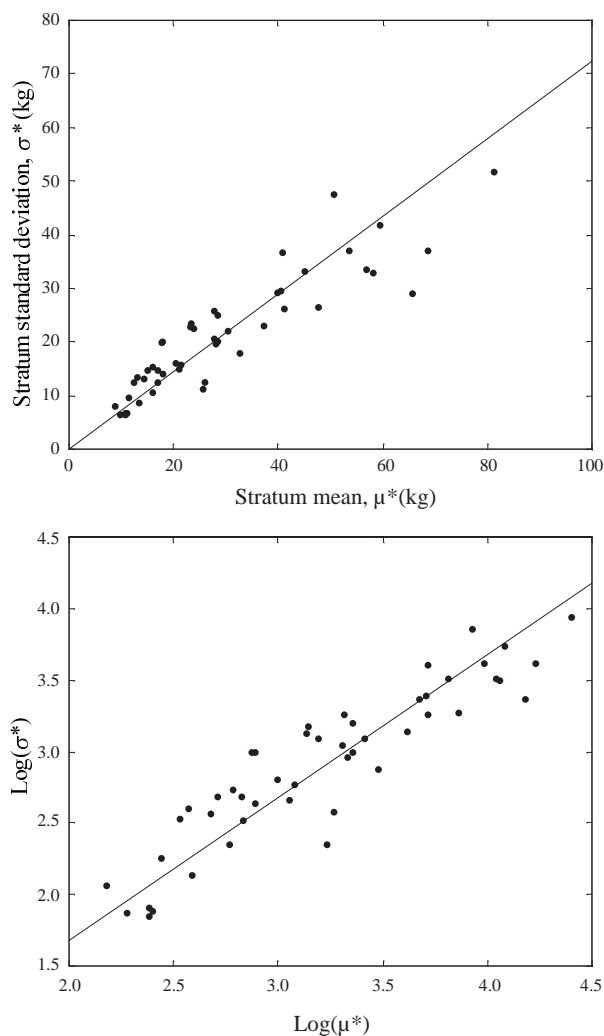
Fig. 2. The relation between empirical stratum means $\mu^* = \overline{y}_{ij}$ and standard deviations $\sigma^* = s_{ij}$ with the fitted curve $\sigma = 0.725 \cdot \mu$ from the shrimp surveys in the Barents Sea 1992–99.

In most cases these estimators have a negligible bias due to the rather large total number of samples from a survey. For the same reason, we may construct reasonable estimators, σ_{rih}^{2*}, for the variances, $\sigma_{rih}^2 = \mathrm{var}(\mu_{rih}^*)$. Some mathematical details are given in the Appendix.

If μ_r has a negligible variation over years, we may improve the precision of the empirical relative strata means, μ_{rh}^*, by applying a linear combination of these from several years. Each survey is weighted proportional to the total number of samples from that survey. In order to assess if μ_r has changed over years, a chi-square test can be applied, where one option is to include only the succeeding last years for which the test is not rejected in the predictor for the next survey.

The subjective prediction distribution for μ

In order to illustrate concepts and terminology, the attention is first restricted to an unknown stratum mean μ to be subjectively predicted. To quantify the prediction ability of a given person, the following simple and practical measures (data) are used: the minimum, L, the most probable (modal), M, and the maximum, U, predicted values for μ are reported. The subjective prediction distribution introduced, $f_s(\mu)$, is conditional on $[L,U]$. L and U correspond to the probabilities $\alpha = P(\mu<L)$ and $\beta = P(\mu>U)$, which define the subjective confidence level, $1-\alpha-\beta$. Two basic assumptions are:

1. The shape of $f_s(\mu)$ and the probabilities α and β are constant for one and the same person, and independent of $[L,U]$, stratum and survey.

2. The subjective distribution for prediction of the outcome from a single trawl haul is transferable to the corresponding distribution for a stratum mean.

The first assumption can be illustrated as follows. A "risk-willing" person will tend to provide narrow intervals, corresponding to a low confidence level. Contrastingly, a "cautious" person will tend to predict large intervals with a high confidence level. In addition, a person will adjust the interval width according to strength of belief in the actual situation. The assumption can be examined by repeated subjective prediction experiments, where the "true" value (or at least an unbiased estimate) can be measured. Due to the many trawl samples during a survey, single trawl haul predictions are suitable to obtain reasonable estimates for α, β and the shape of f_s.

The second assumption implies that if it is found that e.g. a normal distribution is appropriate in the single trawl haul situation, it will also apply to the stratum prediction situation, where estimates for α and β may be applied from the single trawl haul predictions. Again, repeated experiments can be used to assess this assumption.

Motivated by single trawl haul predictions and true values (Fig. 3), it was now assumed that a normal distribution $N(\mu_s,\sigma_s)$ was a reasonable approach for $f_s(\mu)$, that the interval width $U-L$ was approximately proportional to the mid value $\mu_s = (L+U)/2$, and that $\alpha = \beta$. Note that for a given interval $[L,U]$ it is now meaningful to consider $f_s(\mu)$ as a probability density
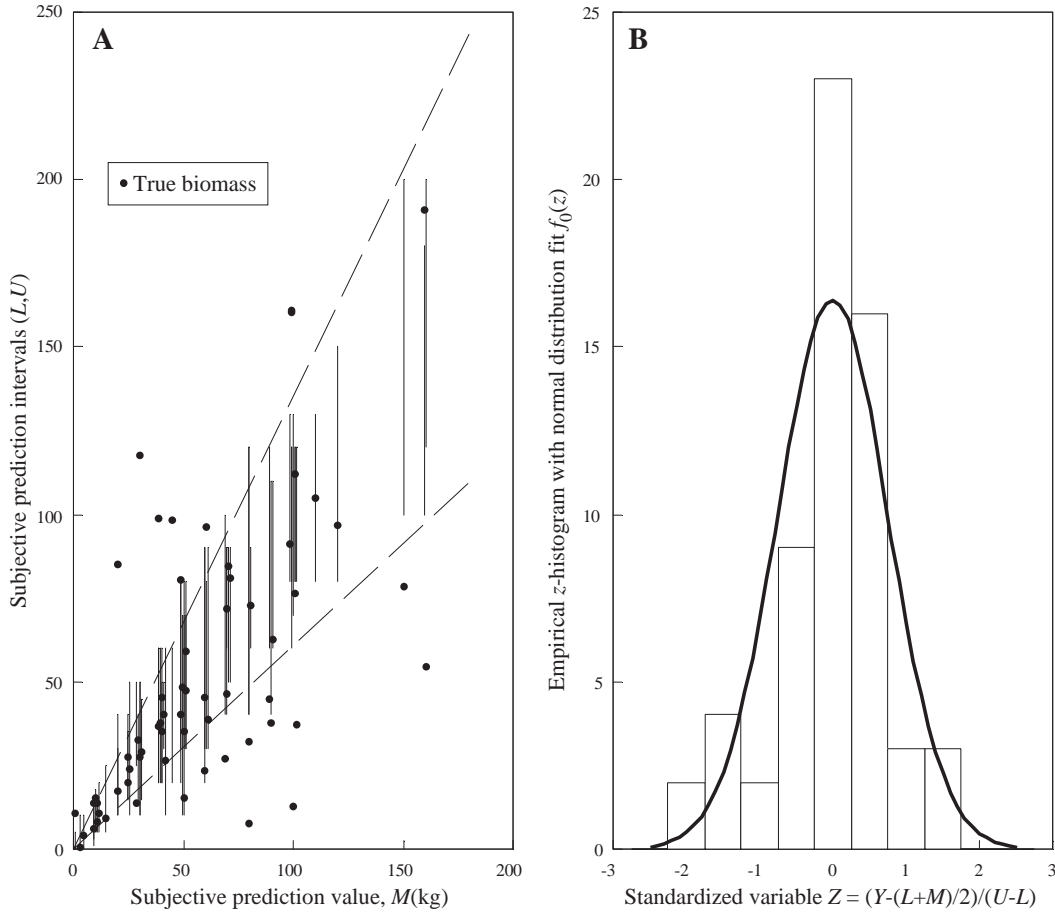
Fig. 3. **A)** Sixty-two subjective prediction intervals $[L,U]$ for biomass in next trawl haul as a function of corresponding modal predicted value, M, for cruise leader in the 1998 subjective prediction experiment. The corresponding true biomasses are plotted as dots. **B)** Histogram of standardised variable z based on data in left panel, with fitted normal distribution (see text for description).

function of μ in a classical frequentistic sense. Over time a long series of predictions can be expected for which $[L,U]$ is constant, where the unknown μ values behave as independent stochastic variables, following the subjective prediction distribution conditional on $[L,U]$. As a consequence of these assumptions, the interval mid value μ_s in addition becomes an approximately unbiased estimator for μ.

The subjective relative prediction vector is

$$\mu_{rs} = [\mu_{r1s},...,\mu_{rms}] = [\mu_{s1},...,\mu_{sm}]/\Sigma\mu_{sl} \qquad (4)$$

where subscript s denotes subjective, and μ_{si} is the mid value of the subjective prediction interval $[L_i,U_i]$ for μ_i.

It is assumed that the subjective prediction intervals for different strata are uncorrelated. For sufficiently narrow intervals relative to the corresponding mid values, approximate variance expressions σ_{ris}^{2*} for $var(\mu_{ris})$ analogous to the historical predictor case can be constructed (see appendix for mathematical details).

The predicted optimal number of trawl stations in each stratum

It is now possible to outline how μ_r is predicted by a combination of historical data and subjective prediction. The elements in the combined relative prediction vector $\left[\mu_{r1C}^{*},...,\mu_{rmC}^{*} \right]/\Sigma_l \mu_{rlC}^{*}$ are determined as a linear combination of μ_{ris} and μ_{rih}^{*}, $i = 1,...,m$, where subscript C denotes "combined". The weights are based on the involved variance estimates, i.e.

$$\mu_{riC}^{*} = K \cdot \left(\frac{\sigma_{ris}^{2*}}{\sigma_{ris}^{2*}+\sigma_{rih}^{2*}} \cdot \mu_{rih}^{*} + \frac{\sigma_{rih}^{2*}}{\sigma_{ris}^{2*}+\sigma_{rih}^{2*}} \cdot \mu_{ris} \right),$$

$$i = 1,...m \qquad (5)$$

where K is a normalising constant to ensure that $\Sigma\mu_{riC}^* = 1$ and the other elements are defined previously. If more than one survey is involved in the historical predictor, the corresponding variance estimators are a linear sum of the variances from each survey, with weights equal to $n_j/\Sigma n_j$, i.e. proportional to the number of samples in survey j. Once the elements of the combined relative prediction vector are determined, eq. (11) in the appendix is applied with μ_{ri} replaced by μ_{riC}^* in order to estimate the optimal number of stations within each stratum. Note that if the variances involving subjective prediction are much larger than the variances based on historical data, the combined predictor is close to the historical predictor, and *vice versa*.

Simulation of the CV reducing effect of subjective prediction

The goal of introducing subjective prediction in the context that were outlined is to reduce the CV of the abundance estimator, $CV(\mu_A^*)$. In order to study the effect of μ_{rC} simulations are performed based on a case study where μ_r has changed from μ_{rh} in the past to μ_{r0} in the future situation where the survey is conducted. It is assumed that the subjective prediction distribution is normal with the true mean. The unstratified situation (proportional allocation) was also included, $N_i \propto A_i$, corresponding to a relative stratum mean vector $\mu_{rA} = [1/m,...,1/m]$, as a reference. The simulation then proceeds as follows (see Appendix for mathematical details):

1. Choose an appropriate stratum relative mean vector $\mu_{rh}^* = [\mu_{r1h}^*,...,\mu_{rmh}^*]$ based on historical data, and calculate the corresponding variance estimates, σ_{rih}^{2*}, for each stratum, i, see eq.(13).

2. Establish $[L_i, U_i]$-data based on subjective prediction, and calculate the subjective prediction vector $\mu_{rs} = [\mu_{r1s},...,\mu_{rms}]$. Estimate the corresponding subjective variances, σ_{ris}^{2*}, for each stratum i, see eq.(20).

3. Determine the combined relative vector μ_{rC} from eq.(5) based on step 1 and 2 and determine the predicted optimal number of stations in each stratum, see eq.(11).

4. Simulate a random true stratum relative mean vector μ_{r0} from the subjective prediction distributions in step 2.

5. Repeat step 4 $n_{boot} = 10\ 000$ times, say, and compare the distributions of $CV(\mu_A^*)$ based on μ_{rA}, μ_{rh}, μ_{rC} and μ_{r0}.

Results

The relationship between the empirical strata means and strata standard deviations are shown in Fig. 2, along with the fitted line $\sigma = k\cdot\mu = 0.725\cdot\mu$.

The hypothesis of a constant μ_{rh} value in the period 1992–99 estimated by applying eq. (16) and a simultaneous 5% test level was rejected for the years 1993 and 1996. Therefore the three last years were used to estimate μ_{rh}, providing the vector

$$\mu_{rh} = [0.073, 0.175, 0.190, 0.109, 0.305, 0.148]\ \ (6)$$

All 4 persons who participated in the subjective prediction experiment appeared to provide intervals which tended to have widths $(U - L)$ proportional to the mid-value $(L + U)/2$. The cruise leader appeared to have the best result in terms of lowest interval factor ($\varepsilon* = 0.74$) combined with the largest subjective confidence level (66%). When translating her $[L_i, U_i]$ intervals by replacing the M_i values with the previous haul value, Y_i-1, or the last year value, $Y_{i,97}$, the subjective confidence level was reduced to 44% and 52%, respectively. This is a promising indication of the potential in applying subjective prediction in this case.

Attention is now restricted to the cruise leader results and the application of these. Based on her 62 (L, M, U)-values the following results were found:

1. $\alpha^* = 0.19$, $\beta^* = 0.15$

2. $U - L \approx 0.74\cdot(L+U)/2$

3. Negligible bias, mean $((Y - (L+U)/2)/(U - L))= -0.004$

4. Negligible deviation between M-values and interval mid values.

The intervals as a function of the modal values, M, are shown in the left panel of Fig. 3, along with the true y-values. In the right panel of Fig. 3 the estimated standardised subjective distribution of $z = (y-(L+U)/2)/(U-L)$ with a fitted normal distribution is shown. Because $\alpha*$ is close to $\beta*$, we let $\alpha = \beta = .17$ in the simulation of the effect of applying a normal subjective prediction distribution in a new survey. The standard deviations σ_{is} to be used are then

$$\sigma_{is} = \frac{\varepsilon_i(U_i + L_i)/2}{2\Phi^{-1}(0.83)} = 0.262\cdot\varepsilon_i(U_i + L_i),$$

$$i = 1,...,m \tag{7}$$

where L_i and U_i are the subjective lower and upper borders of the predicted stratum means and Φ^{-1} denotes the inverse cumulative standard $N(0,1)$ distribution. Based on subjective predictions by the cruise leader, the following relative mean vector was used:

$$\mu_{rs} = [0.093, 0.155, 0.207, 0.078, 0.363, 0.104] \quad (8)$$

with corresponding ϵ−values 0.181, 0.208, 0.250, 0.250, 0.179 and 0.488.

The magnitude of the estimated subjective variances involved appeared to be quite similar to the variances based on historical data.

Random values for the future μ value in each stratum were simulated from the appropriate normal distribution with expectation value $(L+U)/2$ and standard deviation σ_s given by eq. (19). For each simulation the CV values were calculated based on proportional allocation (μ_{rA}), predicted optimal allocation based on μ_{rh}, predicted optimal allocation based on a combination of historical data and subjective prediction (μ_{rC}) and optimal allocation based on true values (μ_{r0}). The results of 10 000 simulations are shown in Fig. 4. The time available for the survey was chosen to be 250 hr, the vessel speed between stations to be 12 naut. mile/hr and the trawling and handling time at each station was set to $t_0 = 1$ hr. As can be seen, a considerable improvement was obtained by applying subjective prediction, despite the rather small difference between μ_{rh} and μ_{rs}.

Discussion

In order to assess the properties of the estimator k^* and the assumption of a chi$^2(m)$ distributed test statistic, Q_j, a gamma distribution may be fitted to the normalized biomass values $y_{ij} / \bar{y}_{ij}$ and then simulated from the corresponding gamma distribution for $\bar{y}_{ij}$ with empirical values used as true parameters. This has in fact been done, where the gamma distribution with shape parameter $c = 1.9$ appeared to be a fairly good approximation to the normalized biomass values, not least due to a rather small number of zero observations (2.8%). The details of this exercise is omitted, partly to shorten the paper and partly because the results have no impact on the results presented in the paper.

A major basis for the approach to quantify the statistical properties of subjective prediction was the assumption of a relatively constant subjective confidence level. This means that a person who becomes more skilled with experience will tend to provide more narrow prediction intervals, while the same subjective confidence level is maintained. Only experience over time can assess the assumption of a constant subjective confidence level.

Another basic assumption was the transferability of the subjective distribution from the simple single-value prediction of Y to the prediction of the stratum mean, μ. An intuitive feeling is that if a normal subjective prediction distribution is appropriate in the single value case, the assumption of a normal distribution is also reasonable for the μ−predictions. Again, this assumption can be assessed by experiments, but a quite long time period is required. The larger number of persons who seriously perform such experiments, however, the shorter will be the time needed to make this assessment.

For simplicity independent subjective predictions between strata were assumed. If the basis for the predictions is related to e.g. migration between strata, this independence assumption is dubious. An overestimate of the emigration effect from one stratum to another will correspond to an overestimate of the immigration effect into the other. As a result, the variances based on independence might become too small, overemphasising the subjective prediction ability.

At least at an early stage many will hesitate in trusting subjective prediction too much. A general approach is then to determine some limitations on maximum influence. As an example one may limit the influence so that the combined vector μ_{rC} is not dominated by the subjective vector μ_{rs}. One way to obtain this is to scale the relative variances σ_{ris}^2 by a common factor so that none of them become smaller than the corresponding historical variances, σ_{rih}^2. In this way the relative subjective variances between strata are maintained.

It is possible to extend the use of subjective prediction as outlined here to adaptive sampling designs. One approach would be to make new subjective predictions before the first trawl haul in each new stratum, to be used in order to determine the optimal number of trawl stations in the next strata. As more strata are finished, the new biomass values can be implemented in the estimate of the historical predictor μ_{rh}. Note that this approach is not in conflict with obtaining unbiased estimates of abundance.

The concept of a subjective prediction distribution can also be extended to make inference of the abundance, μ_A. Let $[L_A, U_A]$ be an unbiased subjective
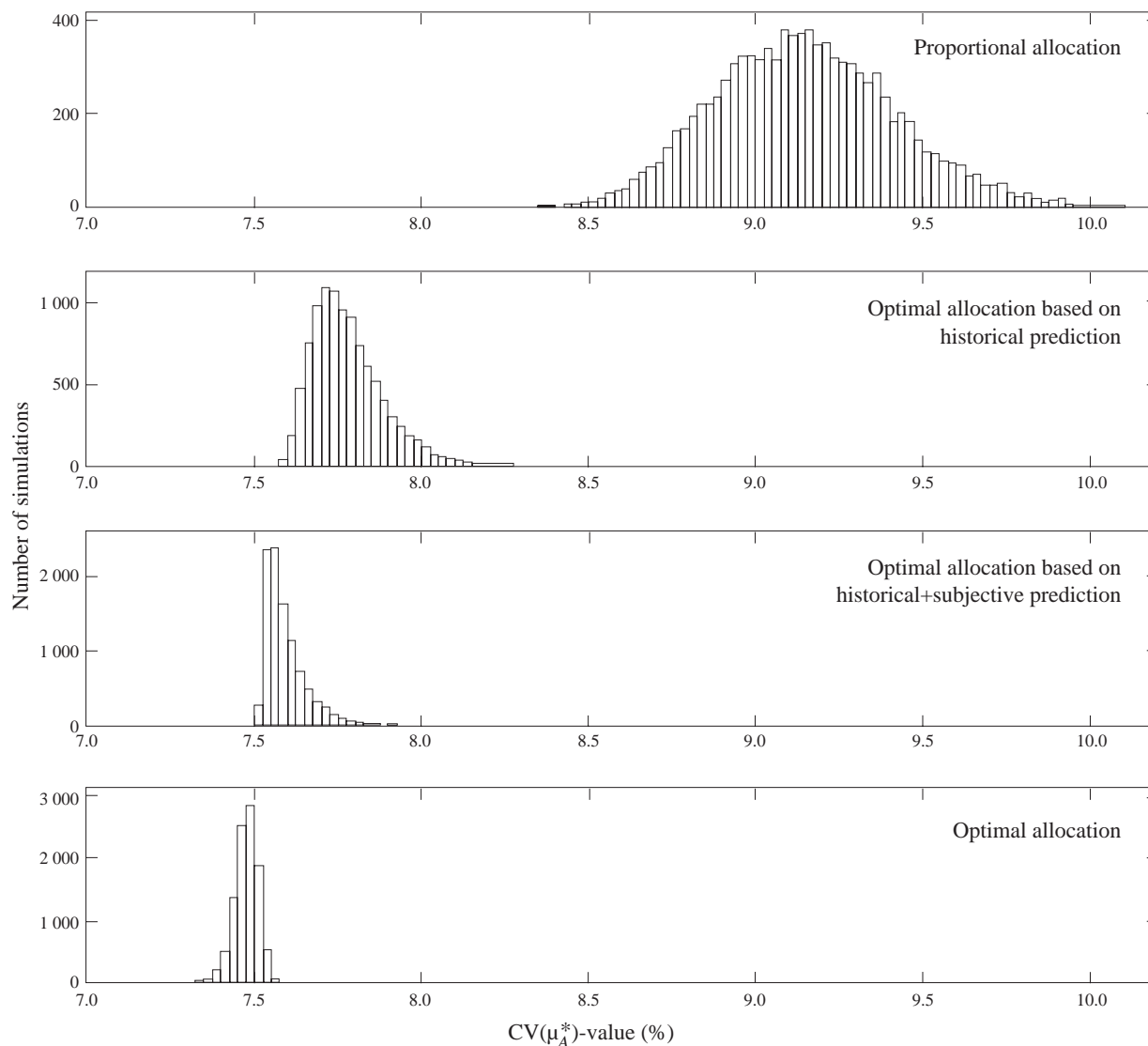
Fig. 4. Simulations of CV (m_A^*) distributions based on different predictions of relative strata mean vector m_r, based on 10 000 simulations of true strata means from cruise leader's subjective prediction distribution (see text).

prediction interval for μ_A consistent with corresponding intervals for each stratum. The subjective prediction distribution $f_s(\mu_A)$ can then be used as a prior in the Bayesian sense, and we may use e.g. a 95% posterior credibility interval for μ_A as an inference measure for μ_A. Because it is reasonable to assume that the distribution of μ_A^* based on data is normal, a normal prior will be a conjugate prior. The construction of a credibility interval is then straight forward. Due to the frequentistic interpretation of $f_s(\mu_A)$, it is now meaningful to say that the posterior expectation value is an unbiased estimator of μ_A.

In principle, the concepts outlined could also be extended to multi-species surveys. In this case, however, a more complicated criterion for optimal allocation must be defined properly, along with an appropriate incorporation of possible co-variation between samples of the different species.

Acknowledgements

First of all the results of her subjective predictions as cruise leader have motivated the belief in the usefulness of the concepts outlined. Secondly, she has contributed through fruitful discussions and careful reading of former drafts of the manuscript. Thanks also to captain Hans Hansen and chief mate Inge Berg of R/V "*Jan Mayen*", along with scientist Jorge dos Santos at Norwegian College of Fisheries, for eager participation in the pilot subjective prediction experiment in 1998. Finally, thanks to the three referees, whose constructive comments have contributed considerably to improve the paper from the earlier draft.

References

ASCHAN, M., and K. SUNNANÅ. MS 1997. Evaluation of the Norwegian Shrimp Surveys conducted in the Barents Sea and the Svalbard area 1980–1997. *ICES C. M. Doc.*, No. Y:7, 24 p.

CARLIN, B. P., and T. A. LOUIS. 1996. Bayes and empirical Bayes methods for data analysis. Chapman & Hall. 399 p.

COCHRAN, W. G. 1977. Sampling Techniques. 3rd edition. New York: Wiley.

GILKS, W. R. 1996. Markov chain Monte Carlo in practice. Chapman & Hall.

HARBITZ, A., M. ASCHAN, and K. SUNNANÅ. 1998. Optimal effort allocation in stratified, large area trawl surveys, with application to shrimp surveys in the Barents Sea. *Fish. Res.*, **37**: 107–113.

MCALLISTER, M. K., and G. P. KIRKWOOD. 1998. Bayesian stock assessment: a review and example application using the logistic model. *ICES J. Mar. Sci.*, **55**: 1031–1060.

NEYMAN, J. 1934. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *J. Roy. Statist. Soc.* **97**, 558–606.

THOMPSON, S.K. 1992. Sampling. Wiley, New York.

THOMPSON, S.K., and A. F. SEBER. 1996. Adaptive Sampling. John Wiley & Sons, Inc.

Appendix

Mathematical Details

Optimal number of trawl stations per stratum

The task considered was to minimise the abundance estimator variance:

$$\text{var}(\mu_A^*) = \sum_{i=1}^{m} \sigma_i^2 \cdot (A_i / A_0)^2 / N_i \tag{9}$$

under the constraint of a given (constant) vessel time, t, available:

$$t = \sum_{i=1}^{m} N_i t_0 + \frac{1}{v} \sum_{i=1}^{m} \sqrt{A_i N_i} \tag{10}$$

The two sums above are due to trawling and sailing between trawl stations, respectively, where t_0 is the (constant) trawl and handling time at each station and v is the (constant) sailing speed between trawl stations. The trawl stations within each stratum are assumed to be located at the intersection points in a square grid.

It turns out (Harbitz *et al*., 1998) that the optimal solutions are very close to the Neumann-allocation $N_i \propto A_i \sigma_i \propto A_i \mu_{ri}$, applying the relationship $\sigma = k\mu$:

$$N_i = \left(\frac{\Sigma_I A_I \sqrt{\mu_{rI}}}{2vt_0 \Sigma_I A_I \mu_{rI}} \right)^2 \left(\sqrt{1 + \frac{4v^2 t \cdot t_0 \Sigma_I \mu_{rI} A_I}{\left(\Sigma_I A_I \sqrt{\mu_{rI}}\right)^2}} - 1 \right)^2 \cdot \mu_{ri} A_i, \quad i = 1,...,m \tag{11}$$

The rather complicated expression in the above equation is due to the non-linear and stratum-dependent expression $t_i = \left(\sqrt{A_i / N_i} \right) / v$ for the sailing time between successive stations in stratum i.

Variance expressions for the historical predictor, μ_{rh}

The following approximate expression for the variance of $\mu_{ri}^* = \bar{Y}_i / \Sigma \bar{Y}_1$ is applied:

$$\text{var}(\mu_{ri}^*) \approx \mu_{ri}^2 \cdot \text{var}\left(\frac{\bar{Y}_i}{\mu_i} - \frac{\Sigma \bar{Y}_I}{\Sigma \mu_I} \right), \quad i = 1,...,m \tag{12}$$

where the sums are over strata. The expression above is based on a Taylor-expansion of μ_{ri}^* to second order with respect to the two ratios within the brackets (Cochran, 1977). The accuracy of the approximation is strongly dependent on the CV of the latter ratio, which involves all observations from a survey. The approximation is therefore assumed to be reasonable.

Based on the relation $\sigma = k\mu$, the variance in the equation above is estimated by

$$\text{var}^*(\mu_{ri}^*) = \mu_{ri}^{*2} \cdot \frac{k^{*2}}{N_i} \cdot (1 - 2\mu_{ri}^* + N_i \Sigma(\mu_{rI}^{*2} / N_I)), \quad i = 1,...,m \tag{13}$$

where the sums are over strata. The estimator k^* for k is a weighted least square estimator based on the natural logarithmic values of the empirical stratum means, $\bar{y}_{ij}$, and standard deviations, s_{ij}:

$$k^* = \exp\left(\sum_{j=1}^{n_{yr}} \sum_{i=1}^{m} \left(\frac{N_{ij}}{\Sigma_{ij} N_{ij}} \right) \log(s_{ij} / \bar{y}_{ij}) \right) \tag{14}$$

where i denotes stratum and j denotes survey. The reason for the log-transformation is the apparently more constant variance of $\log(s_{ij})$ compared to the increasing variance of s_{ij} with respect to $\bar{y}_{ij}$ (Fig. 2).

Hypothesis test of constant μ_{rh} over years

The question of utilising historical data to predict μ_r in a future survey is now faced. First it is assumed as a working hypothesis that μ_r has not changed during the period for which there are data:

$$H_0 : \mu_{rj} = \mu_{rh}, \quad j = 1, ..., n_{yr} \tag{15}$$

where μ_{rh} is constant. As a test statistic, apply

$$Q = \sum_{i=1}^{m} \frac{(\mu_{rij}^* - \mu_{rih})^2}{\text{var}^*(\mu_{rih}^*)}, \quad j = 1, ..., n_{yr} \tag{16}$$

that is assumed to be approximately chi-square distributed chi2(m) with m degrees of freedom. In order to test H_0 an "*ad hoc*" procedure is first used in order to estimate μ_{rh}, which weights each stratum mean proportional to the total number, n_j, of observations from survey j:

$$\mu_{rih}^* = \sum_j \frac{n_j \cdot \mu_{rij}^*}{\sum_j n_j}, \quad i = 1, ..., m \tag{17}$$

where the sums are over surveys. To test for possible differences between years, apply the test statistic Q_j in eq. (16) replacing μ_{rih} with μ_{rih}^* and perform one test per survey. H_0 is rejected if Q_j is larger than e.g. the upper $.05/n_{yr}$ fractile in the chi2(m)-distribution for any of the years. If H_0 is rejected a more detailed analysis is required providing another predictor for μ_r, e.g. μ_r^* from the last survey(s).

Mathematical details for the subjective prediction distribution

As previously noted in the main text, it is assumed that the widths, U–L, of the subjective prediction intervals are proportional to the mid value μ_s:

$$U - L = \varepsilon(L + U)/2 = \varepsilon\mu_s \tag{18}$$

where ε denotes the proportionality constant and is easily estimated from the subjective prediction data. Under the assumptions that $\mu \sim N(\mu_s, \sigma_s)$ and $\alpha = \beta$, the following expression is obtained for $\sigma_s^2 = \text{var}(\mu)$:

$$\sigma_s^2 = \text{var}(\mu) = \left(\frac{\varepsilon(U+L)/2}{2\Phi^{-1}(1-\beta)} \right)^2 \tag{19}$$

An appropriate estimator σ_s^{2*} for σ_s^2 is found by replacing ε and β by estimators based on experimental data. Analogous to the variance estimator for the historical predictors in eq. (13) we now get the following estimators for $\sigma_{ris}^{2*} = \text{var}^*(\mu_{ris})$:

$$\sigma_{ris}^{2*} = \text{var}^*(\mu_{ris}) = \mu_{ris}^2 \cdot \left(\frac{\sigma_{is}^{2*}}{\mu_{is}^2} - 2 \cdot \frac{\sigma_{is}^{2*}}{\mu_{is} \sum_l \mu_{ls}} + \frac{\sum_l \sigma_{ls}^{2*}}{\left(\sum_l \mu_{ls}\right)^2} \right), \quad i = 1, ..., m \tag{20}$$

where the sums are over strata. The expression above is reasonable for sufficiently narrow intervals relative to the mid values, and when the subjective predictions for one stratum are independent of the subjective predictions for any other stratum.

———————————